The Emerging Functionality of Endogenous Lectins: A Primer to the Concept and a Case Study on Galectins including Medical Implications

Hans-Joachim Gabius, PhD; Albert M. Wu[1], PhD


Prof. Hans-Joachim Gabius


Prof. Albert M. Wu

Biochemistry textbooks commonly make it appear that it is a foregone conclusion that the hardware of biological information storage and transfer is confined to nucleotides and amino acids, the letters of the genetic code. However, the remarkable talents of a third class of biomolecules are often overlooked. For example, one of them far surpasses the building blocks of nucleic acids and proteins in terms of theoretical coding capacity by oligomer formation. Although often exclusively assigned to duties in energy metabolism, carbohydrates as part of cellular glycoconjugates (glycoproteins, proteoglycans, glycolipids) have, in fact, other important tasks. Currently, they are increasingly gaining recognition as an operative high-density information coding system. An elaborate enzymatic machinery enables cells to be versatile enough to produce a glycan profile (*glycome*) that is as characteristic as a fingerprint. Moreover, swift modifications during dynamic processes, such as differentiation or malignant transformation, are readily possible. The translation of the information presented in oligosaccharide determinants to biological responses is carried out by lectins. Recognition of foreign glycosignatures in innate immunity, regulation of cell-cell/matrix interactions, cell migration or growth, and intra- and intercellular glycan routing etc represent physiologically far-reaching lectin-carbohydrate functionality. The classification of endogenous lectins is guided by sequence alignments and conservation of distinct structural traits. For example, a jelly-roll-like folding pattern and maintenance of key residue positioning involved in stacking and C-H/π-interactions as well as directional hydrogen bonds to the β-galactoside ligands are common denominators among galectins. Biochemical and biophysical studies are beginning to unravel the intricacies of the selection of a limited set of endogenous ligands, such as certain integrins or ganglioside GM_1, and combined with biological cell experiments, its relevance for cell sociology, e.g. in growth regulation and tumor cell invasion or activated T cell apoptosis. Histopathological monitoring accompanies the biological cell investigations, linking expression of certain family members to tumor progression or suppression. Further insights into the functional consequences of the sugar code's translation are thus expected to have notable repercussions for diagnostic and therapeutic procedures. *(Chang Gung Med J 2006;29:37-62)*

Key words: agglutinin, apoptosis, drug design, galectin, glycoconjugate, lectin, metastasis, sugar code.

Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Germany; [1]Glyco-Immunochemistry Research Laboratory, Institute of Molecular and Cellular Biology, Chang Gung University, Taoyuan.
Received: Jul. 29, 2005; Accepted: Sep. 16, 2005
International Correspondence to: Prof. Hans-Joachim Gabius, Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Veterinarstrasse 13, D-80539 Munich, Germany. Tel.: 49-89-2180-2290; Fax: 49-89-2180-2508; E-mail: gabius@lectins.de or gabius@tiph.vetmed.uni-muenchen.de
Local Correspondence to: Prof. Albert M. Wu, Glyco-Immunochemistry Research Laboratory, Institute of Molecular and Cellular Biology, Chang Gung University, Gueishan Shiang, Taoyuan, Taiwan 333, R.O.C. Tel.: 886-3-211-8966 (Lab); Fax: 886-3-211-8456 (Lab); 886-3-211-8700 (Col); E-mail: amwu@mail.cgu.edu.tw

Introduction

The central dogma of molecular biology has limited our view of the downstream flow of genetic information to proteins. As a result, nucleic acids and proteins have gained a "glamorous" status which "overshadowed" other classes of biomolecules for decades.[1] Speaking of biological information thus automatically invoked the concept of the genetic code. That this situation has become subject to a fundamental change is due to two major lines of accrued evidence. First, a systematic calculation of the upper limits of coding capacity for biomolecules by means of oligomer formation has revealed that carbohydrates are second to no other compound class.[2] While only 64 permutations (code words) are possible for a codon in protein biosynthesis, a total of 38,016 trisaccharides are already theoretically possible when starting with only three different monosaccharides as letters.[2] Ironically, the structural manifestation of this potential to enable high-density coding is exactly at the heart of the problem why recognition of the concept of the sugar code "has apparently lagged so far behind the other fields", i.e. genomics and proteomics.[3] The plethora of oligosaccharides which establishes the glycome of a cell (the equivalent of the genome and proteome) constituted a technical challenge of a new dimension. It required development of efficient strategies to integrate and refine different separation and analytical procedures such as high-resolution anion-exchange chromatography, mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy.[4,5] After having, in principle, mastered this task and successfully established the analytical protocols and guidelines that govern glycan assembly and processing, it has only recently become possible to confidently introduce this topic into courses of basic biochemistry to let the message of the sugar code catch momentum.[6-9]

Second, carbohydrate diversity has a biological meaning. The clinical delineation of a causal relationship between defects in glycosylation and disease, such as in lysosomal storage diseases or leukocyte adhesion deficiency syndrome type II (LAD II), as well as the existence of sugar compound examples as pharmaceuticals attest to the fact that the inferred coding potential is actually utilized and is of medical relevance (for a representative compilation, please see Table 1).[8] It seems reasonable to say that further developments in drug design based on the sugar code can be anticipated by progress in basic research. At this stage, the reader may wonder why coding along the germline is apparently so restricted compared to this property in the glycocalyx.

Looking at nucleic acids, the prime aim to opti-

Table 1. Examples of Sugar Compounds as Pharmaceuticals

Compound	Target	Disease
Acarbose	α-glucosidases (amylases)	Diabetes mellitus
Heparin/heparinoids	Antithrombin III	Thrombosis
Heparin pentasaccharide (Fondaparinux)	Antithrombin III (factor Xa)	Thrombosis
Derivatives or mimetics of 2-deoxy-2,3-dehydro-N-acetylneuraminic acid	Neuraminidase	Viral infection
N-butyldeoxynojirimycin	α-glucosidases (N-glycan processing)	Viral infection
Derivatives or mimetics of milk oligosaccharides	Adhesins and toxins (lectins)	Bacterial infection
GlcN-(2-O-hexadecyl) phosphatidylinositol	GPI-mannosyltransferase I	Protozoan infection (e. g. African sleeping sickness)
Derivatives or mimetics of ialylated/sulfated $Le^{a/x}$-epitopes	Selectins	Inflammatory reaction
D-Man	Phosphomannose isomerase deficiency	Congenital disorder of glycosylation Ib
L-Fuc	GDP-fucose transport	Congenital disorder of glycosylation IIc (LAD II)
N-butyldeoxygalactonojiri-mycin and properly glyco-sylate β-gluco (galacto) cerebrosidase	Glycosphingolipid sylated and enzymatic degradation	Glycosphingolipid storage disorders

From (53[SME1])

mize copying fidelity in heredity accounts for limiting the structure to linearity and the alphabet size to four letters.[10] In contrast, the generation of code words with carbohydrates as letters can take full advantage of all structural levels to achieve chemical diversity, including anomeric variation and branching. Equally noteworthy, each hydroxyl group of a carbohydrate can be engaged in glycosidic bonds, a situation totally different from nucleotides and amino acids. As a consequence, the number of enzymes responsible for generating the apparent glycan diversity must be considerably larger than in DNA/RNA and protein synthesis. Indeed, the elaborate system of glycosyltransferases ensures versatile glycan (code word) generation, although the synthesis of the complete panel of theoretical linkage types is enzymatically not possible for each sugar moiety.[11,12] The overall investment in genetic coding (i.e. genes for glycosyltransferases and supporting proteins, including transporters of nucleotide sugars, such as guanosine diphosphate (GDP)-fucose, a target for mutations causing LAD II) pays off handsomely by rendering synthesis of a wide array of oligomers feasible, especially at branch ends in glycans. However, their availability would be futile without intricate biochemical mechanisms for accurate translation of the high-density information coding of glycan determinants. Towards this end, nature has devised carbohydrate recognition domains (CRDs) organized in the lectin superfamily.[13-15] This term denotes carbohydrate-binding (glyco) proteins. However, those proteins which enzymatically modify the bound carbohydrate ligand (e.g. glycosyltransferases or sulfotransferases), which are immunoglobulins or which physiologically interact with free mono- or disaccharides or their derivatives in transport or chemotaxis, are explicitly excluded.[15] In a nutshell, lectins serve as an essential interface between glycan structure and function. With the major players in information transfer by the sugar code identified, the next section will guide the reader to structural and functional aspects of lectins.

Lectins as effectors in functional glycomics

The documented beginning of research on lectins dates back to 1860 with the observation by Silas Weir Mitchell (1829-1913), a leading expert in neurology and psychiatry at the time and also a successful novelist, that one drop of venom from the rattlesnake (*Crotalus durissus*) and a drop of blood from a pigeon's wounded wing "coagulated firmly within three minutes."[16] More than 100 years after this pioneering observation, Marie Louise Ogilvie and T. Kent Gartner were able to conclusively attribute Ca^{2+}-dependent lectins (later referred to as C-type lectins) in Crotalidae venoms to this activity.[17] Technically, sugar-dependent agglutination was the common assay to detect and monitor lectin activity, and lectins were classified according to their monosaccharide specificity measured by systematic inhibition studies. As summarized in Table 2, the development of affinity chromatography to purify lectins and of protein and gene sequencing to determine their primary structure, as well as the application of biophysical techniques to describe folding patterns, have paved the way for the classification of lectin groups into families. In this respect, sequence alignments have become a powerful tool to identify similarities (or even homologies) in proteins sharing a common sugar target. The extent of identity scoring which is commonly encountered in these computer-supported calculations is exemplarily shown for mammalian proto-type galectins in Figure 1. Such systematic computations in conjunction with biochemical properties, e.g. Ca^{2+}-dependence of sugar binding, led to the definition of five classes of animal lectins (Table 3). Meanwhile, their crystal structures have been solved and support the inherent class separation. The classical route from protein to crystal, its diffraction pattern, and then resolved secondary structure, as nicely shown in the case of a galectin in Figure 2, thus confirmed the validity of this classification system. Accounting for ongoing studies, which are defining new folds for carbohydrate recognition in animals, it is reasonable to conclude that the complexity of code word generation by glycosyltransferases is matched on the level of decoding devices.

The crystallographic analysis of lectins not only aids in the establishment of a knowledge-based classification scheme but also leads to an understanding of the positioning of invariant amino acid residues emerging from sequence alignments. As illustrated for D-galactose, directional hydrogen bonds with distinct side chains and the stacking and CH/π interactions with a Trp ring system govern the ability of galectins (or bacterial toxins, such as cholera toxin; please see below) to distinguish β-galactosides from

Table 2. A Brief Historical Account of Lectinology

Years	Topic	Original author(s)
1860	Observation of blood "coagulation" by rattlesnake venom	S. W. Mitchell
1888	Detection of erythrocyte agglutination by protein fractions from castor beans and other plant seeds	H. Stillmark
1891	Toxic plant agglutinins applied as model antigens	P. Ehrlich
1898	Introduction of the term "haemagglutinin" or phytohaemagglutinin for plant proteins that agglutinate red blood cells	M. Elfstrand
1902	Detection of bacterial agglutinins	R. Kraus
1902	Demonstration that blood "coagulation" by snake venom (later shown to depend on a C-type lectin) observed in 1860 was not due to blood clotting but to cell agglutination	S. Flexner, H. Noguchi
1906	Detection of an agglutinin in bovine serum (later characterized as the C-type lectin conglutinin) acting on activated complement-coated erythrocytes	J. Bordet, F. P. Gay
1907	Detection of non-toxic agglutinins in plants	K. Landsteiner, H. Raubitschek
1913	Use of intact cells for the purification of lectins	R. Kobert
1919	Crystallization of a lectin, concanavalin A	J. B. Sumner
1936	Precipitation of starch, glycogen and mucins by concanavalin A and its interaction with stromata of erythrocytes define carbohydrate as ligand	J. B. Sumner, S. F. Howell
1941	Detection of viral agglutinins	G. K. Hirst
1947/48	Detection of lectins specific for human blood groups	W. C. Boyd, K. O. Renkonen
1952	Carbohydrate nature of blood group determinants proven by lectin-mediated agglutination and its sugar-dependent inhibition	W. M. Watkins, W. T. J. Morgan
1954	Introduction of the term "lectin" for plant agglutinins, primarily for those which are blood group-specific	W. C. Boyd
1960	Detection of the mitogenic potency of lectins toward lymphocytes	P. C. Nowell
1965	Application of affinity chromatography for the isolation of lectins	I. J. Goldstein, B. B. L. Agrawal
1972	Determination of the amino acid sequence and the three-dimensional structure of a lectin, concanavalin A	G. M. Edelman, K. O. Hardman, C. F. Ainsworth et al.
1972-1977	Detection of impaired synthesis of a marker for glycoprotein (lysosomal enzymes) routing as cause for a human disease (mucolipidosis II) and its identification as Man-6-phosphate, the ligand for P-type lectins	E. F. Neufeld and colleagues; W. S. Sly and colleagues
1974	Isolation of a mammalian Gal/GalNAc-specific lectin from liver	G. Ashwell
1978	First conference focusing on lectins and glycoconjugates, termed Interlec	T. C. Bϕg-Hansen
1979	Detection of endogenous ligands for plant lectins	H. Rüdiger
1983	Detection of the insecticidal action of a plant lectin	L. L. Murdock
1984	Isolation of lectins from tumors	H.-J. Gabius; R. Lotan, A. Raz
1985	Immobilized glycoproteins as pan-affinity adsorbents for lectins	H. Rüdiger
1987	Introduction of neoglycoconjugates for localization of tissue lectins for tumor diagnosis	H.-J. Gabius and colleagues
1989	Detection of the fungicidal action of a plant lectin	W. J. Peumans
1992/93	Detection of impaired synthesis of lectin (selectin) ligands by defective fucosylation as cause for leukocyte adhesion deficiency type II	A. Etzioni and colleagues
1995	Structural analysis of a lectin-ligand complex in solution by NMR spectroscopy	J. Jiménez-Barbero and colleagues
1996-1998	Detection of differential conformer selection by plant and animal lectins	H.-J. Gabius and colleagues; L. Poppe and colleagues
2001/02	Advances in lectinology and glycosciences honored by devoting special issues in *Biochim. Biophys. Acta, Biochimie, Biol. Chem., Cells Tissues Organs, Chem. Rev., Curr. Opin. Struct. Biol., J. Agric. Food Chem.*	Liener symposium and *Science* to the topics

From [53], extended and modified[SME2]

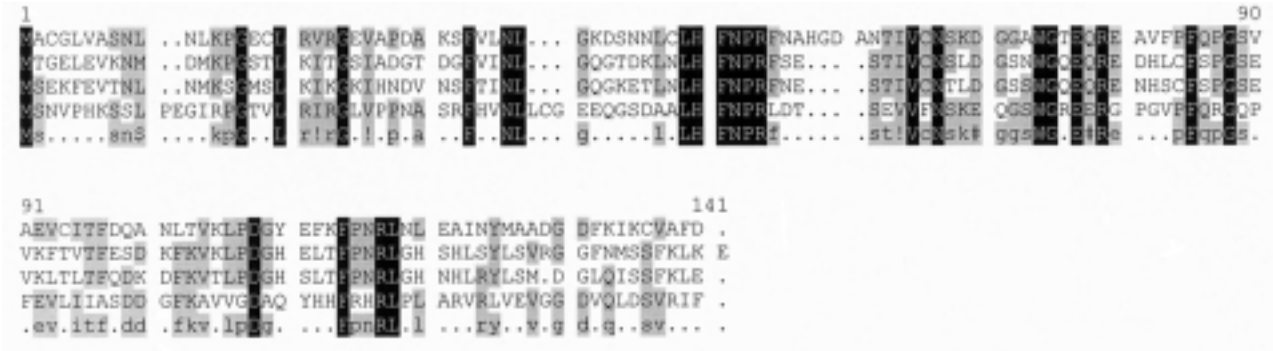


Fig. 1 Intergalectin and interspecies sequence comparison of subgroups of mammalian prototype galectins, including galectins -1, -2 and -7. Amino acid sequences of human galectins -1, -2 and -7, as well as rat galectin-2 were aligned using the Multalin program (http://prodes.toulouse.inra.fr/multalin/multalin.html; version 5.4.1). Identical residues found in all four galectins are indicated as white letters on a black background, whereas residues that are identical or similar between at least two of the sequences are in black letters on a grey background. A consensus sequence calculated from the four galectin sequences is added to the alignment; consensus symbols used: !: I or V; $: L or M; %: F or Y; #: N, D, Q or E.

Table 3. Main Families of Animal Lectins

Family	Structural motif	Carbohydrate ligand	Modular arrangement
C-type	Conserved CRD	Variable (mannose, galactose, fucose, heparin tetrasaccharide)	yes
I-type	Immunoglobulin-like CRD	Variable ($Man_6GlcNAc_2$, HNK-1 epitope, hyaluronic acid, α2,3/α2,6-sialyllactose)	yes
Galectins (formerly S-type)	Conserved CRD	Galβ1,3(4)GlcNAc core structures with species- and galectin type-dependent differences in affinity for extensions to blood group A, B or H epitopes; internal stretches of poly (N-acetyllactosamine) chains	variable
Pentraxins	Pentameric subunit arrangement	4-6-cyclic acetal of β-galactose, galactose, sulfated and phosphorylated monosaccharides	yes
P-type	Conserved CRD	Mannose-6-phosphate-containing glycoproteins	yes

Abbreviations: CRD: carbohydrae recognition domain; from [9] with modification

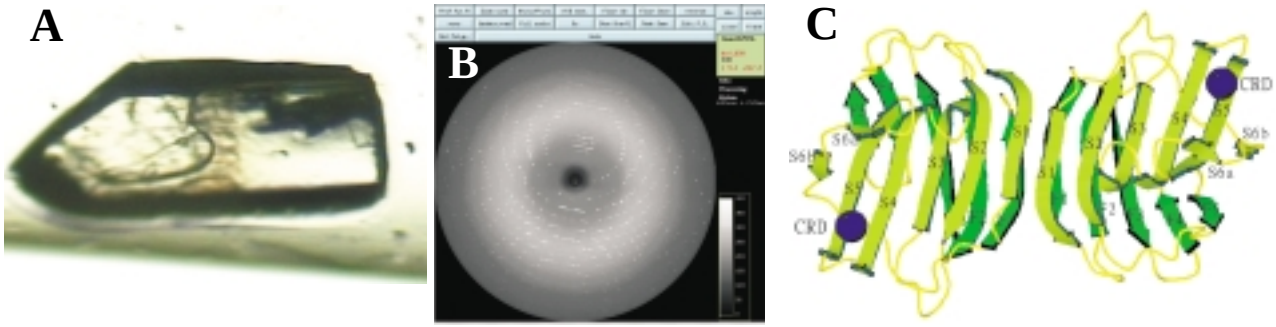


Fig. 2 Orthorhombic crystal ($C222_1$) of the developmentally regulated homodimeric galectin from chicken liver (CG-16) grown in 2 M ammonium sulfate, 5% (v/v) isopropanol and 1% β-mercaptoethanol, pH 5.6, at an estimated final protein concentration of 10 mg/ml. The crystal size is 0.4 x 0.6 mm^3 (A). A 1° oscillation photograph of CG-16 collected on a Rigaku RU200 rotating anode generator operating at 5.4 kW. The diffraction pattern in this film extends to 2.1 Å resolution (B). A ribbon diagram of CG-16, prepared with MOLSCRIPT. The β-strands in the five-stranded (F1-F5) and six-stranded (S1-S6a/S6b) β-sheets are denoted by the letter-number code. The two carbohydrate-binding sites or CRD located at the opposite ends of the homodimer are indicated by spheres (C). This figure is adapted from reference [19].

natural epimers, such as mannose or glucose (please note involvement of the axial 4'-hydroxyl group in the contacts shown in Figure 3).[14,18,19] Beyond these insights, lectin structure investigation has answered another pertinent question regarding the overall binding energy: how does the inherently low affinity of a single protein-carbohydrate contact become strongly bound when a lectin meets a binding partner? Drawing an analogy, antibodies exploit the option of spatial clustering to attain this aim. As noted above, immunoglobulins are definitely excluded from the lectin definition. Nonetheless, similar means to reach the same objective could be envisaged. Indeed, structural and cellular biology studies have taught the amazing lesson that lectin carbohydrate recognition domains (CRDs) often tend to be spatially associated, either by non-covalent or covalent bonds. The spatial topology is thus a factor that matters to a considerable extent, establishing a recurring theme in functional glycomics. Undoubtedly, generation of multiple contacts by clustering is a general key to improve binding activity. To help the reader appreciate this strategy, Figure 4 presents an overview of CRD presentation. This figure is meant to convey how the different modes of presentation of CRDs are tailored by evolution for distinct functions. Lectins used by the innate immune system, such as the serum mannan-binding lectin (and also other collections such as surfactant proteins-A and -D or the ficolins) and the tandem-repeat C-type (macrophage) mannose receptor are model examples. The widely spaced CRDs present in these lectins act as ideal sensors for the non-mammalian surface glycosignatures of infectious bacteria or yeasts, thus explaining their alternate, descriptive name of "pattern recognition receptors".[20,21] Further exemplary cases have been compiled in a recent collection of reviews on animal lectins.[22] As indicated, the presence of a CRD and the functional cooperation of CRDs via spatial positioning (or of a CRD in conjunction with other modules in mosaic-like proteins) are instrumental for lectin functionality. An overview of the current spectrum of documented lectin functions is given in Table 4, which underscores the validity of this section's heading - the idea of lectins as effectors in functional glycomics.

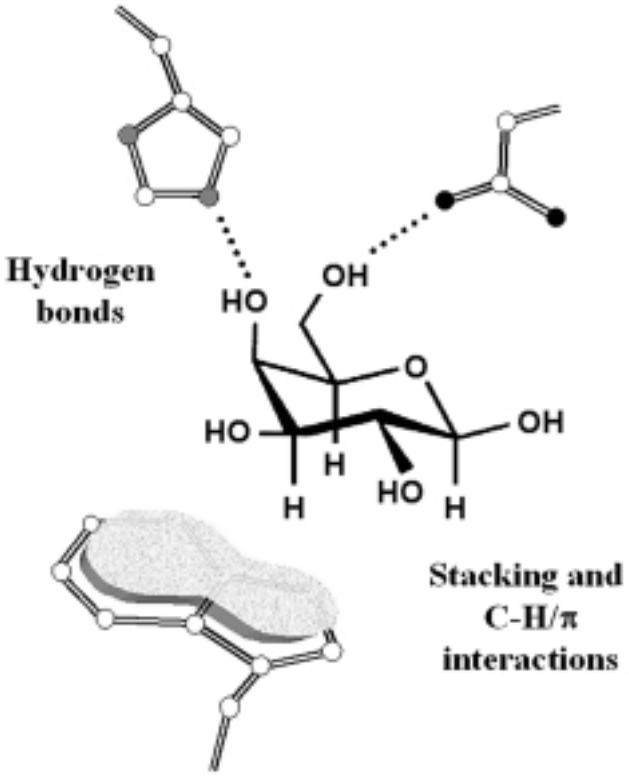


Fig. 3 Main enthalpic features to facilitate binding of a sugar to a protein: hydrogen bonds using lone electron pairs of sugar oxygen atoms as acceptors and the hydrogen atoms of hydroxyl groups as donors, as well as C-H/π-electron interactions between patches of positively polarized character (in this case the B-face of *D*-Gal) and the delocalized π-electron cloud of a suitably positioned aromatic residue (here Trp).

Now that lectin structure has been related to function and guidelines have been defined to allow translation of sequence information of newly discovered lectins into structural motifs, the following question emerges: are there rules of glycan presentation in natural glycoconjugates to regulate lectin affinity? Intuitively, it would mean missing manifold opportunities if this parameter were not to influence affinity. Thus, the aspect complementary to CRD presentation, i.e., the influence of distinct topological presentation of the lectin-reactive glycan(s) by branching or clustering, has become a promising research topic and its importance is in fact being unraveled. Systematic analysis of lectin binding to well-characterized glycoproteins with complex glycosylation profiles including multi-antennary or branched glycan chains proves indispensable in this respect.[23,24] As will be discussed later in this review, we are now becoming aware of different levels of

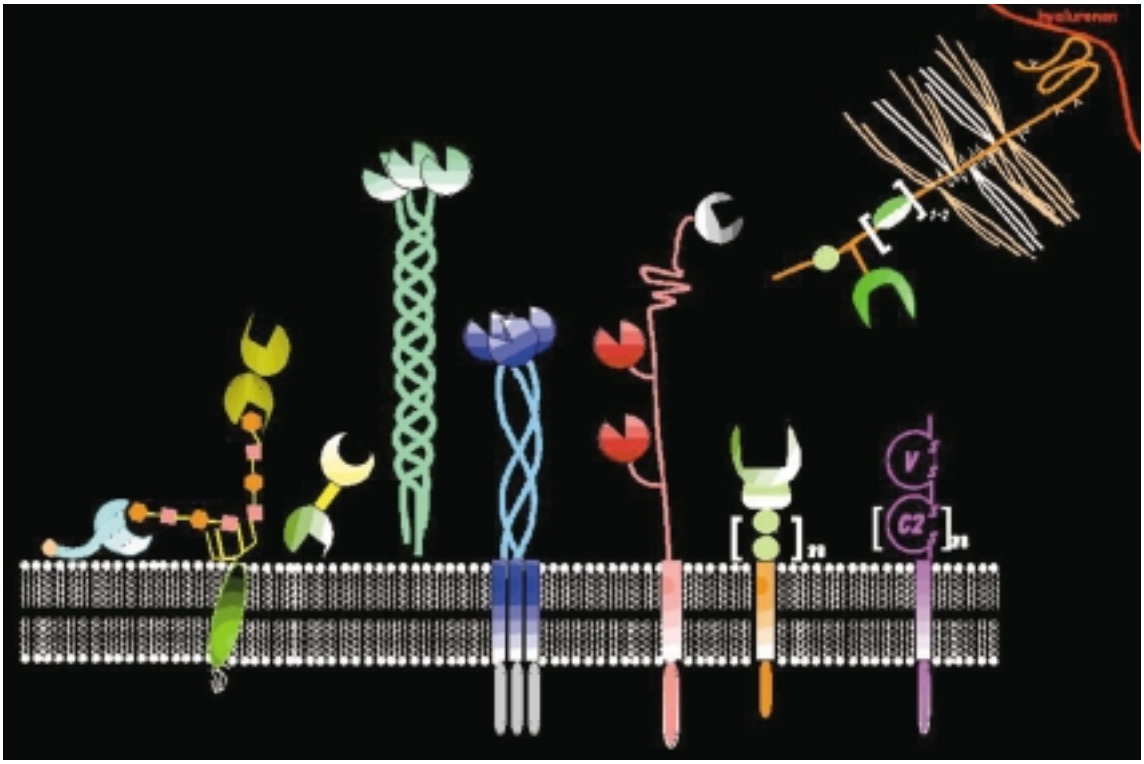


Fig. 4 Illustration of the strategies of how carbohydrate recognition domains (CRDs) in animal lectins are positioned to reach optimal affinity, ligand selection (e.g. to differentiate self from non-self glycan profiles in innate immunity) and accessibility (modified from reference [12]). From left to right, the CRD display in the three subtypes within the galectin family (chimeric, prototype and tandem-repeat-type arrangements), the presentation of CRDs in serum and surfactant collectins or ficolins connected to their collagenous stalks, and the noncovalent association of binding sites in transmembrane C-type lectins by α-helical coiled-coil stalks (e.g. asialoglycoprotein and Kupffer cell receptors, CD23, DC-SIGN or DC-SIGNR). Presented next, the tandem-repeat display in mannose-specific macrophage receptor (also found on dendritic cells, hepatic endothelial cells, kidney mesangial cells, retinal pigment epithelial cells and tracheal smooth muscle cells) and related C-type subfamily lectins (e.g. DEC-205 or Endo-180), as well as in the cation-independent P-type lectin. The occurrence of lectin activity for GalNAc-4-SO_4-bearing pituitary glycoprotein hormones in the cysteine-rich domain, a member of the β-trefoil family of protein modules, in the N-terminal section of the mannose receptor, which is linked via a fibronectin-type-II-repeat-containing module to the tandem-repeat section is also included in the schematic drawing for these two classes of lectins/lectin-like proteins with more than one CRD per protein chain. Further to the right is shown the association of a distal CRD in selectins (attached to an epidermal-growth-factor-(EGF)-like domain and two-to-nine complement-binding consensus repeats) or in the siglec subfamily of I-type lectins using 1-16 C2-set immunoglobulin-like units as spacer equivalents to let the CRD reach out to contact ligands and to modulate capacity to serve in *cis*- or *trans*-interactions on the cell surface. In the matrix, the modular proteoglycans (hyalectans/lecticans: aggrecan, brevican, neurocan and versican) interact with (and also link) hyaluronan via the N-terminal loop assigned to the immunoglobulin superfamily, with receptors binding to the glycosaminoglycan chains in the central region and with carbohydrates or proteins (fibulins-1 and -2 and tenascin-R) via the C-type lectin-like domain flanked by EGF-like and complement-binding consensus repeat modules. For further information on individual lectin groups, please refer to reference [22].

specificity in lectin-carbohydrate recognition.

At this stage, we can reliably conclude that carbohydrates are "ideal for generating compact units with explicit informational properties."[25] Code word generation is achieved by a complex array of glycosyltransferases. The regulation of the expression and activity of these enzymes facilitates dynamic and reversible shifts in the glycome. Glycan diversity is matched on the level of receptor proteins (lectins). Strategic positioning of CRDs in lectins leads to high-affinity binding and a remarkably wide range of lectin functions. The classification of lectins is based

Table 4. Functions of Animal Lectins

Activity	Example of lectin
Ligand-selective molecular chaperones in endoplasmic reticulum	Calnexin, calreticulin
Intracellular routing of glycoproteins and vesicles	ERGIC-53 and VIP-36 (probably also ERGL and VIPL), P-type lectins, comitin
Intracellular transport and extracellular assembly	Non-integrin 67 kDa elastin/laminin-binding protein
Inducer of membrane superimposition and zippering (formation of Birbeck granules)	Langerin (CD207)
Cell-type-specific endocytosis	Hepatic and macrophage asialoglycoprotein receptors, dendritic cell and macrophage C-type lectins (mannose receptor family members (tandem-repeat type) and single CRD* lectins such as langerin/CD207), cysteine-rich domain of the dimeric form of mannose receptor for GalNAc-4-SO_4-bearing glycoprotein hormones in hepatic endothelial cells, P-type lectins
Recognition of foreign glycans(1,3-glucans, LPS)	CR3 (CD11b/CD18), dectin-1, *Limulus* coagulation factors C and G, earthworm CCF
Recognition of foreign or aberrant glycosignatures on cells (incl. endocytosis or initiation of opsonization or complement activation)	Collectins, L-ficolin, C-type macrophage and dendritic cell receptors,/-defensins, pentraxins (CRP, limulin), tachylectins
Targeting of enzymatic activity in multimodular proteins	Acrosin, laforin, *Limulus* coagulation factor C
Intra- and intermolecular modulation of enzyme activities *in vitro*	Porcine pancreatic amylase, galectin-1/2-6-sialyltransferase
Bridging of molecules	Homodimeric and tandem-repeat-type galectins, cytokines (e.g. IL-2:IL-2R and CD3 of TCR), cerebellar soluble lectin
Induction or suppression of effector release (H_2O_2, cytokines etc.)	Galectins, selectins and other C-type lectins such as CD23, BDCA-2 and dectin-1
Cell growth control and induction of apoptosis/anoikis	Galectins, C-type lectins, amphoterin-like protein, hyaluronic acid-binding proteins, cerebellar soluble lectin
Cell migration and routing	Selectins and other C-type lectins, I-type lectins, galectins, hyaluronic acid-binding proteins (RHAMM, CD44, hyalectans/lecticans)
Cell-cell interactions	Selectins and other C-type lectins (e. g. DC-SIGN), galectins, I-type lectins (e.g. siglecs, N-CAM, P_0 or L1)
Cell-matrix interactions	Galectins, heparin- and hyaluronic acid-binding lectins including hyalectans/lecticans, calreticulin
Matrix network assembly	Proteoglycan core proteins (C-type CRD and G1 domain of hyalectans/lecticans), galectins (e. g. galectin-3/hensin), non-integrin 67 kDa elastin/laminin-binding protein

* Carbohydrate recognition domain; from (53), extended and modified.[SME4]

on homology criteria in primary and secondary structures. Examples are given here in Figure 1 and Figure 2 for galectins (lectins showing Ca^{2+}-independent specificity to β-galactosides and the folding pattern shown in Figure 2c). Since various cell adhesion molecules, such as integrins, point to the potential of intrafamily diversity as functional fine-tuning, we next address this question on diversity using galectins as a representative example of a lectin family.

Galectins: structural principles and intrafamily diversity

To serve in information transfer, an epitope should be spatially accessible and subject to an array of substitutions which modulate or even switch off

its ligand activity. Being positioned at the ends of glycan branches, β-galactosides readily fulfill this topological requirement. Moreover, a total of 13 different β-galactosyltransferases are responsible for synthesis of β1,3(4)-galactosides, an indication of the intricate fine-tuning available on the level of the acceptor structure.[26] Adding a variety of potential substitutions introduced by fucosyl-α-galactosyl-, and N-acetylgalactosaminyltransferases and the switch-off signal α2,6-sialylation, β-galactosides also easily satisfy the second requirement for structural diversity defined above.[12] Taken together, it becomes clear that β-galactosides ideally embody the prophetic statement that "glycosyl residues" can "impart a discrete recognitional role to the protein".[25] By extending from a cell or protein surface, these epitopes can act like sensors (or tentacles) in sugar-based communication with galectins. Thus, when looking at Table 4, it comes as no surprise to find that galectins are frequently associated with signaling and adhesion activities. The aforementioned exquisite discrimination against other natural hexoses excludes errors in decoding. On the molecular level, the axial 4'-hydroxyl group of *D*-galactose, as illustrated in Figure 3, is a major contact site for hydrogen bonding, explaining the selection against the naturally occurring epimers glucose and mannose.

As likewise emphasized above in the description of the first observation of activity for an animal lectin which turned out to be a C-type lectin, galectins were also initially detected by the hemagglutination assay. The inhibition of lectin activity by lactose but not unrelated sugars, i.e. agglutination of trypsin-treated rabbit erythrocytes, led to the discovery of the prototype lectin of this family, electrolectin, in extracts of electric organ tissue from *Electrophorus electricus* and various rat, chicken and mouse tissues and cells.[27] Acquisition of the ability to crosslink cells or glycolipids/glycoproteins calls for bivalency with sites spatially separated from each other. The third panel of Figure 2 and the left section of Figure 4 illustrate how the two CRDs are positioned in a homodimeric prototype galectin to generate a cross-linking device.[13] Besides this noncovalent association of modules, two CRDs can be covalently linked by a connecting peptide establishing the tandem-repeat-type organization. The third topological arrangement in the galectin family is found in galectin-3 with its collagenase-sensitive domain, which substantially accounts for oligomer formation (please see Fig. 4, left section). Purification of galectins has become rather easy with custom-made resins.

Affinity chromatography is the method of choice for lectin purification. Owing to the engineering of highly efficient prokaryotic expression vectors, bacterial extracts have replaced tissues as the source for lectins. Since galectins are not glycosylated, it is not necessary to work with mammalian cultures. Pure products are obtained after affinity chromatography as ascertained by two-dimensional gel electrophoresis (Fig. 5). The highly sensitive technique of nano-electrospray ionization mass spectrometry verifies the absence of any substitution by bacterial enzymatic reactions that could confound application of the lectins (Fig. 6). The latter method has even been adapted to determine the quaternary structure of the galectin and is able to pick up dimers without signs of dissociation by sample processing (Fig. 6). With this information on structural aspects and extensive sequence alignments, it was possible to assign each of the 14 currently known mammalian galectins to the three distinct subgroups (Table 5). Although useful in a practical sense, this classification is not sufficiently detailed to mirror evolutionary traits. In order to depict such relationships between

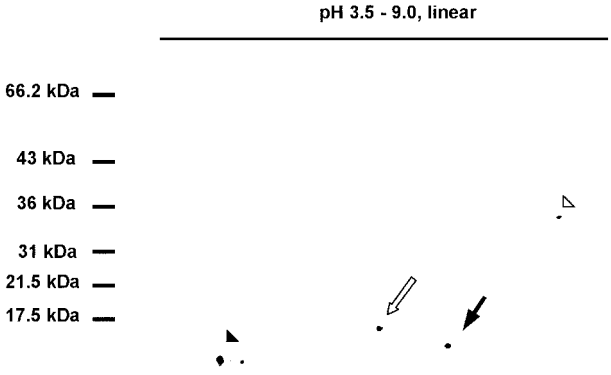


Fig. 5 Two dimensional gel electrophoretic analysis after silver staining showing purity of the three tested galectins: bovine galectin-1 (black arrowhead), rat galectin-5 (open arrow) and mouse galectin-3 (open arrowhead). Human galectin-7 (black arrow) was added as a further control to underscore the resolving power of the method. The occurrence of isoelectric variants was detected for bovine galectin-1, from reference [68].

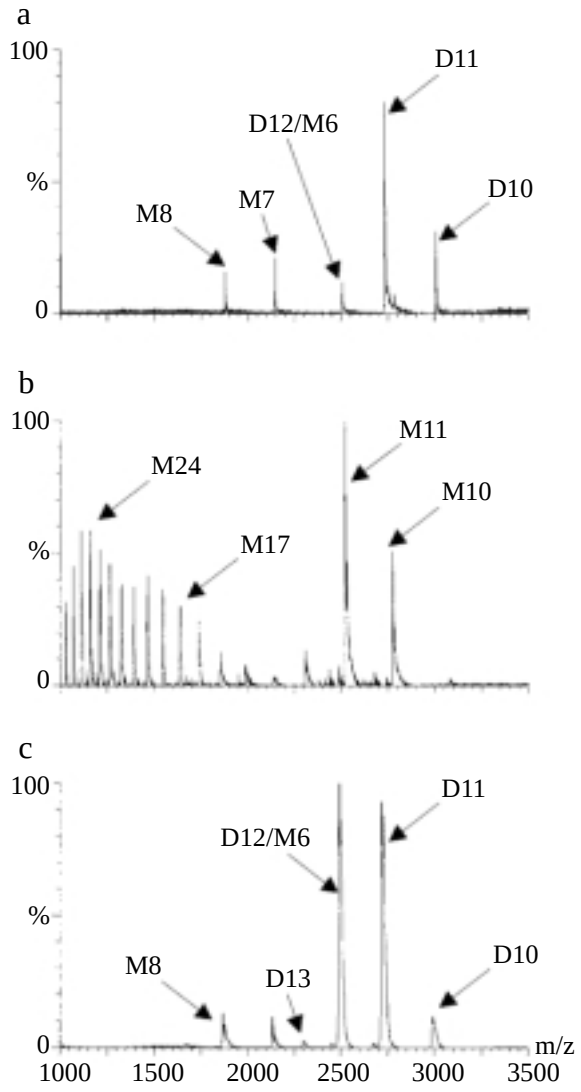


Fig. 6 Nano-ESI mass spectra of human galectin-1 (a), murine recombinant galectin-3 (b), and human recombinant galectin-7 (c) under pseudophysiological conditions to maintain noncovalent oligomer formation. Peak annotations designate the number of positive charges per monomer (M) and dimer (D), from reference [36].

galectins, a genealogical tree was generated, as presented in Figure 7. Of note is the early separation of the prototype galectins-1 and -2 vs -7, the close similarity of the prototype rat galectin-5 to the C-terminal domain of the tandem-repeat-type galectin-9, and the difference between N- and C-terminal CRDs of any tandem-repeat-type galectin. The latter point is taken into account in Figure 4 (left section). Here, the two CRDs of tandem-repeat-type galectins are shown to differ from those in the homodimeric prototype family. It is reasonable to assume that the occurrence of this subgroup can be attributed to the merging of two monomeric modules. This process enables the design of a protein that crosslinks two different ligand types, a special task for which homodimeric prototype galectins are not suited. The position of galectin-3 in the genealogical tree reflects its status as the sole chimera-type family member. At this stage, we can conclude that galectins represent a complex family of endogenous lectins with different degrees of sequence similarity and two principal types of structural organization, that of crosslinking modules with identical or different CRDs (after dissociation or proteolytic cleavage into monomers) or of chimeric proteins with a collagenase-sensitive domain for reversible intermolecular self-aggregation.[13,28] A common feature of galectins is their secretion by a nonclassical pathway. Thus, intracellular association with glycoproteins originating from the Golgi apparatus is precluded.

The information presented in Table 5 not only classifies each mammalian galectin into its appropriate subgroup but also documents progress towards answering the question of tissue-specific regulation by presenting details of tissue expression profiles based on the current literature. The spectrum of individual expression patterns encompasses family members with abundant and frequent tissue presence, such as galectin-1, as well as those closely related members with very restricted tissue presence, such as galectin-2, which is found only in the gastrointestinal tract (please see Figure 1 for a sequence alignment of the prototype galectins-1, -2 and -7). A promising approach to find clues to explain these pronounced differences is to scrutinize the proximal promoter sequences of individual galectin genes. The result of an example of this work is shown in Figure 8 for galectins-1 and -2, and is further explained in the legend to this Figure. Of general medical relevance in this context is the fact that a cell not only expresses a single galectin species but can harbor a characteristic profile of galectin expression (which may often have not been fully explored). To determine this profile, we have introduced fingerprinting by reverse-transcriptase polymerase chain reaction (RT-PCR) or by immunohistochemistry with non-crossreactive galectin-type-specific antibodies.[29-32] Figure 9

Table 5. Members of the Galectin Family of Mammalian Lectins

Name	Occurrence	Structural features
Galectin-1 (galaptin, L-14)	Many cell types	Homodimer; one CRD per subunit (14-15 kDa): prototype
Galectin-2	Gastrointestinal tract; clone from human hepatoma	Homodimer; one CRD per subunit (43% sequence identity to galectin-1; 14 kDa): prototype
Galectin-3 (CBP35, Mac-2 antigen, IgE-binding protein, L-29, L-34)	Many cell types	Monomer with one CRD (oligomer formation in solution and on surfaces); Pro-, Tyr- and Gly-rich repeats in N-terminal section (27-36 kDa): chimera type
Galectin-4	Colon, small intestine, stomach, oral epithelium, esophagus, lung, testis, breast, liver and placenta by RT-PCR	Monomer with two partially homologous but distinct CRDs connected by a link peptide (36 kDa); proteolysis generates truncated prototype-like products: tandem-repeat type
Galectin-5	Reticulocytes, erythrocytes (rat)	Monomer with one CRD (17 kDa): prototype
Galectin-6	Small intestine, colon	Tandem-repeat arrangement of two CRDs (33 kDa)
Galectin-7	Keratinocytes, stratified epithelia, carcinoma cells	Homodimer; one CRD per subunit (15kDa): prototype
Galectin-8	Several tissues; frequent presence in tumor cell lines (link peptide extension possible)	Homologous to galectins-4 and -6 (tandem-repeat arrangement of two CRDs with unique link peptide; 34 kDa)
Galectin-9	Small intestine, liver, lung, kidney, thymus (rat/mouse; small intestinal isoform with 31/32 amino acid extension of link peptide); lymphatic tissue, B cells, T cells and macrophages, pancreas, colon carcinoma cells (human)	Homologous to galectins-4, -6 and -8 (tandem-repeat arrangement of two CRDs with unique link peptide; 36 kDa)
Charcot-Leyden crystal protein (galectin-10)	Major autocrystallizing constituent of eosinophils and basophils	One CRD-like structure with specificity to D-Man (16.5 kDa)
Galectin-11 (ovgal-11)	Sheep gastrointestinal tract induced upon nematode infection	One CRD resembling prototype galectins (14 kDa)
Galectin-12	Several tissues (upregulation in cells synchronized at the G1 phase or G1/S boundary of the cell cycle), adipocytes	Homologous to galectins-4, -6, -8 and -9 (tandem-repeat arrangement of two CRDs with unique link peptide; 35.3 kDa)
Galectin-13	Identical to placental protein 13 (pp13); also expressed in spleen, kidney, bladder and in tumor cells	Homodimer; one CRD per subunit (16.1 kDa); close similarity to galectin-7 and the Charcot-Leyden crystal protein
Galectin-14	Ovine eosinophils, secreted into bronchoalveolar lavage fluid	One CRD resembling prototype galectins (18.2 kDa)

From (69), extended and modified; [SME5] Please note that the presence of human galectins has not been confirmed in all cases (e.g. rat galectin-5)

illustrates the coexpression of several galectins in five different human tumor cell lines (breast, colon, small-cell lung carcinoma, glioblastoma and acute monocytic leukemia) and the normal colonic line hColon. The presence of two bands for the tandem-repeat-type galectin-8 (please also see Table 5) is due to formation of an isoform by an extension of the linker peptide.[31] Intrafamily diversity thus extends beyond the mere presence of several family members. Explicitly, the variability of the length in the

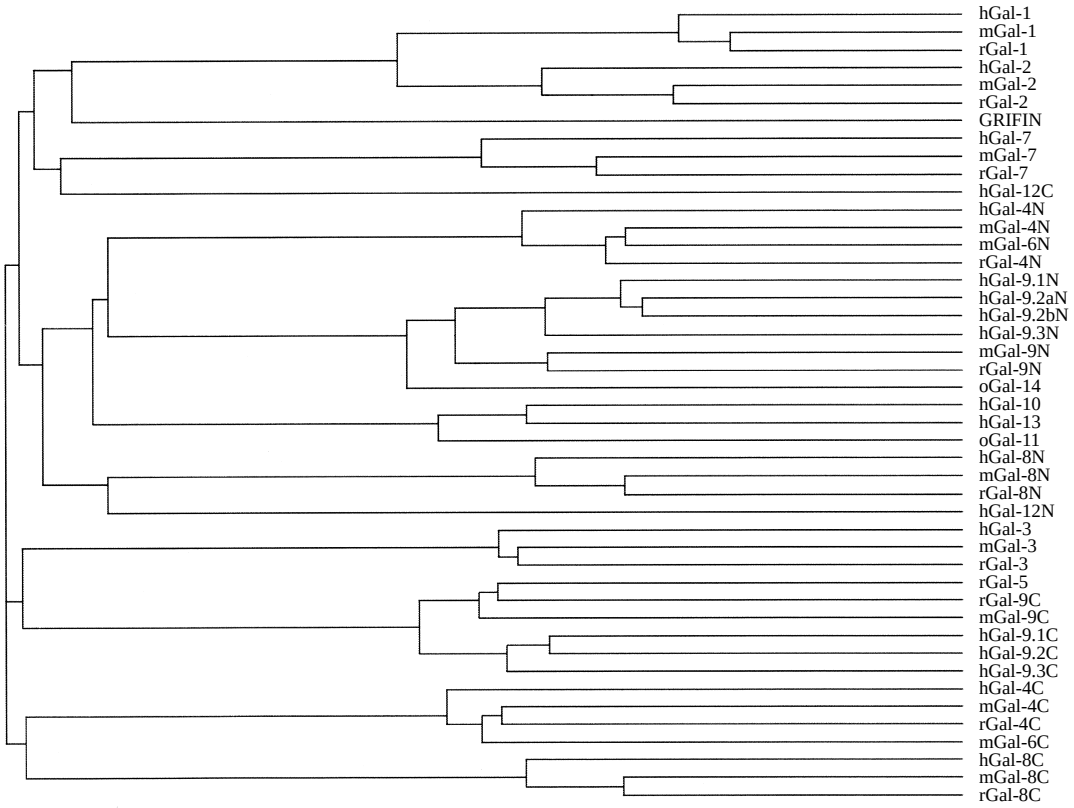


Fig. 7 Genealogical tree of mammalian galectins.

linker peptides, also detected for galectin-9 with an in-frame insertion of 96 base pairs into the coding region of the linker,[31] and alternative splicing (also known to be operative in C-type lectins, such as dectin-1, CD23, DC-SIGN and DCL-1/CD205 (DEC205)) are additional mechanisms that can increase intrafamily diversity.[13,28]

The current status of immunohistochemical fingerprinting not only confirms the RT-PCR data but also provides new prognostic information for tumor diagnosis. In a case study on colon cancer, a low index of expression of galectins-1 and -4 (but not -7) was a favorable prognostic marker for patients with Dukes' A and B tumors.[32] In parallel, studies in animal models for another tumor type provided evidence for the functional basis of galectin-1 in tumor progression. This galectin was found to be negatively correlated with prognosis in patients with glioblastoma and strongly stimulated cell motility and tissue invasion *in vitro* and in xenografts.[33] Before concluding that galectin-1 imparts a generally negative effect on the tumor or patient, it is important to keep in mind that other results have shown that this galectin-induced effect is cell-type specific. An emerging perspective for a new treatment modality in autoimmunity or aberrant cell growth, for instance, is based on the fact that galectin-1 induces apoptosis or reduces cell growth of activated T cells and neuroblastoma (SK-N-MC) cells, respectively.[34-36] The presence of galectin-1 causes a drastic decrease in neuroblastoma cell proliferation *in vitro*, while galectin-3 blocks this effect allowing tumor cells to maintain their growth pattern (Fig. 10). Competitive inhibition of cell surface binding of galectin-1 by galectin-3 accounts for the neutralizing effect by the chimera-type family member, a case of functional divergence among galectins.[35-37] The galectins' major target in these cells is the pentasaccharide of ganglioside GM_1, a molecular rendezvous which will be further explained in the next section. Evidently, galectin-3 is not able to trigger signaling pathways equivalent to the capacity of the cross-link-

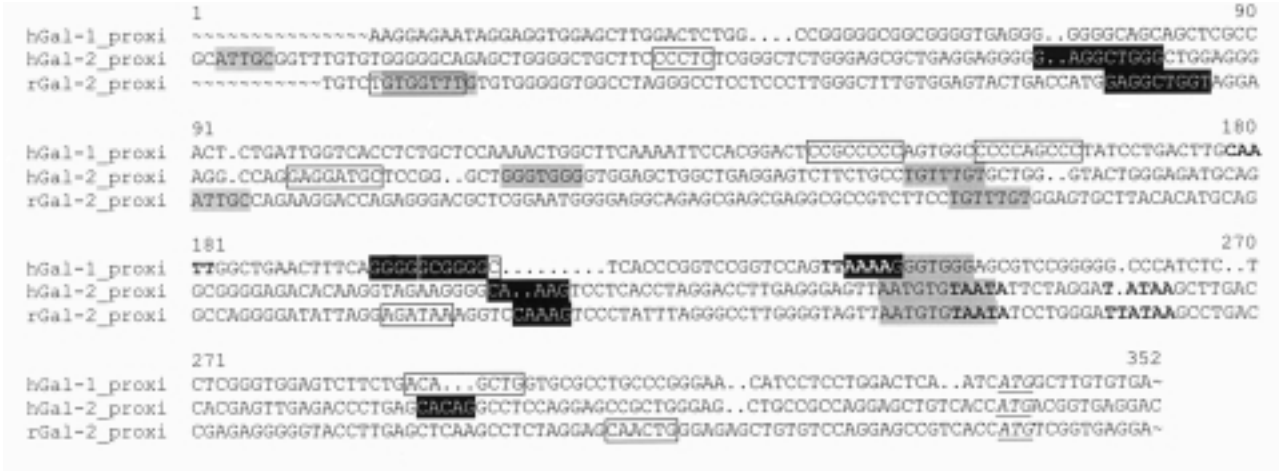


Fig. 8 Intergalectin and interspecies comparison of proximal promoter regions of human galectins-1 (hGal-1; Accession: *Homo sapiens* chomosome 22 contig NT_011520, region: 17380467-17380777) and -2 (hGal-2; Accession: *Homo sapiens* chomosome 22 contig NT_011520, region: complement of 17284952-17285292), as well as rat galectin-2 (rGal-2; Accession: *Rattus norvegicus* chromosome 7 supercontig NW_047780, region: complement of 11420257-11420597). Translation start codons are given in italics and underlined. CAAT and TATA boxes are indicated in bold letters. Putative binding sites for transcription factors are marked according to their occurrence: black boxes indicate sites that occur in all three sequences, grey boxes indicate sites that occur in only two sequences, and framed open boxes indicate sites that are unique to the actual sequence. The sites depicted for transcription factor binding are as follows. *Black boxes*: hGal-1: Sp1-binding site (196-205); TCF-1-binding site (236-240); hGal-2: Sp1-binding site (72-82); TCF-1-binding sites (206-212 and 290-294); rGal-2: Sp1-binding site (78-86); TCF-1-binding site (208-212). *Grey boxes:* hGal-1: PuF-binding site (240-246); hGal-2: C/EBP-binding site (3-7); PuF-binding site (117-123); HNF-5-binding site (151-157); Brn-2-binding site (239-248); rGal-2: C/EBP binding sites (17-24 and 91-95); HNF-5-binding site (155-161); Brn-2-binding site (239-248). *Open boxes:* hGal-1: GCF-binding sites (143-150 and 200-206, note that the latter overlaps almost completely with the Sp1 binding site); AP-2-binding site (157-165); E2A-box (289-298); hGal-2: CTCF-binding site (40-44); c-Ets-1-binding site (99-106); rGal-2: AP-3-binding site (16-23); GATA-1, -2, -3-binding site (197-202); c-Myb-binding site (306-311).

ing galectin-1.[38,39] This example highlights the enormous importance of spatial factors in the realm of protein-carbohydrate recognition. Although already touched upon in the previous section, it is worth discussing this issue in greater detail in the next section of this review.

Galectin-glycoform binding

The binding properties of lectins (galectins) can be analyzed by inhibition of the glycoform-lectin binding assay. The factors involved in the lectin-glycoform interactions, however, are not only limited to the combining sites of lectins and glycotopes of glycoforms but also include other characteristics of the molecules, such as polyvalency, solubility, steric and charge factors. Therefore, studies on the mode of interaction between lectins and various well-defined glycoproteins, glycosphingolipids and polysaccharides should provide further information about actual requirements for binding. Many methods have been used to study the binding properties of lectins.[40-42] Recently, important data used to elucidate carbohydrate specificity of galectin was derived from the results of galectin-glycoform interactions analyzed by the Enzyme-Linked Lectinosorbent Assay (ELLSA) and inhibition of ELLSA.[43-45] For example, the profile of the combining site of recombinant tandem-repeat-type rat galectin-4 (G4-N) has been defined by inhibition of ELLSA as follows:[46] (i) the monosaccharide specificity is Gal/GalNAc > GlcNAc, Glc and L-Fuc, (ii) the disaccharide specificity is Galβ1→4Glc (L) / GalNAcβ1→3Gal (P) > Galβ1→3GalNAc (T) > Galβ1→3GlcNAc (I) > Galβ1→4GlcNAc (II), (iii) the most active ligand is L-Fucα1→2Galβ1→3GlcNAc β1→3Galβ1→4Glc (H active Iβ1-3L) or Galβ1→3 GlcNAcβ1→3Galβ1→4Glc (Iβ1-3L), (iv) the contribution of single multiantennary or small clusters of glycotopes, such as tri-II, Di-II and I and II bi-antennary is very limited, (v) a high density of polyvalent I/II, T and/or GalNAcα1

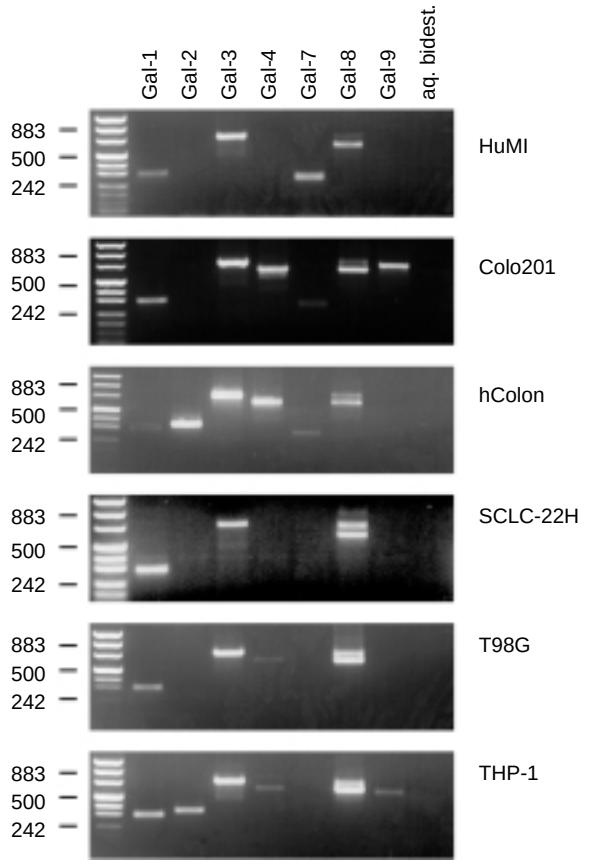


Fig. 9 Detection of galectin-specific transcripts in different human tumor cell lines (breast, colon, small-cell lung carcinoma, glioblastoma and acute monocytic leukemia) and the normal colonic line hColon, by RT-PCR analysis. cDNA preparations of cell lines were subjected to amplification by galectin-specific primers. Numbers on the left indicate the length of molecular weight markers (M). aq. bidest.: negative control in which water was used as a template. This figure is adapted from reference [30].

→Ser/Thr (Tn) strongly favors G4-N/glycoform binding. These glycans were up to $2.3x10^6$, $1.4x10^6$, $8.8x10^5$ and $1.4x10^5$ more active than Gal, GalNAc, monomeric I/II and T, respectively; (vi) while LFuc is a poor inhibitor, its presence as α1-2 linked to terminal Galβ1-containing oligosaccharides, such as H active Iβ1-3L, markedly enhances the reactivities of these ligands; (vii) when blood group A (GalNAcα1→) or B (Galα1→) determinants are attached to terminal Galβ1→3/4GlcNAc (or Glc) oligosaccharides, the reactivities are also increased; (viii) with LFucα1→3/4 linked to sub-terminal GlcNAc, the reactivities of these haptens are reduced; and (ix) short chain

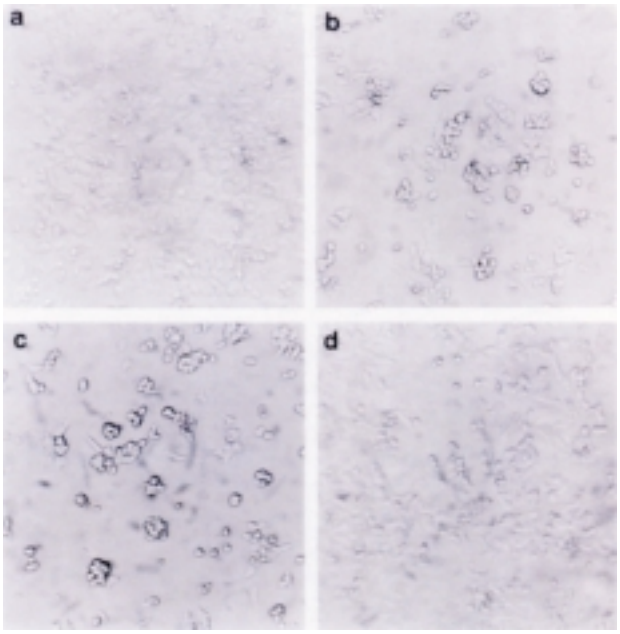


Fig. 10 Photomicrographs of neuroblastoma cell preparations cultured in the absence or presence of galectins. After seeding at an initial density of 10^4 cells/well and culturing for 16 hr to allow cell attachment, aliquots of individual cell batches were cultured for 48 hr in serum-supplemented medium (a) or the same medium containing 125 μg galectin-1/ml (b) as controls. The effect of the addition of 125 μg galectin-7/ml (c), and of galectin-7 in the presence of a tenfold excess of galectin-3 (d) was determined in parallel, under identical experimental conditions (magnification, 125x). This figure is adapted from reference [35].

Galβ1→3(Fucα→4)GlcNAc (Le^a) / Galβ1→4(Fucα1→3)GlcNAc (Le^x) / Fucα1→2 Galβ1→4(Fucα1→3)GlcNAc (Le^y) and the short chains of sialyl Le^a/Le^x are poor inhibitors. These distinct binding features of G4-N establish the important concept of affinity enhancement by high-density polyvalencies of glycotopes (vs. multi-antennary I/II) and by introduction of an ABH key sugar to Galβ1-terminated core glycotopes. The polyvalent ligand binding properties of G4-N may help our understanding of its crucial role for cell membrane raft stability and provide salient information for the optimal design of blocking substances, such as antitumoral glycodendrimers (please see below). In the system of ELLSA and inhibition of ELLSA, the amount of reagents required is greatly reduced as compared to that needed for Quantitative Precipitin and Precipitin-Inhibition assays.[47-49]

Ligand sequence, shape and spatial presentation: the six levels of affinity regulation

The ongoing systematic scrutiny of all aspects of the molecular interaction between a lectin and its ligand has allowed new insights into the enormous talents of glycoconjugate-based carbohydrates for biorecognition. The first level of affinity regulation is embodied by mono- or disaccharide discrimination, detected for example in hemagglutination assays (Table 6). Commercially available substances generally suffice for this first-step characterization. Reflected in the term *galectin*, the β-galactoside lactose (or thiodigalactoside) is a commonly used potent inhibitor for members of this lectin family. However this property by no means implies that the sugar specificity is an absolute constant feature for all galectins. Among them, the selectivity for a distinct disaccharide, e.g. Galβ1,4GlcNAc (LacNAc, N-acetyllactosamine), may already differ markedly, endowing certain lectins with distinctive capacity for ligand selection. This assumption was indeed validated, in this case, by measuring relative potency values of 5.6 and 700 compared to lactose in inhibition assays in two model studies using different galectins.[46-51] While sharing the galactose unit as a major contact point, selectivity exists at the level of the disaccharide. This result signifies that the elongation of the chain length of the sugar ligand guides us to the second level. The current availability of a wide panel of oligosaccharides by chemical synthesis or biochemical purification from natural sources makes it possible to map the fine-specificity of any lectin at level two. For example, it has become evident that human galectin-7 can bind internal LacNAc determinants (e.g. in α2,6-sialyl DiLacNAc), which, by contrast, are not accessible for galectin-1.[52] Constituting level two, the sequence of the carbohydrate structure, starting with linkage-point variations of the dimer, results in affinity modulation. In detail, the selection of the two hydroxyl groups for the glycosidic linkage (1-2, 1-3, 1-4 or 1-6) and also the introduction of substitutions such as α1,2-fucosylation and then addition of α1,3-Gal(GalNAc) to obtain ABH-blood group substances can all account for affinity changes. Increasing contact complementarity beyond the basic structural unit of lactose offers the potential for enthalpic gains by new hydrogen bonds. Notably, the introduction of the term *complementarity* directly leads us to the next level. By moving from the initial depiction of a carbohydrate structure as sequence (in two dimensions) to the third dimension (the shape), we reach level three of affinity regulation (Table 6). Here, a truly remarkable feature of glycans is revealed.

The fact that the chair conformation of the hexopyranose rings is energetically preferred pinpoints the dihedral angles ϕ, ψ of the glycosidic bond as the main source of flexibility of saccharides. One can visualize the basic principle of conformational mobility for a disaccharide by turning each hand (sugar unit) independently while maintaining thumb contact. Configurations where the pyranose rings come into spatial vicinity (clash) will resemble "hills" in a topographical map, giving ϕ, ψ, E-maps contour lines for energy levels (Fig. 11). Modeling studies and experimental monitoring by NMR spec-

Table 6. Six Levels of Affinity Regulation

Level	Definition	Example
1	Mono- and disaccharides	Incl. anomeric position
2	Oligosaccharides	Incl. branching and substitutions
3	Shape of oligosaccharide	Differential conformer selection
4	Spatial parameters of glycans in natural glycoconjugates	
	(a) Shape of glycan chain	Modulation of conformation by substitutions not acting as lectin ligand such as core fucosylation or introduction of bisecting GlcNAc in N-glycans, influence of protein part
	(b) Cluster effect with bi- to pentaantennary N-glycans or branched O-glycans	Incl. modulation by substitutions, please see (a).
5	Cluster effect with different but neighboring glycan chains on the same glycoprotein	In mucins
6	Cluster effect with different glycoconjugates on the cell surface by spatial vicinity	In lipid rafts or microdomains

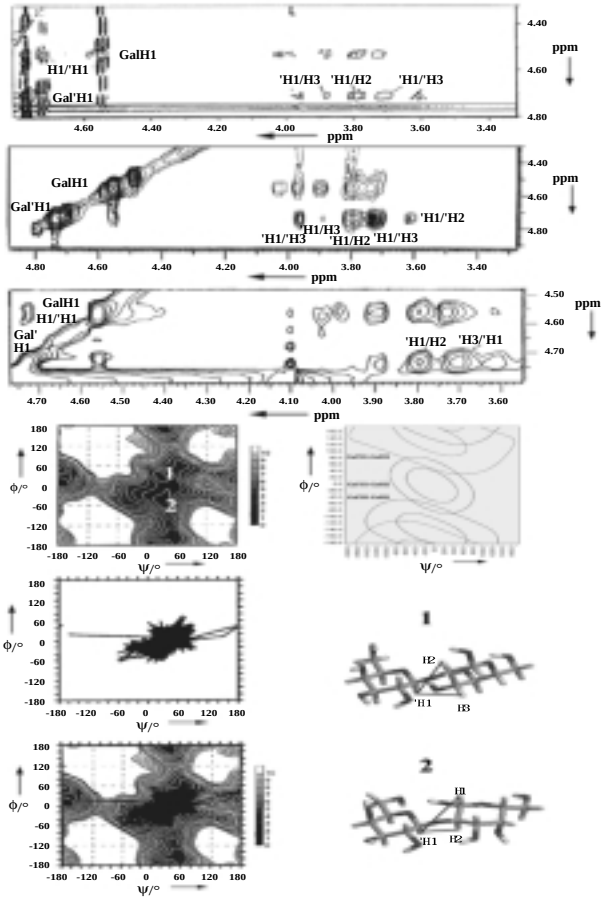


Fig. 11 Illustration of the structural aspects of differential conformer selection of a digalactoside by a plant and an animal lectin. Relevant parts of 2D ROESY/trNOESY spectra recorded at 500 MHz, 298 K and a mixing time of 100 ms of the free disaccharide Gal'β1-2Gal and ligand at a molar 10:1 ratio in the presence of the galactoside-specific mistletoe lectin (*Viscum album L.* agglutinin, VAA) and the chicken liver galectin (CG-16), respectively, show the presence of three interresidual cross-peaks in the free ligand and two such signals in the lectin-ligand complexes, as described previously. Presence of the interresidual H1'/H2 cross-peak is shared by the three spectra, whereas only one of the interresidual H1'/H1 and H1'/H3 cross-peaks is present in the trNOESY spectra with either the plant or the animal lectin (top panel). Molecular mechanics (ε = 4) and molecular dynamics calculations (ε = 80, CVFF, 300 K, 1000 ps) combined with the NMR-spectroscopically based contour line pairs (please see[66a and 67a] for details) revealed the binding of only one of the two conformers present in solution (labeled as 1 and 2 in the ϕ, ψ, E-plot) by these two lectins (bottom panel, left). The plant agglutinin and the animal lectin select different conformers. Their structures are presented in the right part of the bottom panel. This figure is adapted from reference [53].

troscopy have found that the accessible conformational space for a disaccharide is represented by "valleys" so that carbohydrate structures – in contrast to peptides – harbor limited flexibility.[51,53] Only distinct conformations are energetically favorable so that a small set of shapes constitutes a saccharide's three-dimensional structure.[44] In this respect, carbohydrate epitopes resemble peptide motifs in a protein. However, there is an important difference: protein sections are conformationally restrained by the indispensable presence of neighboring residues. Such expenditure to build special surroundings, in order to let a bioactive section of a protein adopt its suitable shape, is not necessary for carbohydrates: their distinct shapes simply come by nature. This property serves an enormous advantage for ligands in biorecognition. In the thermodynamic balance sheet, the entropic penalty incurred by binding such carbohydrates is less than what could be expected during the same process with a highly flexible ligand that loses its conformational freedom. As a result, we have pinpointed another feature rendering glycans predestined for a role in information transfer.

In allegoric terms, drawing on E. Fischer's famous lock-and-key principle,[54] one carbohydrate conformation can be likened to a key. We are accustomed to thinking that one sequence will build only one key. The given results on bioactive carbohydrates have changed this notion completely. This discovery signifies that one sequence can form a "bunch of keys".[55] The attractive consequence of the interconversion between low-energy conformers was described as, "the carbohydrate moves in solution through a bunch of shapes each of which may be selected by a receptor".[55] Indeed, depending on the nature of the receptor, the same sugar sequence can either be bioactive or bioinert. In other words, a certain carbohydrate ligand (such as Galβ1, 2Gal in Figure 11) is present in more than one conformation in solution. When we examined the bound-state conformation of two lectins, each was found exclusively as a single shape ("key"), leaving the other aside.[53] This principle is referred to as *differential conformer selection*.[56] It is explained graphically for Galβ1, 2Gal (its two shapes being shown in the lower right part of Figure 11) by presenting results of NMR spectroscopy measurements, computer-assisted molecular mechanics and dynamics calculations (for further details, please see the legend on Figure 11). The

binding of the pentasaccharide of ganglioside GM_1 by cholera toxin and galectin-1 using this mechanism may have medical application (please note that this process initiates the galectin-1-dependent effects on neuroblastoma growth shown in Fig. 10).[35,37,57] The pentasaccharide can form only a few distinct shapes in solution, which are defined by the measured ϕ, ψ-angle combinations at each glycosidic linkage. The respective numbers, which define the shape of a total of three conformers (keys), are presented in Table 7 (please see first row). The human lectin and the bacterial toxin both bind the pentasaccharide. In this context, the obvious question is whether both will select the same or two different conformations, or alternatively permit the ligand to change shape while remaining bound? As previously explained with experimental detail,[57] the galectin and the toxin select different conformers. These structures are shown in Figure 12. The clinical perspective of this result is evident: if an inhibitor is tailored to be arrested in the toxin-binding shape, a potent pharmaceutical without reactivity to the tissue lectin could be created.

Further structural scrutiny of this lectin-ligand complex also showed that carbohydrate units other than the primary contact site (here GalNAc/sialic acid) can significantly increase the interaction energy in the overall thermodynamics of binding.[57] In this case, enthalpic gains assigned to level two underlie conformer selection at level three. These additional contributions make it difficult to predict ligand affinity for oligosaccharides without experimental input, as exemplified for the ganglioside's pentasaccharide.[57] Thus, caution should be exercised when extrapolating binding data from simple structures (level one) to complex glycans (levels two and three). So far, we have only dealt with oligosaccharide chains, at best from a ganglioside. Besides extending a ligand's sequence and taking shape alterations into consideration, a look at the strategic presentation of CRDs in lectins (please see Fig. 4) intimates that lectin-reactive epitopes could also be displayed with impact on affinity. The choice for the branching mode of an N- or O-glycan would then mean more than a mere structural alteration picked from a panel of synthetic possibilities. This assumption leads us to further levels ligand affinity regulation (Table 6).

The basic structural motif of N-glycans is the biantennary chain. Further dissection of this element by enzymatic modifications yields: (a) substituted versions, e.g. containing a core fucose unit or a bisecting GlcNAc in the stem region, and (b) increase in the number of branches by adding up to three antennae.[7,8] Initially viewed as complication within the process of structural analysis, the question can now be addressed as to whether, and to what extent, these structural refinements serve to modulate affinity for lectins (levels four a, b in Table 6). Progress in chemoenzymatic synthesis of the substituted N-glycans, shown in Figure 13 for the route to the biantennary N-glycan with bisecting GlcNAc, made it possible to start to address this issue.[58-60] The reader may wonder whether the presence of such a substitution at a site distant from the lectin-reactive determinant in the branches can really affect the glycan's conformation. As shown in Figure 14, a substitution (here the bisecting GlcNAc) is, in fact, able to

Table 7. Average Dihedral ϕ, ψ-angles of Glycosidic Linkages in the Pentasaccharide Chain of Ganglioside GM_1 Free in Solution and in Complex with Human Galectin-1 in Solution or with Other Receptor Types in Crystal*

Disaccharide Type/Protein type†	Neu5Acα2-3Gal (ϕ / ψ)	Galβ1-3GalNAc (ϕ / ψ)	GalNAcβ1-4Gal (ϕ / ψ)	Galβ1-4Glc (ϕ / ψ)
Free saccharide	(-150/-10), (-70/30), (80/0)	(50/30), (50/0), (45/-60)	(25/25), (0/30), (-15/-30)	(-30/-30), (30/-15), (45/5)
Human galectin-1	-	(45/30), (45/0), (45/-30)	-	-
Human galectin-1	(-70/10)	-	-	(50/10), (20/-30)
Human galectin-1	(70/15)	(45/-5)	(25/25)	(-30/-30), (30/-15), (45/5)
Human galectin-1	-	(45/30), (45/0), (45/-30)	(25/25), (0/30), (-15/-30)	(-30/-30), (30/-15), (45/5)
Cholera toxin (2CHB, 3CHB)	(-172/-26), (-174/-18)	(54/-9), (53/-8)	(44/1), (33/6)	(36/-27), (-25/-33), (-37/-13), (47/0), (-34/-22), (-35/-21)

* When available, information is also listed on building blocks of the pentasaccharide; in these cases the sign - should be read as absence of the corresponding linkage in the respective GM_1-derived fragment.

† The Brookhaven Protein Data Bank code is given as the source for the listed information from (57)[SME7].

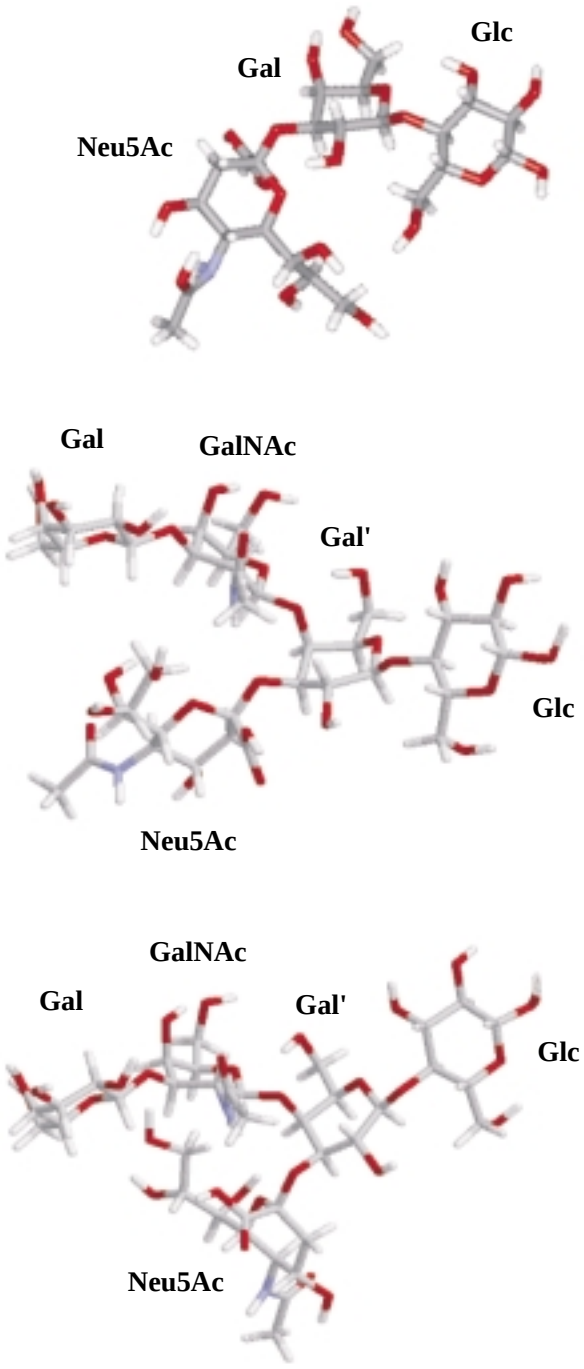


Fig. 12 Illustration of the bound-state conformation of α2,3-linked sialyllactose in complex with human galectin-1 (a), as well as the bound-state conformations of the pentasaccharide of ganglioside GM_1 in complex either with human galectin-1 (b) or with cholera toxin (Brookhaven Protein Bank code no. 2CHB/3CHB) (c), showing the result of differential conformer selection. This figure is adapted from reference [57].

alter the accessible conformational space of the N-glycan markedly, and recent reports substantiate that this effect on the conformational dynamics has an impact on lectin affinity, depending, notably, on the nature of the studied sugar receptor.[59,60] Thus, consideration of the complete chain shape and the topological characteristics of ligand presentation that it influences lead us from level three to level four. As listed in Table 6, this level is subdivided into two subcategories because substitutions can go along with increases in the branching mode. The way this parameter acts upon lectin affinity is convincingly documented by proof-of-principle studies on glycan clearance from circulation *in vivo*. Citing the classical example of the hepatocyte C-type lectin (asialoglycoprotein receptor; please see Table 4 and Fig. 4), an increase in valency from mono- and bi- to trivalent glycans resulted in a geometrical increase in affinity, termed the *glycoside cluster effect*.[61] Complemen-tarity between the presentation topology of this lectin's three CRDs and presentation of the sugar target causes the relative affinity to increase from 1 to 1,000 and finally reach 1,000,000 with triantennary glycans.[61] It goes without saying that this effect is of great potential for drug targeting to the liver.[62] Also, availability of a triantennary glycan in a glycoprotein from human ovarian cyst fluid was instrumental in revealing the strong impact of branching on the avian galectin CG-16 and poor inhibitory capacity for a mammalian CRD, i.e. the N-terminal domain of galectin-4.[50,51]

We previously stated that glycans as information carriers offer the potential for dynamic and reversible structural alterations. Also, the activity profile of glycosyltransferases can be the object of modulation with implications for the glycome, as detected in diseases.[63] So far, changes in glycosylation have primarily been mapped and interpreted phenomenologically. The presented reasoning and evidence enable the hypothesis that they can be functionally relevant. To give an example from clinical chemistry, the glycosylation changes of α_1-acid glycoprotein, with its five N-glycosylation sites in inflammation, have been related to its anti-inflammatory effects by interfering with leukocyte adhesion to endothelium through blockade of the selectins' C-type lectin subfamily.[64] Needless to say, that respective report opens a wide study area with clinical relevance. The take-home message is that the spatial vicinity of

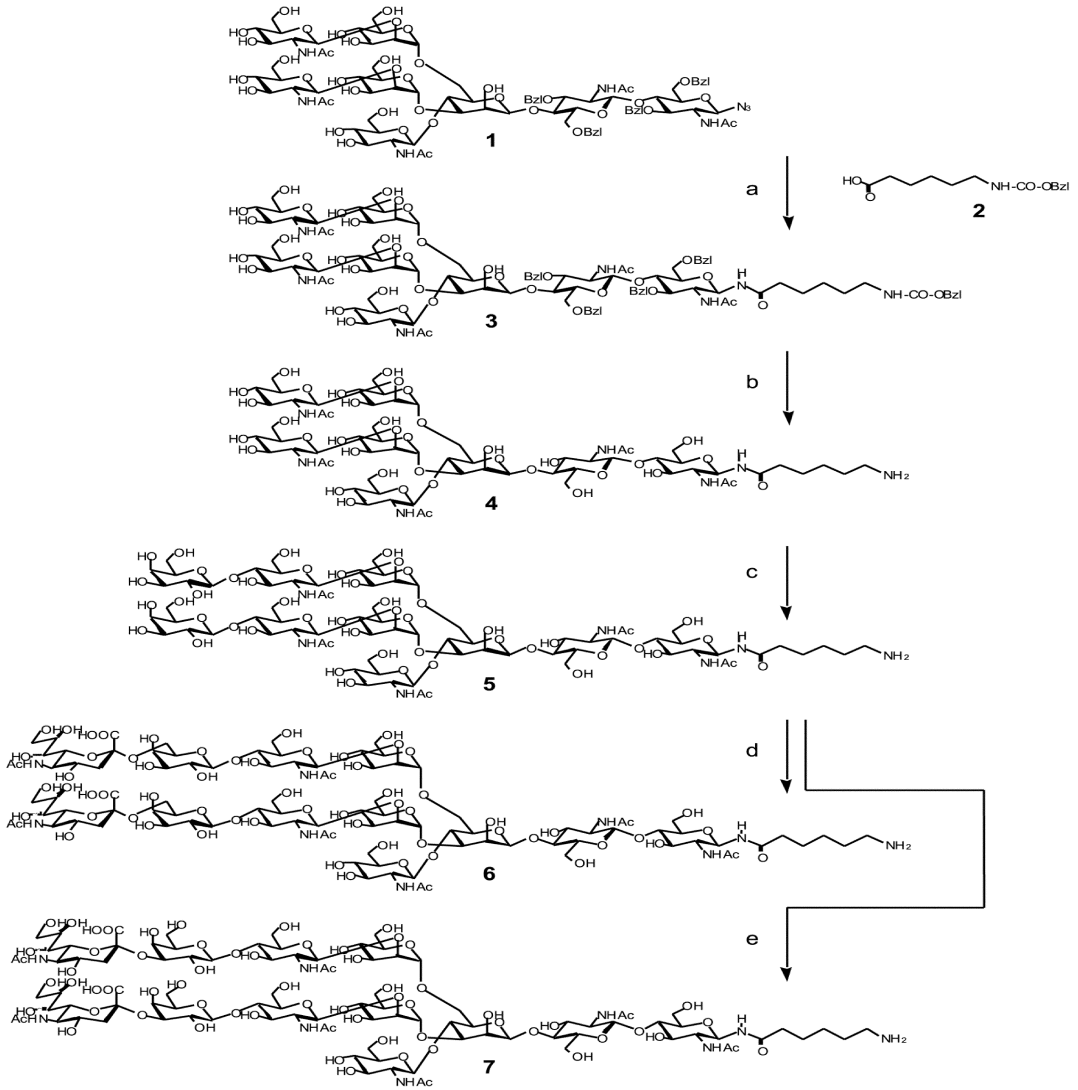


Fig. 13 Chemical and enzymatic steps to produce galactosylated and sialylated *N*-glycans substituted with bisecting GlcNAc. (a) 1. propanedithiol, Et_3N, MeOH; 2. *N*-benzyloxycarbonyl-6-aminohexanoic acid 2, TBTU, 1-hydroxybenzotriazole (HOBT) (1.-2.: 54 %); (b) Pd-H_2, AcOH, MeOH (95 %); (c) β1-4-galactosyltransferase, UDP-Gal, alkaline phosphatase (89 %); (d) α2-6-sialyltransferase, CMP-NeuNAc, alkaline phosphatase (c+d: 80 %); (e) α2-3-sialyltransferase, CMP-NeuNAc, alkaline phosphatase (c+e: 61 %). This figure is adapted from reference [60].

branches from an individual glycan, a factor valid for N- and O-glycans, is capable of having an effect on affinity for animal lectins. This consideration immediately guides us from level four to level five in Table 6. Of course, glycan clustering can also be achieved through spatial vicinity between different N- and/or O-glycans. Instructive examples are mucins with their high density of O-glycan chains, where branches of a single O-glycan and chains of different O-glycans are closely spaced together. Since mucin-like domains in glycoproteins are also not rare, a valid hypothesis would be that this topological aspect of glycosylation has a functional meaning for protein-carbohydrate interactions. Equally important, different N-glycans or antennae from different N- and O-glycans of a glycoprotein can cooperate in such a setting, a property lost when working with free glycans after digestion of the pro-

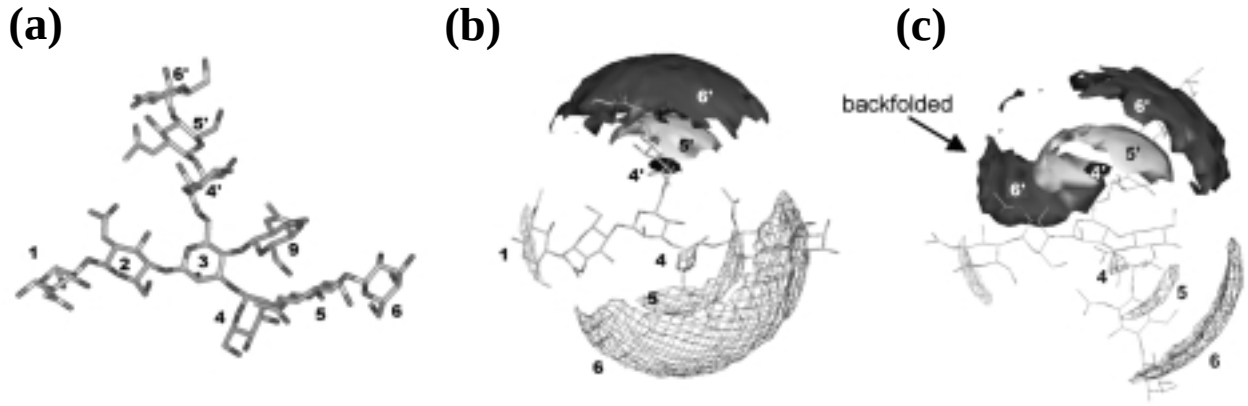


Fig. 14 Illustration of the nomenclature system for the *N*-glycan constituents (a), their inherent flexibility by isocontour plots (derived from the analysis of xyz population densities of each monosaccharide unit) at a constant energy level of 1.5 kcal/mol for the biantennary nonasaccharide (b), and the decasaccharide containing the bisecting GlcNAc (c). For convenient comparison, the conformations were identically positioned in space by superimposing the ring atoms of the mannose units of the pentasaccharide core. Access to the conformational space in the vicinity of the linear part of the core for the terminal galactose moiety of the α1-6 arm is emphasized by introducing the term 'backfolded' into the figure (c). This figure is adapted from reference [60].

tein carrier. In order to examine the extent to which this level of affinity regulation is operative, a panel of well-characterized glycoproteins is indispensable, as are systematic studies of a lectin family, which are currently in progress for galectins.[50,51]

It is at the cell surface that the sixth level of affinity regulation, i.e. the modulation of local density of lectin-reactive glycans, is encountered. Microdomains, lipid rafts or glycosynapses are candidate structures to facilitate cell adhesion or signaling by lectin-carbohydrate recognition.[65] Membrane fluidity, aggregate formation of glycoproteins/glycolipids and changes in glycan structures afford an array of possibilities for a wide range of regulatory events. A topological parameter arrests or shifts conformational equilibria for glycan determinants by altering their density. To let this system work requires a conspicuous level of inherent lectin specificity toward certain cell surface glycans. In other words, despite binding a frequently-present basic element on level one, a lectin is expected to be very selective for "real" cell surface ligands. Indeed, this is the case for galectins. A survey of the literature attests that galectins-1 and -3 react only with distinct glycoproteins or glycolipids to trigger cell-type-specific activities (Table 8). Obviously, it is justified to define level six, hereby giving further research direction and aim. To avoid making this work appear purely academic, it is worth familiarizing the reader with perspectives for medical applications focusing on galectins.

Perspectives for galectin-dependent medical applications

Table 9 is a summary of the means by which basic research on galectins can be turned into clinical application is presented. The monitoring of the presence/absence of distinct lectins or lectin ligands is

Table 8a. Glycan Lligands of Galectin-1

Laminin	Lamp-1
Fibronectin (tissue)	L1
$\alpha_1\beta_1$,$\alpha_7\beta_1$-integrins	Mac-2 BP
Thrombospondin	GI mucin
CEA	CA125
Chondroitin sulfate PG	Hsp90-like GP
CD2, 3, 4, 7, 43, 45	Thy-1
Distinct neutral GL	Ganglioside GM_1

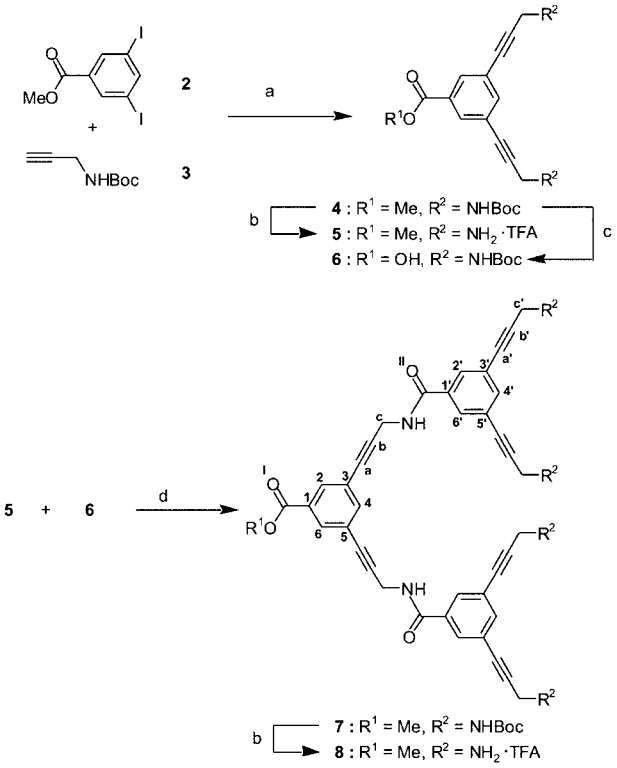
Table 8b. Glycan Ligands ofGalectin-3 (Mac-2)

Laminin	Lamp-1/-2
β_1-integrin	Tenascin
MAG	Hensin (DMBT-1)
Mac-2 BP	CD11b of CD11b/DC18
Cubilin	(Mac-1 antigen)
MP20 (tetraspanin)	Mac-3
CEA	Colon cancer mucin
CD32; CD66a,b; Cd98	LI-cadherin

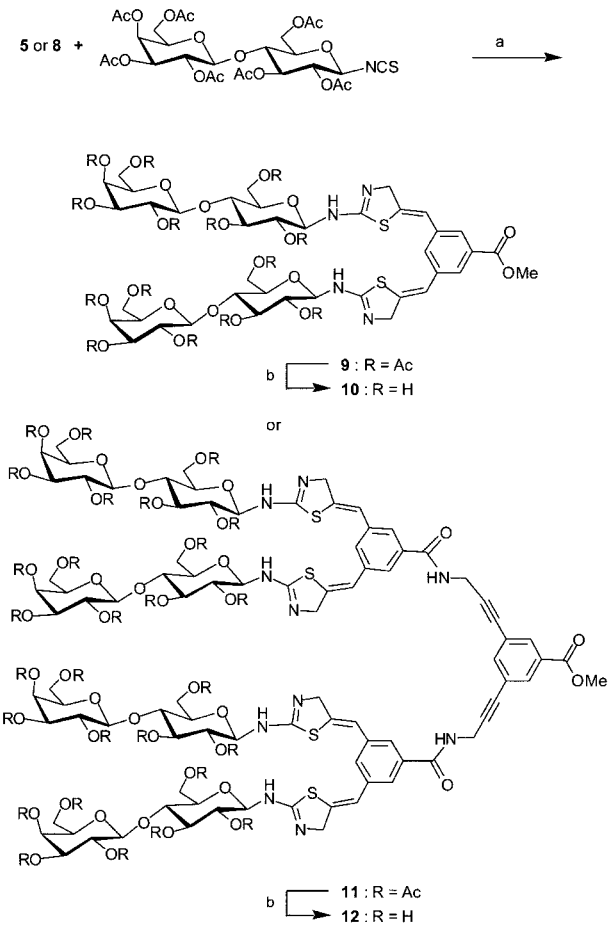
Table 9. Medical Applicability of Galectins and Their Ligands

Method	Aim
Immuno (or glyco-) histochemistry with galectin-specific antibody (or neoglycoconjugate)	Tumor diagnosis and prognostic evaluation
Lectin histochemistry with labeled galectins for ligand localization	Tumor diagnosis and prognostic evaluation
Galectin as therapeutic (suitable expression vector, purified protein or galectin-mimetic peptide)	Regulation of cell proliferation (e.g. activated T lymphocytes or lymphoma cells, neuroblastoma cells, susceptible carcinoma cells)
Downregulation of galectin expression (e.g. by siRNA), application of blocking antibody or glycoclusters with optimal design (exploiting all levels of specificity)	Inhibition of tissue invasion or metastasis (e.g. glioblastoma, colon carcinoma)
Design of mutants (e.g. L11A substitution of galectin-1)	Impairment of functionality of oncogenic H-ras
Rational manipulation of ligand presentation (switch on/off via α2,6-sialylation)	To increase/decrease cell susceptibility to galectin(s)
Rational manipulation of glycoconjugate expression (increase/decrease of cell surface presentation of distinct targets such $\alpha_5\beta_1$-integrin or ganglioside GM_1)	To increase/decrease cell susceptibility to galectin(s)

already a current topic in tumor pathology, with examples for new prognostic information given above. *In vitro* and animal models, including cell clones with deliberate changes in galectin expression, are and will be helpful to explore the relevance of a certain lectin for tumor progression in distinct tumor classes. These experiments have a therapeutic dimension as well. Merging the information on the importance of spatial factors for ligand presentation with the power of synthetic chemistry to prepare custom-made blocking substances, has made it feasible to synthesize target-selective inhibitors. Naturally branched glycans give precedence for devising suitable neoglycoconjugates.[66,67] These tools cannot only be employed to block CRDs but also to localize the sugar-binding capacity on cells and in tissue sections.[68] Current experience in this area has shown that strong multivalency effects are attainable. When the synthetic route given in Scheme 1 and Scheme 2 was followed, an enhancement of almost 4300-fold compared to the monomeric derivative was determined for the chimera-type galectin-3 with a tetravalent scaffold presenting lactose as ligand.[68] The rigid tetravalent glycodendrimer containing lactose-2-aminothiazoline units at the glycodendrimer periphery was prepared from cyclization of the thiourea sulfur onto the triple bond of the spacer (Scheme 2). Notably, mono- and dimeric prototype galectins were considerably less reactive with this glycocluster.[68] Hence, it follows that the design of custom-made inhibitors is furnished by accrued knowledge on each of the relevant levels of specificity (Table 6) in order to tailor optimal glycoclusters for a particular purpose, e.g. interfering with the prometastatic capacity of galectin-3 in colon cancer models. It is not the creativity of the chemist to supply suitable scaffolds that is the limiting factor to turn this concept into experiments.[66,67] While synthetic products are capable of localizing and blocking lectins, the development of such tools through molecular biological engineering is instrumental for another aim – to rationally modulate lectin expression. Down-regulation of galectin-1, which is responsible for tissue invasion in glioblastoma, is a perspective supported by the current experimental evidence presented above. Establishing proapoptotic activity in autoreactive T cells is a concept for therapeutic management of autoimmune disorders. Since any lectin effect is dependent on the specific selection of ligands (Table 8), manipulating ligand presentation can be equally effective. A key role in this respect is played by the α2,6-sialyltransferase which introduces a switch-off signal to ligand properties for galectin-1. Targeting distinct glycosyltransferases is a long-term goal in order to give a deliberate direction to the galectin signaling network with consequent clinical benefit (Table 9). Glycomic profiling with plant lectins or monoclonal antibodies in pathology, which maps disease-associated glycan changes, has thus gone from phenomenology to defining functionally relevant shifts in the pattern of potential ligands for tissue lectins. With a growing understanding of the functional relevance of disease-



Scheme 1. Reaction conditions: (a) $Pd^{(0)}$, CuI, NEt_3, CH_3CN, 2 d (65%); (b) TFA, CH_2Cl_2 20 min (quant.); (c) Tesser's base, 14 h (quant.); (d) BOP, *i*Pr_2NEt, CH_3CN, 14 h (71%). This figure is adapted from reference [68].



Scheme 2. Reaction conditions: (a) first *i*Pr_2NEt, CH_2Cl_2, then AcOH, CH_2Cl_2, 14 h (65%) for 9; CH_3CN; *i*Pr_2NEt, CH_2Cl_2, 14 h (33%) for 10; (b) NaOMe, MeOH 2 h, (70%) for 11; dioxane, MeOH, 1 M NaOH, 3 h (94%) for 12. This figure is adapted from reference [68].

associated alterations in glycosylation, the envisaged interventions are likely to be more than symptomatic treatment modalities.

Conclusions

When asked to deliver input on how to construct an operative artificial cell, special attention will have to be given to the design of cell surface determinants. They are required to act as versatile sensors for communication with the environment. Spatial accessibility and a molecular vocabulary to encode a wide array of messages with high density (the space on the cell surface, after all, is limited) are salient parameters for functionality. When looking at natural models, these prerequisites are completely fulfilled by glycocalyx carbohydrates. In fact, the sugars of cellular glycoconjugates are second to no other class of biomolecules in terms of information storage. Furthermore, the emerging insights into the expression of binding partners for the information-storing sugar epitopes have given the concept of the sugar code its present shape: lectin-carbohydrate interactions transmit the information from glycoconjugate oligosaccharides into diverse biological responses. The initially puzzling degree of glycan structure diversity thereby gains a functional dimension. Topological aspects of the two sides in this recognition system, i.e. glycan and lectin, are presently being related to affinity modulation with medical application perspectives. In fact, new diagnostic tools have already been established, attesting to the current status of results presented here. Moreover, experimental approaches for examining therapeutic perspectives have been inspired. Merging the fields

of synthetic and analytical chemistry, biochemistry and molecular cell biology with medical practices, such as oncology or pathology, is sure to pave the way to a new level of understanding on the role of this recognition system in disease.

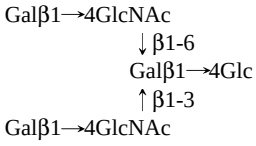
Acknowledgement

This work was supported by grants from: the Chang Gung Medical Research Project (CMRP No. 1028), Kwei-san, Tao-yuan, Taiwan; the National Science Council (NSC 92-2311-B-182-005, 92-2320-B-182-045 and 92-2320-B-182-046), Taipei, Taiwan; and the Mizutani Foundation for Glycoscience.

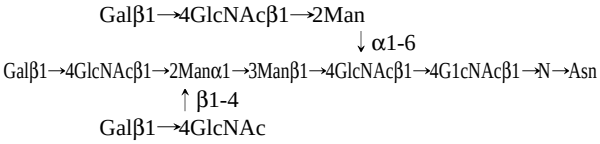
Abbreviations used:

L, Galβ1→4Glc; P, GalNAcβ1→3Gal; Thomsen-Friedenreich disaccharide (T) Galβ1→3GalNAc; Tn, GalNAcα1→Ser/Thr; H, LFucα1→2Gal, human blood group H specific disaccharide; I, Galβ1→3GlcNAc, human blood group type I precursor sequence; II, Galβ1→4GlcNAc, human blood group type II precursor sequence; Iβ1→3L (type 1), Galβ1→3GlcNAcβ1→3 Galβ1→4Glc, lacto-*N*-tetraose; IIβ1→3L (type II), Galβ1→4GlcNAcβ1→3Galβ1→4Glc, lacto-*N*-neo-tetraose; Le^a, Galβ1→3(Fucα1→4)GlcNAc; Le^x, Galβ1→4(Fucα1→3)GlcNAc; Le^y, Fucα1→2Galβ1→4 (Fucα1→3)GlcNAc.

Di-II, Lacto-*N*-hexaose



Tri-II, Tri-antennary Galβ1→4GlcNAc (II) glycopeptides from asialofetuin.[70]



REFERENCE

1. Sharon N. Lectins: from obscurity into the limelight. Protein Sci 1998;7:2042-8.
2. Laine RA. The information-storing potential of the sugar code. In: Gabius HJ, Gabius S, eds. Glycosciences: Status and Perspectives. London-Weinheim: Chapman & Hall, 1997:1-14.
3. Roseman S. Reflections on glycobiology. J Biol Chem 2001;276:41527-42.
4. Hounsell EF. Methods of glycoconjugate analysis. In: Gabius HJ, Gabius S, eds. Glycosciences: Status and Perspectives. London-Weinheim: Chapman & Hall, 1997:15-29.
5. Geyer H, Geyer R. Strategies for glycoconjugate analysis. Acta Anat 1998;161:18-35.
6. Kobata A. Structures and functions of the sugar chains of glycoproteins. Eur J Biochem 1992;209:483-501.
7. Sharon N, Lis H. Glycoproteins: structure and function. In: Gabius HJ, Gabius S, eds. Glycosciences: Status and Perspectives. London-Weinheim: Chapman & Hall, 1997:133-62.
8. Reuter G, Gabius HJ. Eukaryotic glycosylation: whim of nature or multipurpose tool? Cell Mol Life Sci 1999;55:368-422.
9. Gabius HJ. Animal lectins. Eur J Biochem 1997;243:543-76.
10. Szathmáry E. Why are there four letters in the genetic alphabet? Nat Rev Genet 2003;4:995-1001.
11. Brockhausen I, Schachter H. Glycosyltransferases involved in N- and O-glycan biosynthesis. In: Gabius HJ, Gabius S, eds. Glycosciences: Status and Perspectives. London-Weinheim: Chapman & Hall, 1997:79-113.
12. Gabius HJ, André S, Kaltner H, Siebert HC. The sugar code: functional lectinomics. Biochim Biophys Acta 2002;1572:165-77.
13. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 2002;12:43R-56R.
14. Lis H, Sharon N. Lectins: carbohydrate-specific proteins that mediate cellular recognition. Chem Rev 1998;98:637-74.
15. Rüdiger H, Gabius HJ. Plant lectins: occurrence, biochemistry, functions and applications. Glycoconj J 2001;18:589-613.
16. Mitchell SW. Researches upon the venom of the rattlesnake. Smithsonian Contributions to Knowledge 1860;XII:89-90.
17. Ogilvie ML, Gartner TK. Identification of lectins in snake venoms. J Herpetol 1984;18:285-90.
18. Sharon N, Lis H. The structural basis for carbohydrate recognition by lectins. Adv Exp Med Biol 2001;491:1-16.
19. Solís D, Jiménez-Barbero J, Kaltner H, Romero A, Siebert HC, von der Lieth CW, Gabius HJ. Towards defining the role of glycans as hardware in information storage and transfer: basic principles, experimental approaches and recent progress. Cells Tissues Organs 2001;168:5-23.
20. Linehan SA, Marínez-Pomares L, Gordon S. Macrophage lectins in host defence. Microbes Infect 2000;2:279-88.
21. Kilpatrick DC. Mannan-binding lectin: clinical significance and applications. Biochim Biophys Acta

2002;1572:401-13.
22. Gabius HJ. (ed.) Special issue on animal lectins. Biochim Biophys Acta 2002;1572:163-434.
23. Wu AM. Carbohydrate structural units in glycosphingolipids as receptors for Gal and GalNAc reactive lectins. Neurochem Res 2002;27:593-600.
24. Wu AM. Carbohydrate structural units in glycoproteins and polysaccharides as important ligands for Gal and GalNAc reactive lectins. J Biomed Sci 2003;10:676-88.
25. Winterburn PJ, Phelps CF. The significance of glycosylated proteins. Nature 1972;236:147-51.
26. Hennet T. The galactosyltransferase family. Cell Mol Life Sci 2002;59:1081-95.
27. Teichberg VI, Silman I, Beitsch DD, Resheff G. A β-D-galactoside binding protein from electric organ tissue of *Electrophorus electricus*. Proc Natl Acad Sci USA 1975;72:1383-7.
28. Cooper DN. Galectinomics: finding themes in complexity. Biochim Biophys Acta 2002;1572:209-31.
29. Gabius HJ, Brehler R, Schauer A, Cramer F. Localization of endogenous lectins in normal human breast, benign breast lesions and mammary carcinomas. Virchows Arch B Cell Pathol Incl Mol Pathol 1986;52:107-15.
30. Lahm H, André S, Höeflich A, Fischer JR, Sordat B, Kaltner H, Wolf E, Gabius HJ. Comprehensive galectin fingerprinting in a panel of 61 human tumor cell lines by RT-PCR and its implications for diagnostic and therapeutic procedures. J Cancer Res Clin Oncol 2001;127:375-86.
31. Lahm H, André S, Hoeflich A, Fischer JR, Sordat B, Kaltner H, Wolf E, Gabius HJ. Molecular biological fingerprinting of human lectin expression by RT-PCR. Methods Enzymol 2003;362:287-97.
32. Nagy N, Legendre H, Engels O, André S, Kaltner H, Wasano K, Zick Y, Pector JC, Decaestecker C, Gabius HJ, Salmon I, Kiss R. Refined prognostic evaluation in colon carcenoma using immunohistochemical galectin fingerprinting. Cancer 2003;97:1849-58.
33. Camby I, Belot N, Lefranc F, Sadeghi N, de Launoit Y, Kaltner H, Musette S, Darro F, Danguy A, Salmon I, Gabius HJ, Kiss R. Galectin-1 modulates human glioblastoma cell migration into the brain through modifications to the actin cytoskeleton and levels of expression of small GTPases. J Neuropathol Exp Neurol 2002;61:585-96.
34. Gabius HJ. Probing the cons and pros of lectin-induced immunomodulation: case studies for the mistletoe lectin and galectin-1. Biochimie 2001;83:659-66.
35. Kopitz J, von Reitzenstein C, André S, Kaltner H, Uhl J, Ehemann V, Cantz M, Gabius HJ. Negative regulation of neuroblastoma cell growth by carbohydrate-dependent surface binding of galectin-1 and functional divergence from galectin-3. J Biol Chem 2001;276:35917-23.
36. Kopitz J, André S, von Reitzenstein C, Versluis K, Kaltner H, Pieters RJ, Wasano K, Kuwabara I, Liu FT, Cantz M, Heck AJ, Gabius HJ. Homodimeric galectin-7 (p53-induced gene 1) is a negative growth regulator for human neuroblastoma cells. Oncogene 2003;22:6277-88.
37. Kopitz J, von Reitzenstein C, Burchert M, Cantz M, Gabius HJ. Galectin-1 is a major receptor for ganglioside GM_1, a product of the growth-controlling activity of a cell surface ganglioside sialidase, on human neuroblastoma cells in culture. J Biol Chem 1998;273:11205-11.
38. Brewer CF. Lectin cross-linking interactions with multivalent carbohydrates. Adv Exp Med Biol 2001;491:17-25.
39. Fred Brewer C. Binding and cross-linking properties of galectins. Biochim Biophys Acta 2002;1572:255-62.
40. Gabius HJ, André S, Danguy A, Kayser K, Gabius S. Detection and quantification of carbohydrate-binding sites on cell surfaces and on tissue sections by neoglycoproteins. Methods Enzymol 1994;242:37-46.
41. Gargir A, Shtevi A, Gortler M, Dukler A, Dotan N. GlycoChip™- high throughput technology for screening and analysis of protein-glycan interactions. Glycoconj J 2001;18:26, Abstr. C1.20.
42. Bryan MC, Plettenbutg O, Sears P, Rabuka D, Wacowich-Sgarbi S, Wong CH. Saccharide display on microtiter plates. Chem Biol 2002;9:713-20.
43. Duk M, Lisowska E, Wu JH, Wu AM. The biotin/avidin-mediated microtiter plate lectin assay with the use of chemically modified glycoprotein ligand. Anal Biochem 1994;221:266-72.
44. Wu AM, Wu JH, Chen Y, Tsai M, Herp A. Forssman pentasaccharide and polyvalent Galβ1→4GlcNAc as major ligands with affinity for *Caragana arborescens* agglutinin. FEBS Lett 1999;463:225-30.
45. Wu AM, Song SC, Chen YY, Gilboa-Garber N. Defining the carbohydrate specificities of *aplysia gonad* lectin exhibiting a peculiar D-galactouronic acid affinity. J Biol Chem 2000;275:14017-24.
46. Wu AM, Wu JH, Liu JH, Singh T, André S, Kaltner H, Gabius HJ. Effects of polyvalency of glycotopes and natural modifications of human blood group ABH/Lewis sugars at the Galβ1-terminated core saccharides on the binding of domain-I of recombinant tandem-repeat-type galectin-4 from rat gastrointestinal tract (G4-N). Biochimie 2004;86:317-26.
47. Wu AM. Structural concepts of the blood group A, B, H, Le^a, Le^b, I and i active glycoproteins purified from human ovarian cyst fluid. Adv Exp Med Biol 1988;228:351-94.
48. Wu AM, Lin SR, Chin LK, Chow LP, Lin JY. Defining the carbohydrate specificities of *Abrus precatorious* agglutinin as T (Galβ1→3GalNAc) greater than I/II (Galβ1→3/4GlcNAc). J Biol Chem 1992;267:19130-9.
49. Wu AM, Song SC, Chang SC, Wu JH, Chang KS, Kabat EA. Further characterization of the binding properties of a GalNAc specific lectin from *Codium fragile* subspecies *tomentosoides*. Glycobiology 1997;7:1061-6.
50. Wu AM, Wu JH, Tsai MS, Kaltner H, Gabius HJ. Carbohydrate specificity of a galectin from chicken liver (CG-16). Biochem J 2001;358:529-38.
51. Wu AM, Wu JH, Tsai MS, Liu JH, André S, Wasano K, Kaltner H, Gabius HJ. Fine specificity of domain-I of

recombinant tandem-repeat-type galectin-4 from rat gastrointestinal tract (G4-N). Biochem J 2002;367:653-64.
52. Ahmad N, Gabius HJ, Kaltner H, André S, Kuwabara I, Liu FT, Oscarson S, Norberg T, Brewer CF. Thermodynamic binding studies of cell surface carbohydrate epitopes to galectin-1, -3 and -7: Evidence for differential binding specificities. Can J Chem 2002;80:1096-104.
53. Gabius HJ, Siebert HC, André S, Jiménez-Barbero J, Rüdiger H. Chemical biology of the sugar code. Chem Bio Chem 2004;5:740-64.
54. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. Ber Dt Chem Ges 1894;27:2985-93.
55. Hardy BJ. The glycosidic linkage flexibility and timescale similarity hypotheses. J Mol Struct 1997;395-396:187-200.
56. Gabius HJ. The how and why of protein-carbohydrate interaction: a primer to the theoretical concept and a guide to application in drug design. Pharm Res 1998;15:23-30.
57. Siebert HC, André S, Lu SY, Frank M, Kaltner H, van Kuik JA, Korchagina EY, Bovin N, Tajkhorshid E, Kaptein R, Vliegenthart JF, von der Lieth CW, Jiménez-Barbero J, Kopitz J, Gabius HJ. Unique conformer selection of human growth-regulatory lectin galectin-1 for ganglioside GM1 versus bacterial toxins. Biochemistry 2003;42:14762-73.
58. André S, Unverzagt C, Kojima S, Dong X, Fink C, Kayser K, Gabius HJ. Neoglycoproteins with the synthetic complex biantennary nonasaccharide or its α2,3/α2,6-sialylated derivatives: their preparation, assessment of their ligand properties for purified lectins, for tumor cells *in vitro*, and in tissue sections, and their biodistribution in tumor-bearing mice. Bioconjug Chem 1997;8:845-55.
59. Unverzagt C, Andre S, Seifert J, Kojima S, Fink C, Srikrishna G, Freeze H, Kayser K, Gabius HJ. Structure-activity profiles of complex biantennary glycans with core fucosylation and with/without additional α2,3/α2,6 sialylation: synthesis of neoglycoproteins and their properties in lectin assays, cell binding, and organ uptake. J Med Chem 2002;45:478-91.
60. André S, Unverzagt C, Kojima S, Frank M, Seifert J, Fink C, Kayser K, von der Lieth CW, Gabius HJ. Determination of modulation of ligand properties of synthetic complex-type biantennary N-glycans by introduction of bisecting GlcNAc in silico, in vitro and in vivo. Eur J Biochem 2004;271:118-34.
61. Lee RT, Lee YC. Neoglycoconjugates. In: Gabius HJ, Gabius S, eds. Glycosciences: Status and Perspectives. London-Weinheim: Chapman & Hall, 1997:55-77.
62. Yamazaki N, Kojima S, Bovin NV, André S, Gabius S, Gabius HJ. Endogenous lectins as targets for drug delivery. Adv Drug Deliv Rev 2000;43:225-44.
63. Brockhausen I, Schutzbach J, Kuhns W. Glycoproteins and their relationship to human disease. Acta Anat 1998;161:36-78.
64. van Dijk W, Brinkman-Van der Linden EC, Havenaar EC. Glycosylation of α_1-acid glycoprotein (orosomucoid) in health and disease: occurrence, regulation and possible functional implications. Trends Glycosci Glycotechnol 1998;10:235-45.
65. Hakomori S. Glycosylation defining cancer malignancy: new wine in an old bottle. Proc Natl Acad Sci USA 2002;99:10231-3.
66. Bovin NV, Gabius HJ. Polymer-immobilized carbohydrate ligands: versatile chemical tools for biochemistry and medical sciences. Chem Soc Rev 1995;24:413-21.
67. Roy R. A decade of glycodendrimer chemistry. Trends Glycosci Glycotechnol 2003;15:291-310.
68. Vrasidas I, André S, Valentini P, Böck C, Lensch M, Kaltner H, Liskamp RM, Gabius HJ, Pieters RJ. Rigidified multivalent lactose molecules and their interactions with mammalian galectins: a route to selective inhibitors. Org Biomol Chem 2003;1:803-10.
69. Gabius HJ. Glycohistochemistry: the why and how of detection and localization of endogenous lectins. Anat Histol Embryol 2001;30:3-31.
70. Wu JH, Song SC, Chen YY, Tsai MC, Kabat EA, Wu AM. Multi-antennary Galβl-4GlcNAc ard Galβ1-3GalNAc clusters as important ligands for a lectin isolated from the sponge *Geodia cydonium*. FEBs Lett 1998;427:134-8.

內源性凝集素 (Endogenous Lectin) 的新機能：涉及醫學意含的初階概念及以一種半乳醣凝集素 (Galectin) 爲實例的探討

翰斯・加米士 吳明道[1]

一般的生物化學教科書所提及對於生物資訊所儲存和傳送的硬體僅限於遺傳密碼的單位字母核酸和氨基酸。然而在生物資訊上有著顯著能力的第三類分子（碳水化合物）卻因此被忽略。同樣以數個構單元所聚合之結構的碳水化合物（例如寡醣）而言其所具有的編碼能力遠超過核酸和蛋白質所能攜帶。在細胞的醣共軛物質例如醣蛋白 (glycoprotein) 蛋白質黏多糖 (proteoglycans) 醣脂質 (glycolipids) 通常被視爲和能量的新陳代謝相關，但事實上他們有其他更重要的任務。今日我們對碳水化合物的認識爲-有效的高密度生物資訊編碼系統。複雜的酵素使細胞得以多樣化也產生指紋般的醣質體 (glycome)。更甚者，修改細胞認知動力學或了解惡性腫瘤轉移會很快成爲有可能的事情。轉譯寡醣上所存在的資訊是由凝集素 (lectins) 以生物反應的方式進行。天賦的免疫系統對於外來醣訊息的識別細胞對細胞／質間作用力的調控 轉移 成長和細胞內黏多醣的衰退等證明了凝集素-碳水化合物在生理上廣汎的功能性。內源的凝集素系統是受到序列組合和明顯的固定特徵所引導。例如捲心蛋糕的摺疊和關鍵基團安置在堆積中的維持和 C-H／π-電子作用力與有同等重要性的指向β-半乳醣苷側鏈 (galactoside) 之受質的方向性氫鍵，爲半乳醣凝集素的共同結構特徵。生物化學和生物物理學學者開始闡明錯綜複雜的內源性受質受受限制的選擇作用。例如嵌入素 (integrins) 神經節苷脂 (ganglioside) GM_1 和其對於細胞群體生物學方面相關連的生物實驗等…腫瘤細胞的生長和入侵的調控，或 T 細胞所活化的細胞程式性凋亡 (apoptosis)。組織病理學監測和細胞生物學的合作研究某種家族分子成員和腫瘤的發展或抑制的關聯。進一步著眼醣密碼轉譯所形成的結果預期將會有對診斷和療程有明顯的影響。(長庚醫誌 2006;29:37-62)

關鍵字：凝集素，細胞程式性凋亡，藥物設計，半乳醣凝集素，醣共軛物，碳水化合物結合蛋白，轉移，醣密碼。

路德維格-馬克斯米利安大學 德國；[1]長庚大學 醫學院 醣類免疫化學研究室
受文日期：民國94年7月29日；接受刊載：民國94年9月16日
德國通訊作者：翰斯・加米士教授，Institute of Physiology, Physiological Chemistry and Animal Nutrition, Faculty of Veterinary Medicine, Ludwig-Maximilians-University, Munich, Veterinaerstr. 13, 80539 Munich, Germany. Tel.: 002-49-89-2180-2290; Fax: 002-49-89-2180-2508; E-mail: gabius@lectins.de
台灣通訊作者：吳明道教授，長庚大學 醫學院 醣類免疫化學研究室。桃園縣333龜山鄉文化一路259號。Tel.: (03)2118800轉5175或5176; Fax: (03)2118456; E-mail: amwu@mail.cgu.edu.tw