Chapter 10

Biological Information Transfer Beyond the Genetic Code: The Sugar Code*

Hans-Joachim Gabius

Abstract In the era of genetic engineering, cloning, and genome sequencing, the focus of research on the genetic code has received an even further accentuation in the public eye. When, however, aspiring to understand intra- and intercellular recognition processes comprehensively, the two biochemical dimensions established by nucleic acids and proteins are not sufficient to satisfactorily explain all molecular events in, e.g. cell adhesion or routing. To bridge this gap consideration of further code systems is essential. A third biochemical alphabet forming code words with an information storage capacity second to no other substance class in rather small units (words, sentences) is established by monosaccharides (letters). As hardware oligosaccharides surpass peptides by more than seven orders of magnitude in the theoretical ability to build isomers, then the total of conceivable hexamers is calculated. Beyond the sequence complexity application of nuclear magnetic resonance (NMR) spectroscopy and molecular modeling have been instrumental to discover that even small glycans can often reside in not only one but several distinct low-energy conformations (keys). Intriguingly, conformers can display notably different capacities to fit snugly into the binding site of nonhomologous receptors (locks). This process, experimentally verified for two classes of lectins, is termed "differential conformer selection." It adds potential for shifts of the conformer equilibrium to modulate ligand properties dynamically and reversibly to the well-known changes of sequence (including anomeric positioning and linkage points) and of pattern of substitution, e.g. by sulfation. In the intimate interplay with sugar receptors (lectins, enzymes, and antibodies) the message of coding units of the sugar code is deciphered. This communication will trigger postbinding signaling and the intended biological response. Knowledge about the driving forces for the molecular rendezvous, that is, contributions of bidentate or cooperative hydrogen bonds, dispersion forces, stacking and solvent rearrangement, will enable the design

of high-affinity ligands or mimetics thereof. They embody clinical applications reaching from receptor localization in diagnostic pathology to cell-type-selective targeting of drugs and inhibition of undesired cell adhesion in bacterial/viral infections, inflammation, or metastasis.

1 Introduction

Basic biochemical knowledge assigns nucleic acids and proteins the decisive role in information flow in biosystems. Connected by the genetic code the transcribed portions of the genome govern the expression of a complex set of messages on the level of polypeptides. To meet the requirement of flexible regulation of product availability, the synthesis and degradation of proteins can be intimately modulated. Moreover, posttranslational modifications, phosphorylation taking a prominent place in textbooks, assure rapid and reversible fine-tuning of enzyme and receptor activities. With genome sequencing becoming routine practice, the allurement to view biological information as being exclusively epitomized by the genetic code becomes nearly irresistible. To be mindful of substance classes, which would otherwise be unfairly and incorrectly treated as "second class citizens" (von der Lieth et al., 1997b), this review furnishes information to underscore that the current judgment to place primary emphasis in research on the genetic code is unlikely to be final. After introducing first the concept of the sugar code on the level of sequence and conformation and then documenting presence of sophisticated decoding devices (among them endogenous lectins), the versatility of the sugar code will be exemplified leading to the description of perspectives to turn these discoveries into biomedical applications.

2 The Sugar Code: Basic Principles

In order to succeed as hardware for information transfer any substance class must offer the potential for specific coding. The message will have to be deciphered with sufficient biochemical affinity and low probability for ambiguities and misinterpretation. A high-density coding capacity is beneficial to keep the size of the active sections of biomolecules small, thereby reducing the energetic expenses during synthesis. Moreover, spatially easy accessibility and the potential for rapid structural modulations by reversible variations of the chain length and/or introduction of small but decisive substituents are eminent factors in the design of an efficient code system. This set of conditions describes the frame in which the quality of biological coding is to be rated. By performing such calculations on the theoretical storage capacity expressed as the total number of isomers without preconceptions it takes no effort of persuasion to convincingly show that nucleotides and amino acids are surpassed, by far, by another class of biomolecules.

Currently, carbohydrates have their main place in textbooks in chapters on energy metabolism and cell wall composition. The regular repetitive arrangement of monosaccharides in plant, insect, fungal, or bacterial cell walls or coats seduces to underestimate the other inherent talents of carbohydrates. Amazingly, they are readily discernible when closely looking at a simple structural representation (Fig. 1). Each monosaccharide offers various hydroxyl groups for oligomer formation by glycosidic bonds including the anomeric C1-position. In contrast to nucleic acids and proteins branching of chains is a common feature of the glycan part of cellular glycoconjugates (glycoproteins, glycolipids). Taking stock of the peculiarities of monosaccharide structure the total number of isomer permutation for a hexamer with an alphabet of 20 letters (monosaccharides) reaches the staggering number of 1.44×10^{15} (Laine, 1997). Under the same conditions only 6.4×10^7 (20^6) structures can be devised from 20 amino acids, the four nucleotides just yielding 4096 (4^6) hexanucleotides. Allowing two different substitutions in a hexasaccharide, occurring in Nature, e.g. as sulfation in glycosaminoglycan chains, further increases
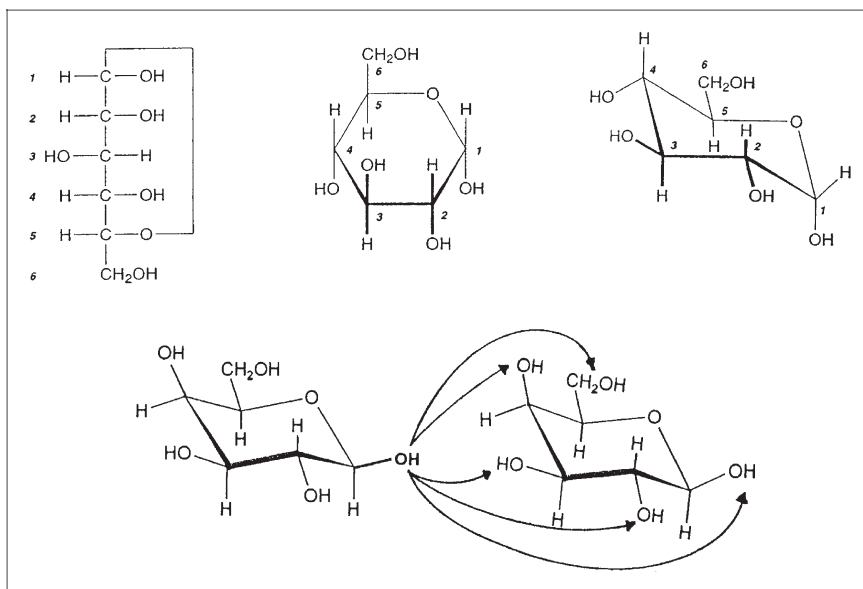

Fig. 1 The different graphic representations of the structure of a hexapyranose using α-D-glucose (Glc) as example (*top*). Commonly, the Haworth formula (*middle*) with the ring being placed perpendicular to the plane is given preference to the traditional Fischer projection of the hemiacetal (*left*). The relative positioning of the axial and equatorial substituents can readily be visualized by drawing the relatively rigid and energetically privileged chair conformation (*right*). For the formation of an acetal (disaccharide) by a glycosidic bond using D-galactose (Gal), the 4′-epimer of glucose as example, the anomeric hydroxyl group of the left monosaccharide can theoretically react with any of the five acceptors present on a second hexopyranose yielding 11 isomers with full consideration of the two anomeric positions (*bottom*). The structure of the β1–3-linked digalactoside is drawn in Fig. 3

the number of isomers by more than two orders of magnitude (Laine, 1997). In the prophetic words of Winterburn and Phelps, "carbohydrates are ideal for generating compact units with explicit informational properties, since the permutations on linkages are larger than can be achieved by amino acids, and, uniquely in biological polymers, branching is possible" (Winterburn and Phelps, 1972).

It is no treading on thin ice to follow the authors to their conclusion that "the significance of the glycosyl residues is to impart a discrete recognitional role on the protein" (Winterburn and Phelps, 1972), and it is not surprising that at least 1.0% of the translated genome in animals is devoted to the generation of code words with as many as 70% of proteins harboring the tripeptide sequon for N-glycosylation (Reuter and Gabius, 1999; Varki and Marth, 1995; Wormald and Dwek, 1999). The core region and complex extensions of this ubiquitous type of protein glycosylation in eukaryotes are shown in Fig. 2. It gives a graphic example how branching sets in and how to read the sugar code. Each linkage is characterized by the anomeric configuration and the positions of the two linkage points, such as β1-4 as opposed to α1-4 or α1-3. Since nucleotide sugars are employed as donors by the glycosyltransferases (Brockhausen and Schachter, 1997; Sears and Wong, 1998), chain growth generally involves the anomeric position restricting the range of products by enzymatic synthesis in relation to all theoretically possible isomers. Nonetheless, the presented staggering complexity of glycan structures has already placed severe obstacles in the way to go beyond merely acknowledging the enormous potential for structural variability towards precise structure determination.

These problems have mainly been solved by the development of sophisticated isolation and analysis methods combining the power of liquid chromatography, capillary zone electrophoresis, mass spectrometry, and NMR spectrometry with that of biochemical reagents such as endo- and exoglycosidases and sugar receptors (Cummings, 1997; Geyer and Geyer, 1998; Hounsell, 1997; Reuter and Gabius, 1999). Application of these techniques has revealed that subtle variations and
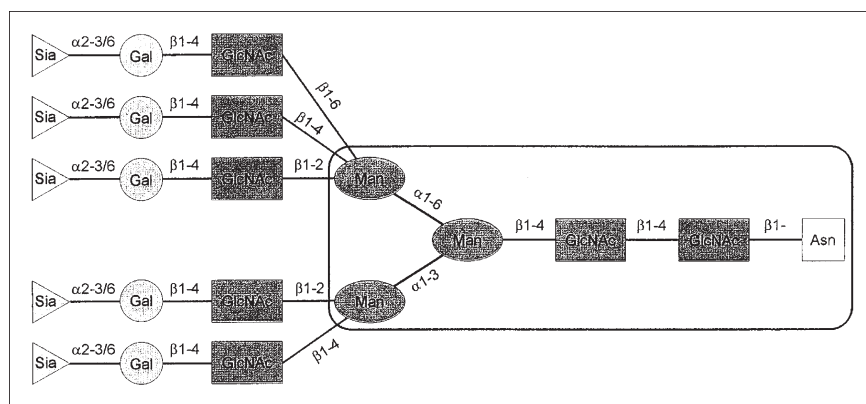

Fig. 2 Structure of the core pentasaccharide of N-glycans given within the frame and the additional branching yielding a penta-antennary complex-type sugar structure (*left*)

modifications are especially frequent in the terminal, spatially accessible sections of the sugar antennae. The strategic placement of distinctive substitutions is expected for a role in information transfer. They are marked by introduction of small substituents (sulfate and *O*-acetyl groups, etc.) into sugar moieties such as *N*-acetylgalactosamine or *N*-acetylneuraminic acid, comparable to the formation of an umlaut in the German language, or by directing a synthetic intermediate to various end products by mutually exclusive refinements, e.g. α1-3 fucosylation, α2-3/6 sialylation, and 4-sulfation (Hooper et al., 1997; Reuter and Gabius, 1996, 1999; Reutter et al., 1997; Sharon and Lis, 1997; Varki, 1996). Intercellular and temporal flexibility turns the available letter repertoire into an array of alternative structures (biosignals). Indeed, the observations that the profile of glycans is not genetically strictly coded but influenced by the presence and relative positioning of the set of enzymes in the assembly line and the actual availability of activated substrates such as nucleotide donors argues in favor of purpose versus randomness (Abeijon et al., 1997; Pavelka, 1997; Varki, 1998). Thus, the prerequisite for rapid and multifarious modulation mentioned in the introductory paragraph is adequately fulfilled in the sugar code.

In view of the assumed importance for maintaining diversity, a multicellular organism with lack of presence of one of the mentioned pathways will allow to probe into the question whether this deficit is accompanied by any remodeling in the overall glycosylation system or not. Assisted by genome sequencing, it can indeed be proposed that absence of sialylation in the nematode *Caenorhabditis elegans* might be compensated by elaboration of another part of the enzymatic machinery. The discovery of 18 different genes for putative fucosyltransferases in the genome of this nematode argues in favor of this notion (Oriol et al., 1999). In these authors' own words, "for some unknown reasons, these nematodes have favored through evolution fucosylation instead of sialylation of their terminal nonreducing oligosaccharide epitopes or glycotopes and since sialic acid and fucose are usually in competition for the same acceptors, the lack of all forms of sialic acid in *C. elegans* fits well with a large expression of different fucosyltransferase genes, making this animal an ideal model for evolutionary studies of fucosyltransferases" (Oriol et al., 1999). All these reactions in glycosylation result in a typical pattern of glycan chains on the level of cells and organs. It is as characteristic as a fingerprint or a signature. Yeast cells, for example, produce mannose-rich surface glycans, while multicellular organisms prominently put histo-blood group epitope-rich complex-type glycans on display. Enzymes for these extensions at the end of antennae (Fig. 2) typically reside in the medial- and *trans*-Golgi regions. Since the number of activities operating upon these sections has especially expanded in the animal kingdom, it is rather unreasonable to assume these refinements to have survived fortuitously. Driving this evolutionary process can be attributed to functions of the glycans ranging from purely physical aspects such as solubility or protection of surface against proteolytic attack to any involvement in recognition (Drickamer and Taylor, 1998; Gagneux and Varki, 1999; Reuter and Gabius, 1999; Sharon and Lis, 1997; Varki, 1996).

A principal comment is warranted on the surmised evolutionary mechanisms of selection of letters for the alphabet of this code system. As insightfully discussed by Hirabayashi (1996), elementary hexose synthesis under prebiotic conditions was most probably facilitated by the following cascade. It started with formol condensation, yielding basic trioses known from glycolysis. The next step is the aldol condensation to 3,4-*trans*-ketoses and a conversion of D-fructose to D-glucose and D-mannose via an enediol-intermediate and the keto-enol tautomerism (Lobry de Bruyn rearrangement). Notably, D-glucose harbors no 1,3-diaxial interactions involving a hydroxyl group (Fig. 1), and the favored "tridymite" water structure is maintained in the presence of equatorial hydroxyl groups (Uedaira and Uedaira, 1985). In mannose as in galactose, a biochemical derivative obtained by the NAD$^+$-dependent epimerization of glucose, only one hydroxyl group is axial, keeping unfavorable 1,3-diaxial interactions and perturbation of solvent structure minimal. In contrast to the 2′- and 4′-epimers, the 3′-epimer has two 1,3-axial interactions. Origin from synthesis under prebiotic conditions and energetic consequences entail the organization of the initial hardware of the sugar code. From them, further letters of the alphabet comprising also the *N*-acetyl derivatives of the 2′-amines of glucose and galactose, L-fucose, D-xylose, and *N*-acetylneuraminic acid are biosynthetically produced. Interestingly, the core section of *N*-glycans (Fig. 2) is composed of basic units derived from a presumably prebiotic origin. This fact invites to speculate on a relationship of evolutionary pathways on the levels of eukaryotic organisms and of glycan complexity. Setting this aspect which is further discussed elsewhere (Drickamer and Taylor, 1998; Gagneux and Varki, 1999; Hirabayashi, 1996; Oriol et al., 1999) aside in this context, it can at least be reliably concluded at this stage that oligosaccharides by their inherent potential for ample sequence permutations including variations in the anomeric position and the linkage groups for a glycosidic bond deserve attention as coding units. Remarkably, recent work extends the capacity for information storage from two dimensions of linear and branched oligosaccharide chains to the third dimension.

3   The Sugar Code: The Third Dimension

The shape of a glycan will be determined by the conformation of the furanose/pyranose rings and the relative positioning of the rings in the chain. Based on x-ray crystallography, neutron diffraction, and homonuclear coupling constant data the $^4C_1$ chair conformer ($^1C_4$ for L-sugars) is the energetically preferred pyranose ring structure (Abeygunawardana and Bush, 1993; Brown and Levy, 1965). In rare cases, for example, for L-iduronic acid as constituent of heparan and dermatan sulfates, and to accommodate mechanical stress, conformational flexibility and elasticity of a pyranose can be generated by chair-boat transitions, which allow L-iduronic acid to acquire the skew-boat form ^2S_o (Casu et al., 1988; Marszalek et al., 1998). Yet the main contribution to define a glycan's shape will generally originate not
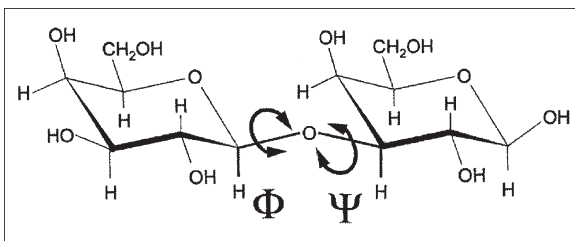
Fig. 3 Depiction of the main source of conformational flexibility of the disaccharide Galβ1-3Gal (see Fig. 1) by independent rotations about the two dihedral angles ϕ and ψ of the glycosidic bond

from this source. In contrast, it will arise from changes of the two dihedral angles ϕ and ψ of each glycosidic bond (Fig. 3).

By letting the thumbs of each hand touch, independent variations of these two parameters by movements of the hands can swiftly be visualized. Since the pyranose rings linked by the glycosidic bond and their exocyclic substituents are rather bulky, their size will impose topological restraints to the intramolecular movements of the oligomer. Compared to oligopeptides with small side chains, the conformational space accessible to the molecule at room temperature will thus be relatively restricted. That this spatial factor limits the range of interchangeable conformations has been inferred by computer-assisted molecular mechanics and dynamics calculations and convincingly documented by experimental evidence primarily from sophisticated NMR-spectroscopy (Bush et al., 1999; Imberty, 1997; Siebert et al., 1999; von der Lieth et al., 1997a, 1998; Woods, 1998). Exploring the actual position(s) of each oligosaccharide on the scale between high flexibility with an ensemble of conformers and almost complete rigidity will definitely have salient implications to predict its role as coding unit. In this respect, it is also worth pointing out that a notable level of intramolecular flexibility is not a favorable factor for crystallization. Indeed, such an extent of unrestrained conformational entropy can contribute to explain the frequently frustrating experience in respective attempts in carbohydrate chemistry. If on the other hand the level of conformational entropy is confined to only very few stable conformers (keys), the presented shape distribution is not only a function of the sequence but also of external factors affecting the actual status of the equilibrium. In this context it should not escape notice that environmental parameters with impact on presentation of the glycan in glycoconjugates might shift the dynamic equilibrium of shape interconversions between attainable positions without requirement to alter the primary structure. Sugar receptors as probes for distinguishing bioactive or bioinert glycan presentation modes on proteins have already given the hypothesis experimental credit (Mann and Waterman, 1998; Noorman et al., 1998; Solís et al., 1987; White et al., 1997). This support brings to view an attractive means to modify shape, which warrants contriving further appropriate experiments to underpin its actual operation beyond any doubt.

As implied by referring to a code system, information stored as sequence and shape will have to be grasped. Translating and transmitting it into intended responses is the task of decoding devices. They should specifically recognize coding units established by glycans. Thus, in addition to physicochemically serving roles to control folding, oligomerization and access of proteolytic enzymes, as already mentioned (Drickamer and Taylor, 1998; Gagneux and Varki, 1999; Reuter and Gabius, 1999; Sharon and Lis, 1997; Varki, 1996), oligosaccharides in glycan chains can be likened to the postal code in an address to convey distinct messages read by suitable receptors. These carbohydrate-binding proteins are classified into enzymes responsible to assemble, modify, and degrade sugar structures, immunoglobulins homing in on carbohydrates as antigens, and, last but not least, lectins. Evidently, the third class encompasses all carbohydrate-binding proteins, which are neither antibodies nor are they enzymes which couple ligand recognition with catalytic activity to process the target (Barondes, 1988; Gabius, 1994). That lectin/glycan recognition has been assigned pivotal duties in an organism can at best be rendered perceptible by aberrations causing diseases. Knowledge accrued from the study of the biochemical basis of human diseases (e.g. mucolipidosis II or leukocyte adhesion deficiency (LAD) type II syndrome) underscores how trafficking of lysosomal enzymes or leukocytes can go awry owing to a lack of generation of the essential carbohydrate signal (Brockhausen et al., 1998; Lee and Lee, 1996; Paulson, 1996; Reuter and Gabius, 1999; Schachter, 1999; von Figura, 1990).

3.1   Lectins: Translators of the Sugar Code

The concept of a recognitive interplay between a sugar ligand and a lectin readily receives support, when the assumed ligand properties can be ascertained. As compiled in Table 1, various experimental approaches exploit the lectin's binding specificity in assays for their detection and characterization. The success in establishing these techniques and the power of affinity chromatography together with expression cloning and homology searches have spurred the transition from the early phase to categorize lectins according to their monosaccharide specificity and requirement for cations to the era to draw genealogical trees of lectin families. Having its roots in the structural definition of the folding pattern and architecture of the carbohydrate recognition domain, the classification scheme is currently agreed upon with five distinct families of animal lectins, i.e. C-type lectins, galectins, I-type lectin, P-type lectins, and pentraxins (Drickamer, 1988, 1993; Gabius, 1997a; Powell and Varki, 1995; Rini and Lobsanov, 1999). That this compilation is unlikely to be final is implied by the description of lectin sequences lacking invariant characteristics of any of the five classes (e.g. the chaperones calnexin and calreticulin mentioned in Table 2) and the detection of new folding arrangements (e.g. the five-bladed β-propeller in the invertebrate lectin tachylectin-2 (Beisel et al., 1999)).

Table 1 Methods used in the search for lectins. (Modified from Gabius, 1997a.)

Tools	Parameter
Multivalent glycans and (neo)glyco-conjugates or defined cell populations	Carbohydrate-dependent inhibition of lectin-mediated glycan precipitation or cell agglutination
Labelled (neo)glycoconjugates and matrix-immobilized extract fractions or purified proteins	Signal intensity
Cell populations	Labeling intensity
Tissue sections	Staining intensity
Animal (neo)glycoconjugatedrug chimera and cell populations	Biodistribution of signal intensity cellular responses (cell viability etc.)
Matrix-immobilized (neo)glycoconju-gates and cell populations	Carbohydrate-inhibitable cell adhesion
Cell extracts	Carbohydrate-elutable proteins
Homology searches with computer programs (e.g. Gene-finder or Blast), expressed sequence tags and knowledge of key structural aspects of carbohydrate recognition domains	Homology score in sequence alignment or knowledge-based modeling
Lectin motif-reactive probe (antibody, primer sets)	Extent of cross-reactivity

In each lectin family sequence alignments and homology searches have so far been conducive to unravel the divergent pathway from an ancestral gene to the current diversity. The intrafamily genealogy of mammalian C-type lectins has elegantly been traced back in a dendrogram to common ancestors for the seven subfamilies (Drickamer, 1993). To illustrate that such domains, often a part of modular arrangements, are no rare peculiarity in animal genomes, it is telling to add that a current database lists 389 C-type lectin-like sequences in animals (Sonnhammer et al., 1998). Yeast lacks this module in its domain collection. In the nematode *C. elegans*, whose elaborate enzymatic system for fucosylation has already been referred to (Oriol et al., 1999), this domain is ranked on the seventh place in frequency of occurrence, excelling for example the abundance of the EGF-like motif (The *C. elegans* Sequencing Consortium, 1998). At present, 183 C-type lectin-like domains have been traced in 125 proteins (Drickamer and Dodd, 1999). However, it is presently unclear how many of these proteins will be actually operative in Ca^{2+}-dependent sugar (or peptide) binding (Drickamer, 1999). Also, at least eight functional galectin genes and a tentative total of 28 candidate galectin genes among the approximately 20,000 genetic reading frames (current number predicted: 19,099) in its genome were identified in the nematode (Hirabayashi et al., 1997; Cooper and Barondes, 1999). These new insights into lectin abundance further increase the percentage of the coding genome devoted to glycan production and recognition.

Table 2 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, VIP-36, P-type lectins, comitin
Intracellular transport and extracellular assembly	Non-integrin 67 kDa elastin/laminin-binding protein
Cell type-specific endocytosis	Hepatic asialoglycoprotein receptor, macrophage C-type lectins, hepatic endothelial cell receptor for GalNAc-4-SO$_4$-bearing glycoproteins
Recognition of foreign glycans (β1,3-glucans, LPS)	CR3 (CD11b/CD18), *Limulus* coagulation factors C and G
Recognition of foreign or aberrant glycosignatures on cells (incl. endocytosis or initiation of opsonization or complement activation)	Collectins, C-type macrophage receptors, pentraxins (CRP, limulin), L-ficolin, tachylectins
Targeting of enzymatic activity in multimodular proteins	Acrosin, *Limulus* coagulation factor C
Bridging of molecules	Homodimeric and tandem-repeat galectins, cytokines (e.g. IL-2:IL-2R and CD3 of TCR), cerebellar soluble lectin
Effector release (H$_2$O$_2$, cytokines etc.)	Galectins, selectins, CD23
Cell growth control and apoptosis	Galectins, C-type lectins, amphoterin-like protein, cerebellar soluble lectin
Cell routing	Selectins, I-type lectins, galectins
Cell–cell interactions	Selectins and other C-type lectins, galectins, I-type lectins
Cell–matrix interactions	Galectins, heparin- and hyaluronic acid-binding lectins
Matrix network assembly	Proteoglycan core proteins (C-type CRD), galectins, non-integrin 67 kDa elastin/laminin-binding protein

For further information, see Gabius (1997a), Gabius and Gabius (1993, 1997), Kaltner and Stierstorfer (1998), Kishore et al. (1997), Vasta et al. (1999), Zanetta (1998) for recent reviews.

Equaling the strides being taken in the structural research on lectins, elucidation of their in vivo significance has steadily moved forward in the last decade. Summarized in Table 2, our present status of knowledge bears witness to the versatility to ply glycan recognition for a variety of purposes. In addition to mediating a physical contact between molecules and cells their initial recognition can trigger postbinding signaling with impact, for example, on growth regulation (Villalobo and Gabius, 1998). With focus on the homodimeric galectin-1 its mediation of downregulation of cell growth of responsive human neuroblastoma cells and of T-cell apoptosis to alleviate collagen-induced arthritis depicts representative examples with potential clinical relevance (Kopitz et al., 1998; Rabinovich et al., 1999).

Albeit necessarily centered in basic science, such cases illustrate the conceivable future potential to turn an endogenous lectin into a pharmaceutical.

Having already moved closer to applied science, the participation of lectins and glycoconjugates in cell adhesion has prompted attempts to rationally interfere with the molecular rendezvous, conceptually visualized as anti-adhesion therapy in Fig. 4. This approach mimics the natural strategy for success achieved with the complex cocktail of milk glycoconjugates.

They are protective by blocking docking of pathogens such as enteropathogenic and hemorrhagic *Escherichia coli, Campylobacter jejuni*, or rotavirus (Newburg, 1999). Although realization of this approach can prove tedious, because the pattern of recognition pairs is often not restricted to very few lectins (*Helicobacter pylori* with at least ten different carbohydrate-binding activities compared to the single type of influenza sialidase whose inhibition will noticeably affect virus propagation (Karlsson, 1999; Lingwood, 1998; von Itzstein and Thomson, 1997)), the custom-made design of tools, drawn as symbols in the strategy-outlining Fig. 4, justifies efforts to first localize binding partners and then to interfere with their activity aimed at therapy.

Notably, the first method can be used independently, e.g. in diagnostic procedures to characterize cell features. The visualization of carbohydrate-specific activities is commonly performed with carrier-immobilized sugar structures. Covalent attachment of a suitable derivative furnishes the versatility to produce neoglycoconjugates tailored to the experimental requirements (Bovin and Gabius, 1995; Lee Y.C. and Lee, 1994). Compared to a single carbohydrate unit the affinity of the multivalent ligand "is often beyond that expected from the increase in sugar concentration due to the presence of multiple residues on the protein (or polymeric backbone). Such an affinity enhancement is termed the glycoside cluster effect" (Lee R.T. and Lee, 1994). The geometrical increase in affinity with a numerical increase in valence for
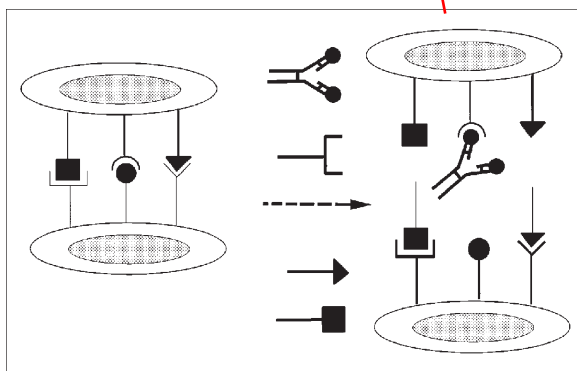

Fig. 4 Interference in lectin-mediated cell contact formation or recognition processes in general with target-specific blocking reagents, i.e. antibodies, sugar receptors, and oligosaccharides or mimetics thereof. (Kindly provided by Priv.-Doz. Dr. H. Kaltner, Munich. Details on the current status of anti-adhesion therapy are given by Cornejo et al., 1997; Gabius, 1997b; Gabius and Gabius, 1997; Karlsson, 1998; Simon, 1996; Zopf and Roth, 1996.)

mono-, bi-, and trivalent Gal-terminated oligosaccharides and mammalian asialoglycoprotein receptor, a C-type lectin, has been attributed to the topological complementarity between multiple ligand and receptor sites (Lee and Lee, 1997). Membrane solubilization by detergent treatment will in this case disrupt the essential spatial arrangement. An important caveat for approaches to detect the cluster effect concerns the use of agglutination assays. In contrast to affinity measurements in direct binding assays, the ongoing aggregation of multivalent receptors and ligands in solution can lead to erroneous conclusions. Indeed, under these circumstances isothermal titration calorimetry failed to record enhancements of Gibbs' free energy of binding but measured an endothermic, entropically favored process, its extent correlating with the inhibitory potency (IC_{50}-values) of tetra- and hexavalent ligands (Dimick et al., 1999).

Adding a label to the neoglycoconjugates renders them serviceable for detection of ligand-specific sites in cells and tissues, as listed in Table 1 with special practical emphasis being currently placed in tumor diagnosis (Gabius et al. 1995; Danguy et al., 1998; Gabius et al., 1998; Kayser and Gabius, 1999). In view of common lectin histochemistry with plant agglutinins, this method has been designated as "reverse lectin histochemistry" (Gabius et al., 1993). Following the description of a relevant clinical correlation, e.g. binding of histo-blood group A- and H-trisaccharides to lung cancer cells and survival of patients (Kayser et al., 1994), further work will aim to define the tissue target and to refine the ligand for optimal selectivity and specificity (Mammen et al., 1998) en route to running assays to unveil, if possible, therapeutic benefit in lectin-directed anti-adhesion therapy (see references given in legend for Fig. 4) and drug targeting (Gabius, 1989, 1997b). To attain this objective, it is indispensable to comprehend the how and why of protein–carbohydrate recognition. Thus, it is inostructive to proceed with a brief outline of these principles relevant for drug design.

4   Principles of Protein–Carbohydrate Recognition

Basically, typical contributions to the Gibbs' free energy of ligand binding originate from hydrogen bonding, van der Waals forces and the consequences of the hydrophobic effect. Factors to be reckoned with to predict the affinity of a ligand further include any alterations of the geometry and motional dynamics of the receptor and/or the ligand and/or the solvent molecules. As experimentally readily accessible parameters by calorimetric techniques, the determination of the reaction enthalpy and entropy delineates the global driving force towards complex formation. These parameters have, for example, been measured for an array of mono- and disaccharides in the cases of a plant and an animal lectin sharing specificity to D-galactose (Bharadwaj et al., 1999), and the plot of the data (Fig. 5) according to the equation:
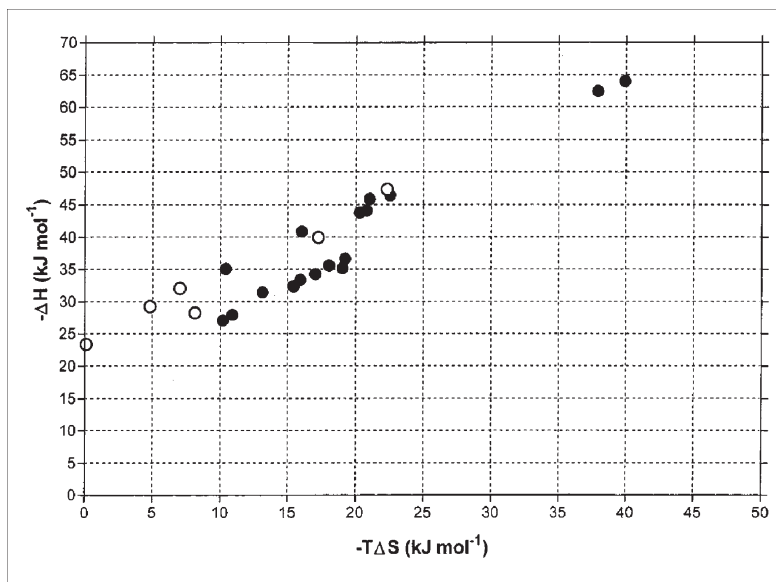
$$-\Delta H = -\Delta G - T\Delta S$$

Fig. 5 Enthalpy–entropy compensation plot for the binding of a panel of mono- and disaccharides to the galactoside-specific mistletoe lectin (•) and the galectin from adult chicken liver (o)

reveals a slope near unity and intercepts of −16.45 kJ/mol (plant lectin) and −23.12 kJ/mol (animal lectin).

This figure conveys a fundamental message on the relation between enthalpic and entropic factors attributed to the participation of weak intermolecular forces. An increase in enthalpy for ligand binding is inherently balanced by an entropic penalty (or vice versa), an obvious example of common enthalpy–entropy compensation (Dunitz, 1995; Gilli et al., 1994; Lumry and Rajender, 1970). Its illustration automatically poses an ambitious question. The major challenge is to assign events on the level of the molecules in the course of association to the global enthalpic and entropic factors to bridge the gap between the demand for rules to optimize shape recognition and the thermodynamics. With this knowledge in hand, it might be feasible to intentionally shift the specificity and selectivity of derivatives. As the controversial discussion on the positive or negative role of water molecules for the enthalpy of complexation illuminates (Gabius, 1998; García-Hernández and Hernández-Arana, 1999; Lemieux, 1996; Toone, 1994), it will be essential to scrutinize the behavior of each participant of the molecular rendezvous in detail. Consequently, quick complete answers should not be expected but stepwise advances by the combination of computer-assisted calculations, spectroscopic techniques in solution, chemical tinkering with the ligand structure towards potent mimetics, and x-ray crystallography. An impression into the practical implementation of this interdisciplinary approach is given in the next paragraph.

5   How to Define Potent Ligand Mimetics

Taking the meaning of the word "carbohydrate" (C $(H_2O)_n$) literally, the abundant display of hydroxyl groups with their sp^3-hybridized oxygen atoms acting as acceptors with two lone electron pairs and the protons as donors nourishes the view that hydrogen bonds will dominate the spectrum of binding forces. When the spacing between two hydroxyl groups or the axial 4′-hydroxyl group and the ring oxygen atom matches that of an amino acid side chain (amide or carboxylate), two neighboring sites on the ligand can well be engaged in bidentate hydrogen bonds. The necessity for topological complementarity to yield the intricate network, schematically shown in Fig. 6, may not only be a source for enthalpy but also for selectivity, distinguishing anomers such D-Gal versus D-Man/D-Glc. It can thus be expected that the axial 4′-position for recognition of D-Gal and the equatorial 3′,4′-positions for binding D-Man/D-Glc play decisive roles. This assumption is strikingly verified by x-ray crystallography and in solution by chemically engineered ligand derivatives (Rini, 1995; Solís et al., 1996; Weis and Drickamer, 1996; Gabius, 1997a, Solís and Díaz-Mauriño, 1997; Gabius, 1998; Lis and Sharon, 1998; Loris et al., 1998; Rüdiger et al., 1999;). With this structural explanation it becomes obvious why the change of the position of one hydroxyl group to form an epimer discussed during the presentation of the individual members of the mon-osaccharide alphabet unmistakably has the effect of creating a new letter. By the way, the same principle holds true for the characteristic formation of two coordination
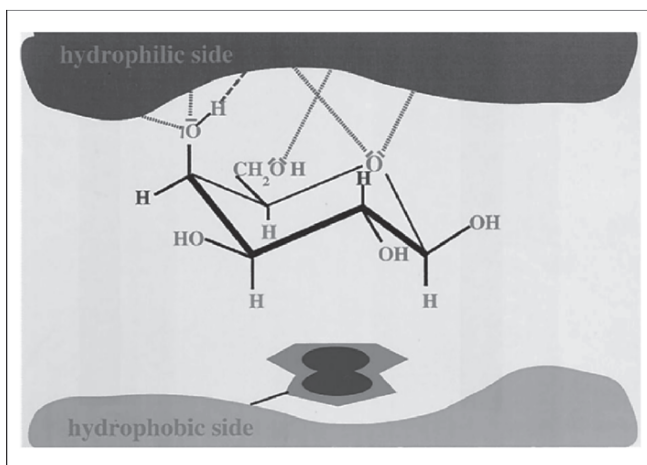

Fig. 6   The potential of D-galactose (see Fig. 1, bottom, and Fig. 3) for establishing interactions with constituents of the binding pocket of a sugar receptor. While the rather polar upper side can be engaged in frequent hydrogen bonds exploiting lone electron pairs of oxygen atoms as acceptors and the protons of appropriately positioned hydroxyl groups as donors (and also coordination bonds with a Ca^{2+}-ion in the case of C-type lectins), C-H/π-electron interactions and entropically favorable stacking can be engendered by an intimate contact of an aromatic (here: indolyl) amino acid side chain and the sugar's less polar bottom section. (Kindly provided by Dr. C.-W. von der Lieth, Heidelberg.)

bonds with the central Ca^{2+}-ion in the mentioned C-type lectins. Thereby, any wrong combination for the two adjacent hydroxyl groups involved in contacting the metal ion is excluded and sugar specificity is assured, unless the access-restricting impediment by a constraining loop close to the metal ion is lifted (Weis and Drickamer, 1996; Gabius, 1997a; Lis and Sharon, 1998; Loukas et al., 1999).

Inspecting Fig. 6 more closely, another important feature to drive ligand binding can be discovered. While the upper side of D-Gal is rather polar, the B-face exhibits a hydrophobic character. Stacking to the bulky aromatic amino acid side chain in the binding pocket removes both nonpolar surfaces from solvent accessibility, although the two rings may not be perfectly aligned in parallel. In fact, their positioning can tolerate distortions with angles between 17° and 52° in lectins (Weis and Drickamer, 1996). Nonetheless, this alignment will still contribute to complex stability and also to ligand selection despite its lower degree of directionality relative to hydrogen bonds (Quiocho, 1988; Vyas, 1991). The ensuing shielding of the indolyl side chain by the ligand is reflected for galectins in molecular dynamics calculations as well as differential UV, fluorescence, and laser photo chemically induced dynamic nuclear polarization (CIDNP) spectra (Levi and Teichberg, 1981; Siebert et al., 1997). Beyond this impact on solvent molecules by reducing the apolar surface area the π-electron cloud of the aromatic ring is likely to interact with the aliphatic D-Gal protons which harbor a net positive charge (Dougherty, 1996; Nishio et al., 1995; Weis and Drickamer, 1996). That the ensuing hydrophobic effect and van der Waals interactions do not deserve to be underestimated for impinging on the overall Gibb's free energy gain is underscored by the analysis of dominant forces in tight ligand binding for a variety of cases, where these factors can even surpass by far the contribution of hydrogen bonds (Davis and Teague, 1999; Kuntz et al., 1999).

This observation illustrates the complexity of the question how to account for the global enthalpic and entropic parameters on the level of molecules. For that galectin, whose data set from isothermal titration calorimetry is given in Fig. 5, it has recently been described by crystallographical work that six structural water molecules occupy the binding site in the ligand-free state stabilizing its topology and yielding a not yet precisely quantitated contribution to the Gibbs' free energy change upon displacement (Varela et al., 1999). In the case of a related galectin from the conger eel one additional water molecule even takes place of D-Gal's B-face substituting stacking by forming a π-electron hydrogen bond with a distance of 3.36 Å and an angle of 6.5° between the vector of the weight center of the five-membered section of the indole ring to the water molecule and the vector perpendicular to the ring plane (Shirai et al., 1999). The total exchange of the water molecules with the ligand will not only directly affect these solvent molecules but may also have a bearing on the proteins' intramolecular motions in solution. Remarkably, also the impact of ligand binding on protein flexibility is to be reckoned with. An increase in its vibrational entropy (14.6 kJ/mol for binding of one water molecule to bovine pancreatic trypsin inhibitor as model (Fischer and Verma, 1999)) can offset a substantial portion of the entropic penalty of the immobilization. The extent of this factor will certainly depend on the inherent mobility dynamics of the carbohydrate ligand free in solution. This parameter has already been inferred above to be often restricted due to spatial interference of the rather bulky rings and

substituents. Graphically drawing on E. Fischer's (1894) classical "lock and key" paradigm, the metaphor has tentatively been introduced for this ligand type to view certain oligosaccharides as "bunch of keys" moving in solution through a limited set of shapes (Hardy, 1997). Only one of them may be selected by a receptor.

With a digalactoside (Galβ1-2Gal) as model, the formation of two "keys" from the same sequence is displayed in Fig. 7. Based on the ϕ, ψ, E-plot, shown in its left section, molecular dynamics calculations and nuclear Overhauser effect (NOE) NMR-spectroscopy (Siebert et al., 1996, 1999; von der Lieth et al., 1998), two distinct conformations are present in solution, each molecule rapidly fluctuating between these two topological constellations (Fig. 7, right side). Due to the inability to acquire spectroscopic snapshots with a resolution in the ps range, spectroscopical monitoring will be subject of time and ensemble averaging (Carver, 1991; Jardetzky, 1980). Since the term "key" implies its accurate fit into an appropriate lock, monitoring of transferred NOE signals, reflecting through space dipolar interactions between two protons in the bound ligand in double-resonance experiments, will resolve the gripping question as to which ligand topology will be accommodated in the binding pocket (Gabius, 1998; Jiménez-Barbero et al., 1999; Peters and Pinto, 1996; Poveda and Jiménez-Barbero, 1998; Rüdiger et al., 1999; von der Lieth et al., 1998).
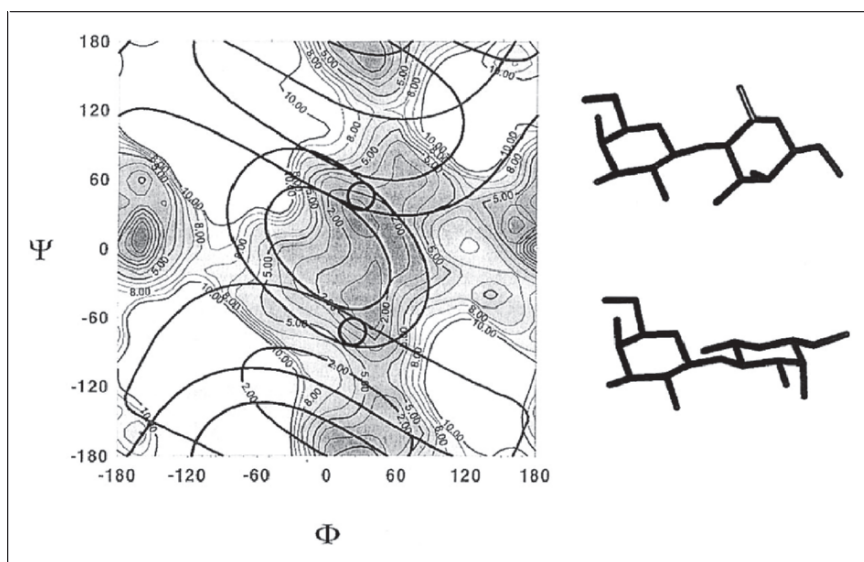

Fig. 7 Illustration of the principle of differential conformer selection. Based on NMR-spectroscopical analysis and molecular mechanics/dynamics calculations the disaccharide Galβ1-2Gal can adopt two distinct conformations in solution, which reside in energetically preferred regions of the ϕ, ψ, E-plot, symbolized by circles (*left*). Keeping the topological positioning of the nonreducing Gal-unit constant, the two sets of ϕ, ψ-values are readily visualized to translate into two significantly different conformers (*right*) which harbor disparate ligand properties. (Kindly provided by Priv.-Doz. Dr. H.-C. Siebert, Munich, and Dr. C.-W. von der Lieth, Heidelberg.)

These experiments provide two captivating answers for the studied case of lectins. Firstly, a lectin can actually select a distinct conformer, as seen for galactoside-binding lectins and selectins (Asensio et al., 1999; Espinosa et al., 1996; Gilleron et al., 1998; Harris et al., 1999; Poppe et al., 1997; Siebert et al., 1996; von der Lieth et al., 1998). Despite the same sequence the shape of other conformers renders them unsuitable for binding. Of course, a wrong key will not open a nonadaptable (rigid) lock designed for a different shape. Secondly, different receptors even with the same saccharide specificity harbor the capacity to bind different conformers. Thus, freezing a distinct conformation should have a dramatic impact on receptor binding as alluded to above. This principle is referred to as "differential conformer selection." It is visualized in Fig. 7 by noting that the conformer defined by the upper ϕ, ψ-combination is exclusively bound by a plant (mistletoe) agglutinin, while the tested galectin homes in on the second conformer (Gabius, 1998; Gilleron et al., 1998; Siebert et al., 1996; von der Lieth et al., 1998). Thus, not only the hydrogen-bonding patterns of these lectins toward D-Gal differ, as delineated by chemical mapping with deoxy and fluoro derivatives (Lee et al., 1992; Rüdiger et al., 1999; Solís et al., 1996), but also the pair of ϕ, ψ-torsion angles of β-Gal-terminated disaccharides. Because the importance of the intramolecular flexibility of the free ligand and conformer selection is only gradually explored as factor to be rationally manipulated, this result together with insights into favorable energetic interactions between the binding partners including solvent molecules warrants consideration for the design of mimetics. Thereby, they can eventually meet the high expectations for potency expressed in Fig. 4. When the geometry of crucial groups is maintained or even improved, the obtained substances do not even need to belong to the class of carbohydrates. To grant adequate heed to mimetics is probably a means to open a wide field for rational drug design, currently, for example, exploited for the influenza A/B neuraminidase and selectins (Sears and Wong, 1999; Simanek et al., 1998; von Itzstein and Thomson, 1997). As caveats to caution against prematurely advocating clinical effectiveness of anti-adhesion therapy in inflammation or of sugar-based drugs in epidemic flu, detrimental long-term effects in an animal model mimicking both acute and chronic intestinal inflammation has been reported (McCafferty et al., 1999). Similarly, stress has been laid upon the necessity to prove clinical benefit for an elegantly invented but costly antiflu drug in terms of an obvious impact on mortality beyond that of common, less expensive medications including vaccination (Cox and Hughes, 1999; Institut für Arzneimittelinformation, 1999; Yamey, 1999).

6   Conclusions

Elucidation of the structural basis of the genetic code and its translation into peptide sequences with milestones set by J.D. Watson and F. Crick (1953) and M.W. Nirenberg and J.H. Matthei (1961) has paved the way for medical applications

more than three decades after the pioneering work in basic science. Today, nearly 20% of the new drugs tested in final phases are based on the technology of genetic engineering, up from 12% last year. To fathom the intricacy of the sugar code and transfer this knowledge on how sugar words are formed and these messages are decoded by receptors to applied science should therefore not be anticipated to be a matter of only a few years. The versatility of exploiting oligosaccharides as carriers of biological information presented by nature should first be thoroughly unraveled. Building this solid basis of experimental data will most likely entail to be able to venture into newly defined areas of glycan functionality and then to launch further projects of interdisciplinary research leading from basic to applied science. In view of the current focus on genomics, the presented evidence and reasoning offer well-grounded arguments to shake the conviction that as source for understanding of recognitive and regulatory processes in normal and disease states fruitful work will solely be confined to handling the data bank of the human genome.

References

Abeijon C, Mandon EC, Hirschberg CB (1997) Transporters of nucleotide sugars, nucleotide sulfate and ATP in the Golgi apparatus. Trends Biochem Sci 22:203–207

Abeygunawardana C, Bush CA (1993) Determination of the chemical structure of complex polysaccharides by heteronuclear NMR spectroscopy. Adv Biophys Chem 3:199–249

Asensio JL, Espinosa JF, Dietrich H, Cañada FJ, Schmidt RR, Martín-Lomas M, André S, Gabius H-J, Jiménez-Barbero J (1999) Bovine heart galectin-1 selects a distinct *(syn)* conformation of C-lactose, a flexible lactose analogue. J Am Chem Soc 121:8995–9000

Barondes SH (1988) Bifunctional properties of lectins: lectins redefined. Trends Biochem Sci 13:480–482

Beisel H-G, Kawabata S-i, Iwanaga S, Huber R, Bode W (1999) Tachylectin-2: crystal structure of a specific GlcNAc/GalNAc-binding lectin involved in the innate immunity host defense of the Japanese horseshoe crab *Tachypleus tridentatus*. EMBO J 18:2313–2322

Bharadwaj S, Kaltner H, Korchagina EY, Bovin NV, Gabius H-J, Surolia A (1999) Microcalorimetric indications for ligand binding as a function of the protein for galactoside-specific plant and avian lectins. Biochim Biophys Acta 1472:191–196

Bovin NV, Gabius H-J (1995) Polymer-immobilized carbohydrate ligands: versatile chemical tools for biochemistry and medical sciences. Chem Soc Rev 24:413–421

Brockhausen I, Schachter H (1997) Glycosyltransferases involved in N- and O-glycan biosynthesis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp 79–113

Brockhausen I, Schutzbach J, Kuhns W (1998) Glycoproteins and their relationship to human disease. Acta Anat 161:36–78

Brown GM, Levy HA (1965) a-D-glucose: precise determination of crystal and molecular structure by neutron-diffraction analysis. Science 147:1038–1039

Bush CA, Martin-Pastor M, Imberty A (1999) Structure and conformation of complex carbohydrates of glycoproteins, glycolipids, and bacterial polysaccharides. Annu Rev Biophys Biomol Struct 28:269–293

Carver JP (1991) Experimental structure determination of oligosaccharides. Curr Opin Struct Biol 1:716–720

Casu B, Petitou M, Provasoli M, Sinay P (1988) Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans. Trends Biochem Sci 13:221–225

Cooper DNW, Barondes SH (1999) God must love galectins; He made so many of them. Glycobiology 9:979–984

Cornejo CJ, Winn RK, Harlan JM (1997) Anti-adhesion therapy. Adv Pharmacol 39:99–142

Cox NJ, Hughes JM (1999) New options for the prevention of influenza. N Engl J Med 341:1387–1388

Cummings RD (1997) Lectins as tools for glycoconjugate purification and characterization. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/ Weinheim, pp 191–199

Danguy A, Decaestecker C, Genten F, Salmon I, Kiss R (1998) Application of lectins and neogly-coconjugates in histology and pathology. Acta Anat 161:206–218

Davis AM, Teague SJ (1999) Die Bedeutung der Balance von Wasserstoffbrückenbindungen und hydrophoben Wechselwirkungen im Wirkstoff-Rezeptor-Komplex. Angew Chem 111:778–792

Dimick SM, Powell SC, McMahon SA, Moothoo DM, Naismith JH, Toone EJ (1999) On the meaning of affinity: cluster glycoside effects and concanavalin A. J Am Chem Soc 121:10286–10296

Dougherty DA (1996) Cation-p interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science 271:163–168

Drickamer K (1988) Two distinct classes of carbohydrate recognition domains in animal lectins. J Biol Chem 263:9557–9560

Drickamer K (1993) Evolution of Ca^{2+}-dependent animal lectins. Progr Nucl Acid Res Mol Biol 45:207–233

Drickamer K (1999) C-type lectin-like domains. Curr Opin Struct Biol 9:585–590

Drickamer K, Dodd RB (1999) C-type lectin-like domains in *Caenorhabditis elegans*: predictions form the complete genome sequence. Glycobiology 9:1357–1369

Drickamer K, Taylor ME (1998) Evolving views of protein glycosylation. Trends Biochem Sci 23:321–324

Dunitz JD (1995) Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. Chem Biol 2:709–712

Espinosa JF, Cañada FJ, Asensio JL, Martin-Pastor M, Dietrich H, Martin-Lomas M, Schmidt RR, Jiménez-Barbero J (1996) Experimental evidence of conformational differences between C-glycosides and O-glycosides in solution and in the protein-bound state: the C-lactose/O-lactose case. J Am Chem Soc 118:10862–10871

Fischer E (1894) Einfluß der Configuration auf die Wirkung der Enzyme. Ber Dt Chem Ges 27:2985–2993

Fischer S, Verma CS (1999) Binding of buried structural water increases the flexibility of proteins. Proc Natl Acad Sci USA 96:9613–9615

Gabius H-J (1989) Endogene Lektine in Tumoren und ihre mögliche Bedeutung für Diagnose und Therapie von Krebserkrankungen. Onkologie 12:175–181

Gabius H-J (1994) Non-carbohydrate binding partners/domains of animal lectins. Int J Biochem 26:469–477

Gabius H-J (1997a) Animal lectins. Eur J Biochem 243:543–576

Gabius H-J (1997b) Concepts of tumor lectinology. Cancer Investig 15:454–464

Gabius H-J (1998) The how and why of protein-carbohydrate interaction: a primer to the theoretical concept and a guide to application in drug design. Pharm Res 15:23–30

Gabius H-J, Gabius S (eds) (1993) Lectins and Glycobiology. Springer Verlag, Berlin/ New York

Gabius H-J, Gabius S (eds) (1997) Glycosciences: Status and Perspectives. Chapman & Hall, London/Weinheim

Gabius H-J, Gabius S, Zemlyanukhina TV, Bovin NV, Brinck U, Danguy A, Joshi SS, Kayser K, Schottelius J, Sinowatz F, Tietze LF, Vidal-Vanaclocha F, Zanetta J-P (1993) Reverse lectin

histochemistry: design and application of glycoligands for detection of cell and tissue lectins. Histol Histopathol 8:369–383

Gabius H-J, Kayser K, Gabius S (1995) Protein-Zucker-Erkennung. Grundlagen und medizinische Anwendung am Beispiel der Tumorlektinologie. Naturwissenschaften 82:533–543

Gabius H-J, Unverzagt C, Kayser K (1998) Beyond plant lectin histochemistry: preparation and application of markers to visualize the cellular capacity for protein-carbohydrate recognition. Biotech Histochem 73:263–277

Gagneux P, Varki A (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. Glycobiology 9:747–755

García-Hernández E, Hernández-Arana A (1999) Structural bases of lectin-carbohydrate affinities: comparison with protein-folding energetics. Protein Sci 8:1075–1086

Geyer H, Geyer R (1998) Strategies for glycoconjugate analysis. Acta Anat 161:18–35

Gilleron M, Siebert H-C, Kaltner H, von der Lieth C-W, Kozár T, Halkes KM, Korchagina EY, Bovin NV, Gabius H-J, Vliegenthart JFG (1998) Conformer selection and differential restriction of ligand mobility by a plant lectin. Conformational behavior of Galb1-3GlcNAcb1-R, Galb1-3GalNAcb1-R and Galb1-2Galb1-R′ in the free state and complexed with mistletoe lectin as revealed by random walk and conformational clustering molecular mechanics calculations, molecular dynamics simulations and nuclear Overhauser experiments. Eur J Biochem 252:416–427

Gilli P, Ferretti V, Gilli G, Borea PA (1994) Enthalpy-entropy compensation in drug-receptor binding. J Phys Chem 98:1515–1518

Hardy BJ (1997) The glycosidic linkage flexibility and time-scale similarity hypotheses. J Mol Struct 395–396:187–200

Harris R, Kiddle GR, Field RA, Milton MJ, Ernst B, Magnani JL, Homans SW (1999) Stable-isotope-assisted NMR studies on ^{13}C-enriched sialyl Lewis^x in solution and bound to E-selectin. J Am Chem Soc 121:2546–2551

Hirabayashi J (1996) On the origin of elementary hexoses. Quart Rev Biol 71:365–380

Hirabayashi J, Arata Y, Kasai K-i (1997) Galectins from the nematode *Caenorhabditis elegans* and the genome project. Trends Glycosci Glycotechnol 9:113–122

Hooper LV, Manzella SM, Baenziger JU (1997) The biology of sulfated oligosaccharides. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 261–276

Hounsell EF (1997) Methods of glycoconjugate analysis. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 15–29

Imberty A (1997) Oligosaccharide structures: theory versus experiment. Curr Opin Struct Biol 7:617–623

Institut für Arzneimittelinformation (eds) (1999) Neue Konzepte gegen Virusgrippe: vielversprechend? arznei-telegramm 30:98–100

Jardetzky O (1980) On the nature of molecular conformations inferred from high-resolution NMR. Biochim Biophys Acta 621:227–232

Jiménez-Barbero J, Asensio JL, Cañada FJ, Poveda A (1999) Free and protein-bound carbohydrate structures. Curr Opin Struct Biol 9:549–555

Kaltner H, Stierstorfer B (1998) Animal lectins as cell adhesion molecules. Acta Anat 161:162–179

Karlsson K-A (1998) Meaning and therapeutic potential of microbial recognition of host glycoconjugates. Mol Microbiol 29:1–11

Karlsson K-A (1999) Bacterium-host protein-carbohydrate interactions and pathogenicity. Biochem Soc Transact 27:471–474

Kayser K, Gabius H-J (1999) The application of thermodynamic principles to histochemical and morphometric tissue research: principles and practical outline with focus on glycosciences. Cell Tissue Res 296:443–455

Kayser K, Bovin NV, Korchagina EY, Zeilinger C, Zeng F-Y, Gabius H-J (1994) Correlation of expression of binding sites for synthetic blood group A-, B-, and H-trisaccharides and for sarcolectin with survival of patients with bronchial carcinoma. Eur J Cancer 30A:653–657

Kishore U, Eggleton P, Reid KBM (1997) Modular organization of carbohydrate recognition domains in animal lectins. Matrix Biol 15:583–592

Kopitz J, von Reitzenstein C, Burchert M, Cantz M, Gabius H-J (1998) 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 273:11205–11211

Kuntz ID, Chen K, Sharp KA, Kollman PA (1999) The maximal affinity of ligands. Proc Natl Acad Sci USA 96:9997–10002

Laine RA (1997) The information-storing potential of the sugar code. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 1–14

Lee RT, Lee YC (1994) Enhanced biochemical affinities of multivalent neoglycoconjugates. In: Lee YC, Lee RT (eds) Neoglycoconjugates. Preparation and Applications, Academic Press, San Diego, pp. 23–50

Lee YC, Lee RT (eds) (1994) Neoglycoconjugates. Preparation and Applications. Academic Press, San Diego

Lee YC, Lee RT (1996) Glycobiology in medicine. J Biomed Sci 3:221–237

Lee RT, Lee YC (1997) Neoglycoconjugates. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 55–77

Lee RT, Gabius H-J, Lee YC (1992) Ligand-binding characteristics of the major mistletoe lectin. J Biol Chem 267:23722–23727

Lemieux RU (1996) How water provides the impetus for molecular recognition in aqueous solution. Acc Chem Res 29:373–380

Levi G, Teichberg VI (1981) Isolation and physicochemical characterization of electrolectin, a b-D-galactoside-binding lectin from the electric organ of *Electrophorus electricus*. J Biol Chem 256:5735–5740

Lingwood CD (1998) Oligosaccharide receptors for bacteria: a view to kill. Curr Opin Chem Biol 2:695–700

Lis H, Sharon N (1998) Lectins: carbohydrate-specific proteins that mediate cellular recognition. Chem Rev 98:637–674

Loris R, Hamelryck T, Bouckaert J, Wyns L (1998) Legume lectin structure. Biochim Biophys Acta 1383:9–36

Loukas A, Mullin NP, Tetteh KKA, Moens L, Maizels RM (1999) A novel C-type lectin secreted by a tissue-dwelling parasitic nematode. Curr Biol 9:825–828

Lumry R, Rajender S (1970) Enthalpy-entropy compensation phenomena in water solutions of proteins and small molecules: an ubiquitous property of water. Biopolymers 9:1125–1227

Mammen M, Choi S-K, Whitesides GM (1998) Polyvalente Wechselwirkungn in biologischen Systemen: Auswirkungen auf das Design und die Verwendung multivalenter Liganden und Inhibitoren. Angew Chem 110:2908–2953

Mann PL, Waterman RE (1998) Glycocoding as an information management system in embryonic development. Acta Anat 161:153–161

Marszalek PE, Oberhauser AF, Pang Y-P, Fernandez JM (1998) Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. Nature 396:661–664

McCafferty D-M, Smith CW, Granger DN, Kubes P (1999) Intestinal inflammation in adhesion molecule-deficient mice: an assessment of P-selectin alone and in combination with ICAM-1 or E-selectin. J Leukoc Biol 66:67–74

Newburg DS (1999) Human milk glycoconjugates that inhibit pathogens. Curr Med Chem 6:117–127

Nishio M, Umezawa Y, Hirota M, Takeuchi Y (1995) The CH/p interaction: significance in molecular recognition. Tetrahedron 51:8665–8701

Noorman F, Barrett-Bergshoeff MM, Rijken DC (1998) Role of carbohydrate and protein in the binding of tissue-type plasminogen activator to the human mannose receptor. Eur J Biochem 251:107–113

Oriol R, Mollicone R, Cailleau A, Balanzino L, Breton C (1999) Divergent evolution of fucosyl-transferase genes from vertebrates, invertebrates and bacteria. Glycobiology 9:323–334

Paulson JC (1996) Leukocyte adhesion deficiency type II. In: Montreuil J, Vliegenthart JFG, Schachter H (eds) Glycoproteins and Disease, Elsevier, Amsterdam, pp. 405–411

Pavelka M (1997) Topology of glycosylation - a histochemist's view. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 115–120

Peters T, Pinto BM (1996) Structure and dynamics of oligosaccharides: NMR and modeling studies. Curr Opin Struct Biol 6:710–720

Poppe L, Brown GS, Philo JS, Nikrad PV, Shah BH (1997) Conformation of sLe^x tetrasaccharide, free in solution and bound to E-, P-, and L-selectin. J Am Chem Soc 119:1727–1736

Poveda A, Jiménez-Barbero J (1998) NMR studies of carbohydrate-protein interaction in solution. Chem Soc Rev 27:133–143

Powell LD, Varki A (1995) I-type lectins. J Biol Chem 270:14243–14246

Quiocho FA (1988) Molecular features and basic understanding of protein-carbohydrate interactions: the arabinose-binding protein-sugar complex. Curr Top Microbiol Immunol 139:135–148

Rabinovich GA, Daly G, Dreja H, Tailor H, Riera CM, Hirabayashi J, Chernajovsky Y (1999) Recombinant galectin-1 and its genetic delivery suppress collagen-induced arthritis via T cell apoptosis. J Exp Med 190:385–397

Reuter G, Gabius H-J (1996) Sialic acids. Structure, analysis, metabolism, and recognition. Biol Chem Hoppe-Seyler 377:325–342

Reuter G, Gabius H-J (1999) Eukaryotic glycosylation - whim of nature or multipurpose tool? Cell Mol Life Sci 55:368–422

Reutter W, Stäsche R, Stehling P, Baum O (1997) The biology of sialic acids: insights into their structure, metabolism and function in particular during viral infection. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 245–259

Rini JM (1995) Lectin structure. Annu Rev Biophys Biomol Struct 24:551–577

Rini JM, Lobsanov YD (1999) New animal lectin structures. Curr Opin Struct Biol 9:578–584

Rüdiger H, Siebert H-C, Solís D, Jiménez-Barbero J, Romero A, von der Lieth C-W, Díaz-Mauriño T, Gabius H-J (1999) Medicinal chemistry based on the sugar code: fundamentals of lectinology and experimental strategies with lectins as targets. Curr Med Chem (in press)

Schachter H (ed) (1999) Molecular basis of glycoconjugate disease. Biochim Biophys Acta 1455:61–418

Sears P, Wong C-H (1998) Enzyme action in glycoprotein synthesis. Cell Mol Life Sci 54:223–252

Sears P, Wong C-H (1999) Kohlenhydratmimetika: ein neuer Lösungsansatz für das Problem der kohlenhydratvermittelten biologischen Erkennung. Angew Chem 111:2446–2471

Sharon N, Lis H (1997) Glycoproteins: structure and function. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 133–162

Shirai T, Mitsuyama C, Niwa Y, Matsui Y, Hotta H, Yamane T, Kamiya H, Ishii C, Ogawa T, Muramoto K (1999) High-resolution structure of the conger eel galectin, congerin I, in lactose-liganded and ligand-free forms: emergence of a new structure class by accelerated evolution. Structure 7:1223–1233

Siebert H-C, Adar R, Arango R, Burchert M, Kaltner H, Kayser G, Tajkhorshid E, von der Lieth C-W, Kaptein R, Sharon N, Vliegenthart JFG, Gabius H-J (1997) Involvement of laser photo CIDNP (chemically induced dynamic nuclear polarization)-reactive amino acid side chains in ligand binding by galactoside-specific lectins in solution. Similarities in the role of tryptophan/tyrosine residues for ligand binding between a plant agglutinin and mammalian/avian galectins and the detection of an influence of single-site mutagenesis on surface presentation of spatially separated residues. Eur J Biochem 249:27–38

Siebert H-C, Gilleron M, Kaltner H, von der Lieth C-W, Kozár T, Bovin NV, Korchagina EY, Vliegenthart JFG, Gabius H-J (1996) NMR-based, molecular dynamics- and random walk molecular mechanics-supported study of conformational aspects of a carbohydrate ligand (Galb1-2Galb1-R) for an animal galectin in the free and in the bound state. Biochem Biophys Res Commun 219:205–212

Siebert H-C, von der Lieth C-W, Gabius H-J (1999) Der Zuckerkode - Oligosaccharide als Träger biologischer Information. Dtsch Apotheker Ztg 139:272–282

Simanek EE, McGarvey GJ, Jablonowski JA, Wong C-H (1998) Selectin-carbohydrate interactions: from natural ligands to designed mimics. Chem Rev 98:833–862

Simon PM (1996) Pharmaceutical oligosaccharides. Drug Discovery Today 1:522–528

Solís D, Díaz-Mauriño T (1997) Analysis of protein-carbohydrate interaction by engineered ligands. In: Gabius H-J, Gabius S (eds) Glycosciences: Status and Perspectives, Chapman & Hall, London/Weinheim, pp. 345–354

Solís D, Estremera D, Usobiaga P, Díaz-Mauriño T (1987) Differential binding of mannose-specific lectins to the carbohydrate chains of fibrinogen domains D and E. Eur J Biochem 165:131–138

Solís D, Romero A, Kaltner H, Gabius H-J, Díaz-Mauriño T (1996) Different architecture of the combining sites of two chicken galectins revealed by chemical-mapping studies with synthetic ligand derivatives. J Biol Chem 271:12744–12748

Sonnhammer EL, Eddy SR, Birney E, Bateman A, Durbin R (1998) Pfam: multiple sequence alignment and HMM-profiles of protein domains. Nucl Acids Res 26:320–322

The C. elegans Sequencing Consortium (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. Science 282:2012–2018

Toone EJ (1994) Structure and energetics of protein-carbohydrate complexes. Curr Opin Struct Biol 4:719–728

Uedaira H, Uedaira H (1985) The relationship between partial molecular heat capacities and the number of equatorial hydroxyl groups of saccharides. J Sol Chem 14:27–34

Varela PF, Solís D, Díaz-Mauriño T, Kaltner H, Gabius H-J, Romero A (1999) The 2.15 Å crystal structure of CG-16, the developmentally regulated homodimeric chicken galectin. J Mol Biol 294:537–549

Varki A (1996) "Unusual" modifications and variations of vertebrate oligosaccharides: are we missing the flowers for the trees? Glycobiology 6:707–710

Varki A (1998) Factors controlling the glycosylation potential of the Golgi apparatus. Trends Cell Biol 8:34–40

Varki A, Marth J (1995) Oligosaccharides in vertebrate development. Sem Develop Biol 6:127–138

Vasta GR, Quesenberry M, Ahmed H, O'Leary N (1999) C-type lectins and galectins mediate innate and adaptive immune functions: their roles in the complement activation pathway. Dev Comp Immunol 23:401–420

Villalobo A, Gabius H-J (1998) Signaling pathways for transduction of the initial message of the glycocode into cellular responses. Acta Anat 161:110–129

von der Lieth C-W, Kozár T, Hull WE (1997a) A (critical) survey of modeling protocols used to explore the conformational space of oligosaccharides. J Mol Struct 395–396:225–244

von der Lieth C-W, Lang E, Kozár T (1997b) Carbohydrates: second-class citizens in biomedicine and in bioinformatics? In: Hofestädt R, Lengauer T, Löffler M, Schomburg D (eds) Bioinformatics, Springer, Berlin/New York, pp. 147–155

von der Lieth C-W, Siebert H-C, Kozár T, Burchert M, Frank M, Gilleron M, Kaltner H, Kayser G, Tajkhorshid E, Bovin NV, Vliegenthart JFG, Gabius H-J (1998) Lectin ligands: new insights into their conformations and their dynamic behavior and the discovery of conformer selection by lectins. Acta Anat 161:91–109

von Figura K (1990) Mannose-6-phosphat-Rezeptoren: ihre Rolle beim Transport lysosomaler Proteine. Naturwissenschaften 77:116–122

von Itzstein M, Thomson RJ (1997) Sialic acids and sialic acid-recognising proteins: drug discovery targets and potential glycopharmaceuticals. Curr Med Chem 4:185–210

Vyas NK (1991) Atomic features of protein-carbohydrate interactions. Curr Opin Struct Biol 1:732–740

Weis WI, Drickamer K (1996) Structural basis of lectin-carbohydrate recognition. Annu Rev Biochem 65:441–473

White KD, Cummings RD, Waxman FJ (1997) Ig N-glycan orientation can influence interactions with the complement system. J Immunol 158:426–435

Winterburn PJ, Phelps CF (1972) The significance of glycosylated proteins. Nature 236:147–151

Woods RJ (1998) Computational carbohydrate chemistry: what theroretical methods can tell us. Glycoconjugate J 15:209–216

Wormald MR, Dwek RA (1999) Glycoproteins: glycan presentation and protein-fold stability. Structure 7:R155–R160

Yamey G (1999) Anti-flu drug may not reduce death rate. Br Med J 319:659

Zanetta J-P (1998) Structure and functions of lectins in the central and peripheral nervous system. Acta Anat 161:180–195

Zopf D, Roth S (1996) Oligosaccharide anti-infective agents. Lancet 347:1017–1021