ChemBioChem

Review
doi.org/10.1002/cbic.202100327

Chemistry Europe
European Chemical Societies Publishing

www.chembiochem.org

What is the Sugar Code?

Hans-Joachim Gabius,[a] Maré Cudic,[b] Tammo Diercks,[c] Herbert Kaltner,*[a] Jürgen Kopitz,[d] Kevin H. Mayo,[e] Paul V. Murphy,*[f] Stefan Oscarson,[g] René Roy,[h] Andreas Schedlbauer,[c] Stefan Toegel,[i] and Antonio Romero[j]

Dedicated to Prof. Samuel H. Barondes and to the memory of Prof. Dr. Hans-Joachim Gabius.

A code is defined by the nature of the symbols, which are used to generate information-storing combinations (e.g. oligo- and polymers). Like nucleic acids and proteins, oligo- and polysaccharides are ubiquitous, and they are a biochemical platform for establishing molecular messages. Of note, the letters of the sugar code system (third alphabet of life) excel in coding capacity by making an unsurpassed versatility for isomer (code word) formation possible by variability in anomery and linkage position of the glycosidic bond, ring size and branching. The enzymatic machinery for glycan biosynthesis (writers) realizes this enormous potential for building a large vocabulary. It includes possibilities for dynamic editing/erasing as known from nucleic acids and proteins. Matching the glycome diversity, a large panel of sugar receptors (lectins) has developed based on more than a dozen folds. Lectins 'read' the glycan-encoded information. Hydrogen/coordination bonding and ionic pairing together with stacking and C–H/π-interactions as well as modes of spatial glycan presentation underlie the selectivity and specificity of glycan-lectin recognition. Modular design of lectins together with glycan display and the nature of the cognate glycoconjugate account for the large number of post-binding events. They give an entry to the glycan vocabulary its functional, often context-dependent meaning(s), hereby building the dictionary of the sugar code.

1. Introduction

"To an observer trying to obtain a bird's eye view of the present state of biochemistry – life may until very recently have seemed to depend on only two classes of compounds: nucleic acids and proteins."[1] They are connected by the genetic code. The sequence of three symbols (letters) of the first alphabet of life (nucleotides) stands either for an amino acid (one of the letters of the second alphabet of life) or a stop signal so that nucleic acid becomes the template for protein biosynthesis. In this special (though fundamental) case of using the term 'code' in life sciences, the information stored in a (nucleotide) sequence has the biological meaning of a sequence of a protein: the dictionary for the 64 entries of the vocabulary of trinucleotides provides their translation into an amino acid or a stop signal. In other cases of using the term 'code', the information encoded in combinations of biochemical symbols (the molecular messages or code words) is 'understood' (decoded) by a 'reader' (receptor). It then initiates the translation of this information into biofunctionality by post-binding ('reading') events, and here sugars come into play. That they have for example been assumed to be the letters of "a potential carbohydrate "language" involved in intercellular interactions"[2] or a molecular basis of the cell-surface code[3] illustrates their status as third alphabet of life.

Originating from the analytical milestones of the identification of the biochemical nature of (snail) mucin as glycoprotein[4] and of "Glycosamin" (*N*-acetylglucosamine, Glc*N*Ac) as building block of the polysaccharide chitin,[5] the focus of work on glycoproteins continued to be on elucidating structural and synthetic aspects for a long time.[6] This situation changed after having documented the abundance, ubiquitous presence and structural diversity of glycans on cell surfaces (and also in a polysaccharide-rich ("sugary") coating termed glycocalyx[7]) and after having realized the enormous potential of the described structural complexity of oligosaccharides for information coding.[2,8] Glycocompounds obviously appeared to have more talents than to store energy and to be a molecular concrete for cell wall stability. The catchword summarizing the resulting hypothesis of their involvement in cellular processes on a broad scale as molecular messages simply connected 'sugar' and 'code'. Historically, serendipity (of local vicinity of two labs) helped to do so. At the time when the genetic code was

[a] *H.-J. Gabius, H. Kaltner*
Institute of Physiological Chemistry, Faculty of Veterinary Medicine
Ludwig-Maximilians-University Munich
Veterinärstr. 13, 80539 Munich (Germany)
E-mail: kaltner@lmu.de
[b] *M. Cudic*
Department of Chemistry and Biochemistry
Charles E. Schmidt College of Science, Florida Atlantic University
777 Glades Road, Boca Raton, Florida 33431 (USA)
[c] *T. Diercks, A. Schedlbauer*
Center for Cooperative Research in Biosciences (CIC bioGUNE)
Basque Research and Technology Alliance (BRTA)
Bizkaia Technology Park, Building 801 A, 48160 Derio, Bizkaia (Spain)
[d] *J. Kopitz*
Institute of Pathology, Department of Applied Tumor Biology
Faculty of Medicine, Ruprecht-Karls-University Heidelberg
Im Neuenheimer Feld 224, 69120 Heidelberg (Germany)
[e] *K. H. Mayo*
Department of Biochemistry, Molecular Biology & Biophysics
University of Minnesota
Minneapolis, MN 55455 (USA)
[f] *P. V. Murphy*
CÚRAM – SFI Research Centre for Medical Devices
and the School of Chemistry
National University of Ireland Galway
University Road, Galway H91 TK33 (Ireland)
E-mail: paul.v.murphy@nuigalway.ie
[g] *S. Oscarson*
Centre for Synthesis and Chemical Biology
University College Dublin
Belfield, Dublin 4 (Ireland)
[h] *R. Roy*
Département de Chimie et Biochimie
Université du Québec à Montréal, Case Postale 888
Succ. Centre-Ville Montréal (Québec), H3C 3P8 (Canada)
[i] *S. Toegel*
Karl Chiari Lab for Orthopaedic Biology
Department of Orthopedics and Trauma Surgery
Medical University of Vienna
Vienna (Austria)
[j] *A. Romero*
Department of Structural and Chemical Biology
CIB Margarita Salas, CSIC
Ramiro de Maeztu 9, 28040 Madrid (Spain)

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cbic.202100327

cracked, the term 'sugar code' was suggested by drawing the following analogy: "just as Marshall (Nirenberg) was working on a *nucleic acid code* that determined the structure of proteins, Vic(tor) Ginsburg [a member of the Gordon Tomkins laboratory, whose lab was across the hall from Marshall's at NIH Building 10] believed there is a *sugar code* that determines intercellular interactions" (p. 25).[9]

When applying the code concept to glycans, the first step on the way to prove that there is every reason to believe in their position in the flow of biological information is to explain why "carbohydrates are ideal for generating compact units with explicit informational properties".[8b] Using these biochemical symbols, a sophisticated system of 'writers', together with 'editors & erasers', leverages sugars to generate the large vocabulary of the carbohydrate language (the glycome).[10] Since carbohydrates are equipped with ample chemical means for molecular recognition such as their hydroxyl groups (the stoichiometric proportion of $C_n(H_2O)_m$ with $n \leq m$ in hexoses and the etymological roots for carbon (Latin '*carbo*' = coal) and water (Greek '*hydor*') explain the origin of the term 'carbohydrate'; for further etymological information, please see Ref. [11]), the 'reading' of 'words' written in sugar is easy. Tissue receptors (lectins) are a link between the glycan-encoded information and the actual process such as cell adhesion so that a primer to mammalian lectins will follow.

It starts by highlighting their diversity on the level of protein folds. Once a peptide fold has acquired ability to bind sugar, this structure is a starting point for ensuing evolutionary diversification. Variations in lectin sequence and modular design as well as selectivity and specificity of their pairing with cellular glycoconjugates underlie the translation of the glycan-encoded message into a distinct cellular response. The long-term aim of work on the sugar code, i.e. to compile a dictionary for the vocabulary (listing the functional meaning(s) for glycan words), is finally sketched by describing biomedical activity of glycan-lectin recognition exemplarily (in its cellular context). Since "only in recent years have we begun to appreciate how deeply glycan functions pervade all aspects of organismic biology, molecular biology, and biochemistry",[12] this introduction to the concept of the sugar code can be of interest for a broad readership.

2. Letters of the Sugar Alphabet

Prebiotic conditions on earth are generally assumed to have allowed the synthesis of glyceraldehyde and its keto-tautomer, which then formed ketohexoses by aldol condensation.[13] Lobry de Bruyn rearrangement led to the aldohexoses D-glucose (Glc) and D-mannose (Man), primitive metabolism then to D-galactose (Gal), which have no or just one 1,3-diaxial interaction, so that their presence in polysaccharides and eukaryotic glycans is thermodynamically favored (the mystery why Glc is not present in glycans of mature glycoproteins will be solved below).[14] The structures of often used carbohydrate letters are shown in Supporting Information, Figure S1. Intriguingly, with Gln as donor, Glc can be converted to the amino sugar glucosamine (Glc*N*), which like Gal*N* is then *N*-acetylated to yield Glc*N*Ac and Gal*N*Ac (Figure S1). A biochemical letter (see also below for the case of 5-methylcytosine) is specifically modified, through which the collection of symbols (alphabet of life) is extended in number. The impact of a modification on a letter's meaning is nicely demonstrated by the following analogy: these derivatives can be considered as the equivalent of an Umlaut in the German language used for the letters A (i.e. Ä), O (i.e. Ö) and U (i.e. Ü).

In order to assemble oligo- and polymers, monosaccharides - like nucleotides or amino acids – must first be activated, and they are most reactive at their anomeric center for this purpose.[15] Physiologically, the resulting conjugate with a nucleoside mono- or diphosphate acts as the glycosyl donor for the enzymatic transfer of sugar to an acceptor, and this specifically in either the α- or the β-anomeric position. This is the first source for structural variability of glycans beyond the sequence. The next one is due to the chemical equivalence of the other hydroxyl groups besides the one at the anomeric center.

In contrast to making phosphodiester and peptide bonds, glycan biosynthesis is not restricted to connecting fixed positions between donor and acceptor. Instead, it can engage more than one hydroxyl group of a sugar used as an acceptor when an oligosaccharide is built in steps by glycosyltransferases. To give an example, a single diglycoside would be expected as sole product when a donor-acceptor pair is linked in nucleic acid/protein style.

The possibility of the enzymatic transfer of a sugar to more than one acceptor site, however, let it become clear that more than a single product will be obtained, and we illustrate this principle in a figure. Using L-fucose as a graphic example, its naturally occurring α1,2-, α1,3-, α1,4- or α1,6-linkages are drawn in Figure 1 (top panel). Impaired fucosylation, to underline clinical significance, is the cause of a leukocyte adhesion deficiency (LAD II/CDG IIc) and, in mouse models with engineered deficiencies in α1,3- or α1,6-fucosylation, of disorders in leukocyte trafficking (by lowering production of ligands for cell adhesion molecules, i.e. selectins; see below) or diminished growth factor signaling, respectively.[16] To showcase the various positions of Fuc residues in natural glycans, the structures of the histo-blood group ABH(0) and Lewis (Le^x/Le^y) determinants and of the *N*-glycan stem with its core fucosylation are presented in the bottom panel of Figure 1 (please note that the presence of Fuc in the blood-group H(0) epitope and its property as ligand were the prerequisites to demonstrate inhibition of hemagglutination by a sugar, here a derivative of Fuc: see below).

Inspecting the structures of the oligosaccharides shown in Figure 1 now closely makes evident that – in contrast to nucleic acid and proteins – a branch is installed into a linear glycan by fucosyltransferases (a second example for branching by glycosyltransferases is shown in Supporting Information, Figure S2 and is explained below).[17] Moreover, the Fuc moiety can also be transferred from its donor (GDP-Fuc) directly to the hydroxyl of serine or threonine in *O*-fucosylation of epidermal growth

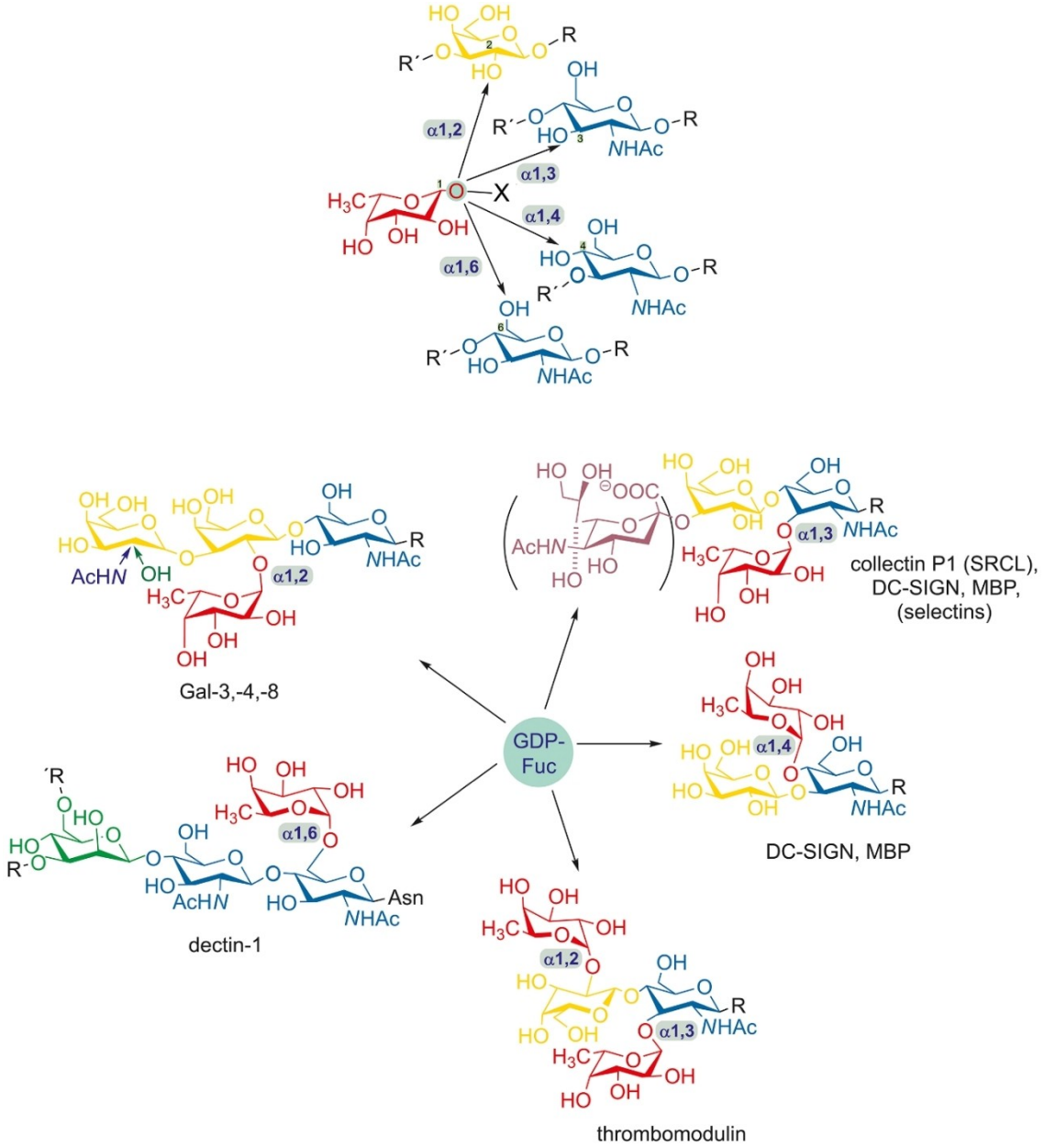

Figure 1. Illustration of the four routes to transfer the Fuc moiety from its GDP-Fuc donor (GDP given as X) to the 2, 3, 4, or 6 position of glycan acceptors by mammalian fucosyltransferases (top panel), examples of resulting oligosaccharides in *N*- and *O*-glycans that define the R′ position are presented in the bottom panel. This part shows glycans with α1,2-fucosylation (histo-blood group ABH(0) epitopes), with α1,3-fucosylation ((sialyl) Le^x), with α1,4-fucosylation (Le^a) and with α1,2/4-fucosylations (Le^y) as well as the *N*-glycan stem with α1,6-fucosylation termed core fucosylation (examples for lectins that bind the respective structure are named.

factor (EGF)-like and thrombospondin type 1 repeats in the endoplasmic reticulum.[17b,18]

Overall, our case study of fucosylation thus teaches the lesson that the chemical properties of monosaccharides make activation at the anomeric center and natural variability of where to add the sugar to an acceptor possible. Also in contrast to nucleic acids and proteins, branching is common in glycans. On the side of the enzymes, the availability of a group of acceptor site-specific glycosyltransferases for each letter of the sugar alphabet, the fucosyltransferase family consisting of 13 members in mammals,[17,18] ensures to realize the enormous inherent potential of carbohydrates to yield glycan diversity. Since the enzymatic apparatus for glycan biosynthesis with a total of at least 167 glycosyltransferases has developed the required complexity during evolution to prepare many more than a few isomers like writing with letters of an alphabet does, coding by glycans will reach the comparatively highest capacity. Clearly, it would mean missing manifold opportunities if doing so were without physiological significance. It is thus fair to

conclude that it is making a snap judgement when underestimating sugars as code symbols.

Intriguingly, even a further structural feature has been detected that increases diversity among glycans, i.e. the ring size (5-membered furanose *vs* 6-membered pyranose). The frequent presence of galactofuranose (Gal*f*, shown in Figure S1) in polysaccharides and glycoconjugates of bacteria, fungi and parasitic protozoa is a proof-of-principle case.[19] Its profile of distribution in Nature therefore predestines the occurrence of the five-membered ring for Gal as an indicator of non-self origin of a complex carbohydrate, to which host defense can be directed (see below). Similarly, *O*-methylated Man/Fuc residues resulting from *S*-adenosylmethionine-dependent modification (as in methylation of cytosine or histones) offer such a target site, because they are absent in mammals.[20] This qualitative difference nourishes the expectation for an exploitation of this sugar-based trait in host defense, too (see below).

In summary, adding the ring size to the status of anomery and to the ability of all hydroxyl groups as possible acceptor incl. the frequent occurrence of branching accounts for the unsurpassed level of structural permutations within glycans among biomolecules. When expressing the information-storing potential of the sugar code in numbers, i.e. the coding capacity, a set of six letters from the sugar alphabet (shown in Supporting Information, Figure S1) will theoretically build 1.05×10^{12} linear and branched hexasaccharides, quite favorably comparing in pool size to the 6.4×10^{7} hexapeptides from 20 amino acids,[21] and there is more. Evolution has developed even more means to further increase the coding capacity by biochemical symbols. This is done by introducing biochemical modifications after the assembly of oligo- and polymers, that is post-synthetically.

This natural diversification strategy is common to all three types of biomolecular alphabet. Diverse types of substitution of a basic structure like the addition of a phosphate (in two steps) or a sulfate are known to occur in letters of the sugar alphabet, as diverse as they are for example known from nucleotides,[22] and they are clinically relevant as the case of deficiency in Man phosphorylation as cause of the I-cell disease (mucolipidosis II) underscores.[23] The initial placement of a modification into a glycan can be followed by further enzymatic processing. As can happen with 5-methylcytosine, the fifth letter in DNA, by hydroxylation, the methyl group of an *N*-acetyl substituent can similarly be oxidized to the hydroxylmethyl (in *N*-acetylneuraminic acid to yield the *N*-glycol(o)yl as shown in Figure S1 (bottom), one of up to nearly 50 ways to create a sialic acid from this parental compound).[24]

In general, post-synthetic modifications give letters a new meaning. Phosphorylation (in the 6-position of Man labeling glycans of lysosomal enzymes) or sulfation (at the 4-position of a branch-end Gal*N*Ac in *N*-glycans of distinct glycoproteins such as certain pituitary glycohormones (LH, TSH) or at the 3-position of the sulfatides' Gal headgroup) at specific sites in glycans have been likened to a postal-code writing for transport processes (see below). It now becomes clear why carbohydrates had rightly been judged to be "ideal" for this purpose.[8b]

Ironically, exactly this property had been responsible to slow down progress of research. "In this remarkable age of genomics, proteomics, and functional proteomics, I am often asked by my colleagues why glycobiology has apparently lagged so far behind the other fields. The simple answer is that glycoconjugates are much more complex, variegated, and difficult to study than proteins or nucleic acids."[25] Interestingly, this already holds true for individual letters: the elucidation of the structure of the cited *N*-acetylneuraminic acid, for which a total of 11 structures were proposed over time, took 25 years.[26] After having surveyed the structural basis for reading high-level versatility within the sugar code, we now move from the alphabet of sugar letters to the vocabulary of glycans.

3. The Vocabulary of the Sugar Code

The presented proof-of-principle case of fucosylation has illustrated the existence of an elaborate system for enzymatic assembly to turn the described potential of sugars for structural glycan diversity into reality. The members of the team of the sets of glycosyltransferases with their genuine specificities for donor and acceptor pairs as well as for the status of anomery (α or β), for linkage positions and for ring size are called the writers. Products of glycogenes for example for sugar activation and transport assist and feed the assembly line. Fluctuations in the status of substrate and enzyme availabilities will dynamically modulate characteristics of the product panel, as for instance work with toxin (lectin)-resistant cell mutants and detection of compensatory responses within the glycome to an engineered deficiency for a glycosyltransferase *in vivo* revealed.[27] Writing proceeds in principle in a stepwise manner to generate linear and branched glycans. Chain elongation can produce repeats of a building block, i.e. the *N*-acetyllactosamine (Lac*N*Ac) unit building oligo- or polyLac*N*Ac sequences in (β-1,6-linked) *N*- and (core 2/4) *O*-glycan branches or the glycosaminoglycan keratan sulfate; branched structures reach an up to penta-antennary design in the case of the complex-type *N*-glycans (the names of the six involved β1,2/4/6-Glc*N*Ac transferases (GnTs) along with product designations after sequential Glc*N*Ac transfer to the *N*-glycan core pentasaccharide are graphically displayed in Supporting Information, Figure S2; the 11 GnTs for the β1,3-linkage are involved in other pathways of glycan biosynthesis; see below).[28] Carbohydrate chemistry has succeeded to develop elegant strategies for production of such structures, as exemplarily shown for a Lac*N*Ac dimer (DiLac*N*Ac) in Figure 2 (top part; for details on synthesis, please see Supporting Information, Scheme S1). Such synthetic oligosaccharides can then be used for interaction analysis such as calorimetry or spectroscopy (please see below).

Alternatively, such glycan derivatives can be conjugated to proteins and dendrimer scaffolds. What started with neoglycoproteins used as antigens has led to diverse applications of the products as sensors for the presence of sugar receptors and the elucidation of their special binding properties.[29] This way, biomimetics of cellular glycoconjugates with up to multi-

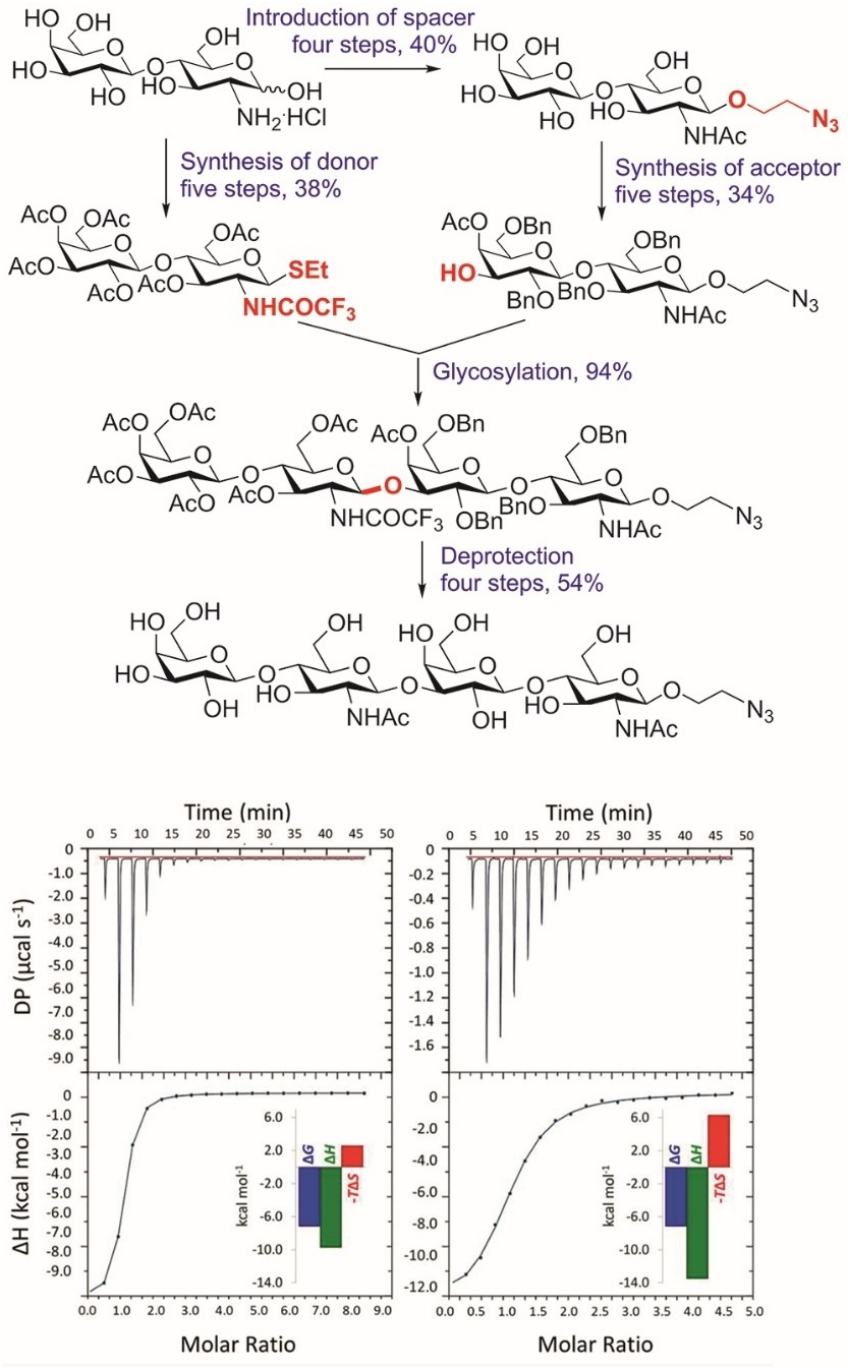

Figure 2. Overview on DiLac*N*Ac synthesis (for details, please see Supporting Information, Scheme S1) and the calorimetric titration profiles of its interaction with human galectin-3 in H_2O (bottom, left) and D_2O (bottom, right). For details, please see Ref. [63e].

antennary *N*-glycans or of the clustered appearance of *O*-glycans in mucins are obtained.

Examples of a trivalent glycocluster and of a (starburst) 16mer glycodendrimer are presented in Figure 3 (details given in Supporting Information, Scheme S2). Synthetic glycoclusters and glycodendrimers are valuable tools to answer questions on the relevance of topological features of glycan presentation for their biological meaning, and, therefore, their successful application spurs continued vigorous synthetic efforts (for work to coin the common term 'glycoside cluster effect', please see below).[30] With glycodendrimers at hand, not only the natural branching of glycans in glycoconjugates can be mimicked but models for experimental study can be brought to the level of cell surface (microdomain)-like glycoconjugate presentation. The synthesis of amphiphilic Janus glycodendrimers, capable to self-associate to various types of nanoparticles as glycosphingolipid assembly underlies the classical liposomes, paved the way to prepare fully surface-programmable vesicle-like models to

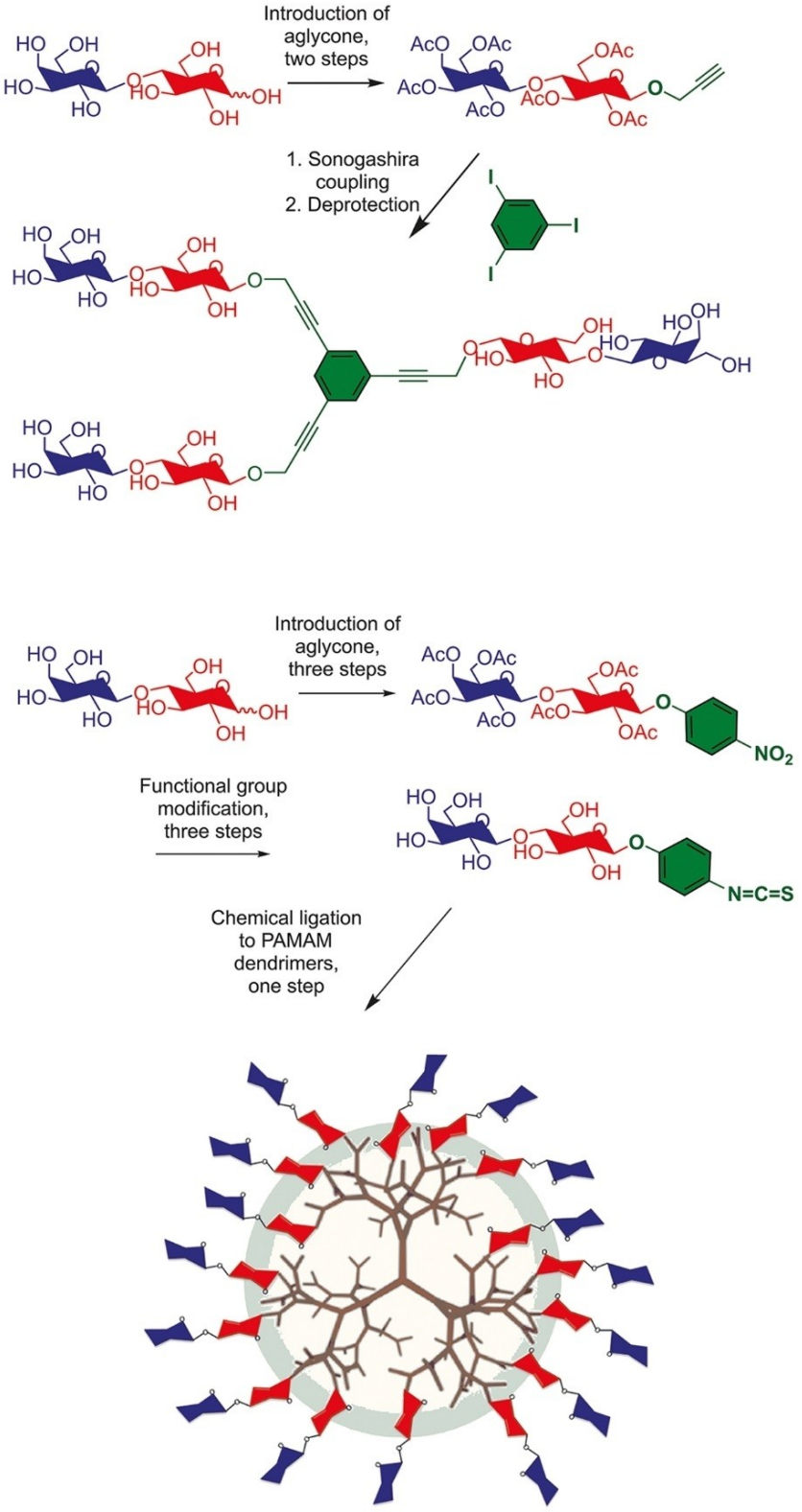

Figure 3. Illustrations of the syntheses of a triiodobenzene-based trivalent glycocluster (top) and of a 16mer starburst glycodendrimer (bottom). For further information on the syntheses, please see Supporting Information, Scheme S2 and the original reports with details on results of lectin assays.[96]

systematically study bridging phenomena by sugar receptors in a bottom-up manner.[31]

After the initial writing process, specific letters within glycans can be modified by the equivalents of editors, as already indicated above: sulfotransferases for respective *N*- and *O*-substitutions or an epimerase for converting D-glucuronic acid into L-iduronic acid in glycosaminoglycans belong to this group (see Figure S1 and below). These enzymes can cooperate to implement the enormous diversity in the disaccharide unit of the glycosaminoglycan chains of proteoglycans. Creating different patterns of substitutions is an intriguing strategy to let this simple structural platform acquire complexity, as shown in Supporting Information, Figure S3.[32] In space and time, glycan structures and their modification patterns do not remain unchanged. Erasers remove added groups from the 'message' up to the size of a letter, this for example seen for sialic acids from oligosaccharides of distinct gangliosides such as GD1a upon cell activation or differentiation (see below) and for Glc and Man residues during the maturation of *N*-glycans in the endoplasmic reticulum.[33] In appealing analogy to the chain of events when shaping the vocabulary of nucleic acids or proteins, writers intimately team up with editors and erasers to increase sequence variability and add dynamic reshaping to information coding. Cycles of post-synthetic modification and removal of substitutions by coordinated editor & eraser activities thus are a hallmark of all three coding platforms in the flow of biological information. The case study of the events to shape the histone code highlights these principles.[34] The fundamental lesson thus is that fine-tuning of vocabularies by post-synthetic processing is a common feature of biochemical codes that are based on each of the three alphabets of life.

Has the analytical technology to define the glycan vocabulary (glycome) reached the necessary level to perform its detailed mapping on the level of cells? Starting from the stepwise characterization of cellular glycans and of the parts of the enzymatic machinery to generate them, global profiling of the products of glycosylation pathways of wild-type and genetically engineered (for glycogenes) eukaryotic cells has indeed been achieved.[35] Glycan analysis at the glycome level (for recent review on experimental approaches, please see Ref. [36]) enables respective profile monitoring at high-level sensitivity, big-data glycomics then leading to its integration into systems biology.[37]

The step from detection and characterization of glycans to their localization in cells and tissues is facilitated by cyto- and histochemistry using cells and sections as assay platform. The initial approach of monitoring sugar presence by performing chemical visualization protocols such as the periodic acid-Schiff stain[38] has been replaced by using sugar receptors, and this with a considerable gain of specificity. Figure 4 documents how distribution profiles of distinct glycan determinants in sections of fixed and paraffin-embedded tissues look like. The systematic application of this technique has revealed clearly non-uniform/-random patterns of presence of glycan determinants with spatiotemporal dynamics of expression.[39] Fine-tuned regulatory mechanisms on genetic, epigenetic, transcriptional and post-transcriptional/-translational levels in combination with

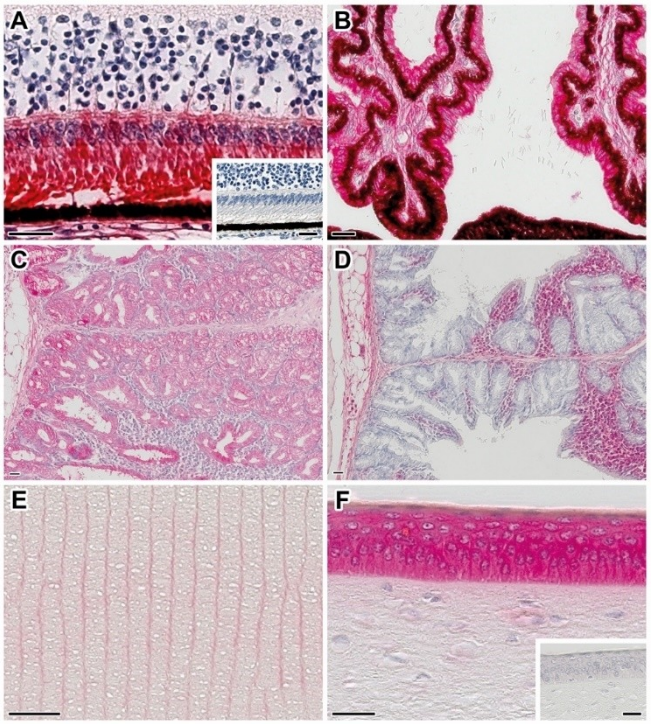

Figure 4. Lectin histochemical localization of glycans in sections through retina and anterior segments of fixed adult chicken eyes. Detection of the mucin-type *O*-glycan core 1 disaccharide (TF antigen) in retina's photoreceptor layer (inset: inhibition control with cognate sugar) (A), of Lac*N*Ac oligomers in connective tissue and epithelial cells of the ciliary body (B), of α2,3 (C)- or α2,6 (D)-sialylated *N*-glycans in immune cell aggregates of Haderian gland, of β1,6-branched *N*-glycans between lens fibers (E) and of β-galactosides (bound by the labeled chicken galectin CG-1B) in corneal epithelium (inset: inhibition control with cognate sugar) (F). Scale bars: 20 μm (Reproduced with permissions from Ref. [97b] Copyright 2017 John Wiley and Sons and from Ref. [97c] Copyright 2018 Elsevier; for technical details and information on the lectins used as tools, please see Ref. [97]).

the noted factor acceptor/donor availabilities turn the large potential for glycan diversity into reality, such mechanisms recently unpicked in the cases of the two α2,6-sialyltransferases (an example for detection of α2,6-sialylated *N*-glycans by a fungal lectin is shown in Figure 4D).[40] Swift coordinated reactions to external factors, e.g. a stressor such as tunicamycin that blocks the route toward *N*-glycosylation at its first step, here to safeguard homeostasis by the unfolded protein response,[41] support the conclusion of a broad-scale physiological significance of protein glycosylation, as emphasized by an explicit statement from the literature given above.[12] Looking at the transcriptional regulation, the multitude of permutations of individual control elements for expression of glycogenes is comparable to what a multi-dimensional switchboard can achieve and a challenge for explorations. In our context, it is imperative to underline that specific glycan-protein recognition underlies this method for detecting saccharides and hereby monitoring spatiotemporal expression patterns. The letters of the sugar alphabet are well-suited to make this interaction selective and specific because they offer regions of considerable size for molecular complementarity, which we will look at next.

Hydroxyl groups are readily accessible to establish directional hydrogen or coordination (with Ca^{2+} in proteins) bonding. The position of each substituent is checked in this process for complementarity: if either the equatorial hydroxyl of Glc(*N*Ac)/Man or the axial hydroxyl of Gal at the C4-position (see Supporting Information, Figure S1) together with a second OH group is engaged for example in coordination bonding (in analogy to bidentate H-bonding involving side chains of Arg and others), then the lectin's specificity to Man or Gal is readily explained. Hydrophobic complementarity is achieved with methyl groups present for example in Fuc or Glc*N*Ac/Gal*N*Ac. C–H/π-interactions are possible between the B-face of Gal (that presents the slightly polarized C–H groups) and the π-electrons of a Trp residue. Last but not least, strong ionic pairing brings sialic acids and charged sugar derivatives such as sulfated epitopes in contact with strategically positioned basic amino acids (or H-bonding donors, mostly main chain N–H groups) in the receptor's binding site (see below). A network of electrostatic bonds, for example, will be an efficient molecular brake when a leukocyte in the bloodstream needs to be brought to a stop to adhere to inflamed endothelium (in the case of the three distinct lectins of the C-type family called selectins), will be molecular glue in transport processes, e.g. apical or axonal sulfatide-dependent glycoprotein routing (for C- and P-type lectins or a galectin) or for contact building to trigger outside-in signaling (for siglecs). This enumeration underscores the potential of glycans to generate affinity, selectivity and specificity by different molecular modes of complementarity in a binding ('reading') process. The inherent requirement for mutual docking let us become aware of another favorable feature of glycans in information coding and transfer that is explained next.

The thermodynamics of the association of a glycan to a receptor would not include a large entropic penalty, if glycans had a low degree of intramolecular flexibility around glycosidic bonds, and this often is the case. Since the conformational space of oligosaccharides then is well-structured like a landscape with energetically privileged valleys, E. Fischer's famous lock-and-key analogy can be applied to view these conformers as bioactive keys.[42] Interestingly, the conversion between conformers of "the bunch of keys – each of which can be selected by a receptor"[43] (and differential conformer selection is a common phenomenon among lectins) is often a rapid process and hereby an impediment to crystallize glycans. When a rather rigid oligosaccharide such as sialyl Le^x meets a preformed docking site in a selectin, association driven by ionic interaction will even have a high k_{on}-value so that we reach the following fundamental conclusion: on the level of molecules and cells, "a universal biological principle, namely, molecular key-lock configuration as a mechanism of selectivity" is operative.[44]

The inherent ability to select, choose or read (*legere* in Latin) was the reason to call (glyco)proteins from plants, which agglutinated erythrocytes depending on their blood group ABH (0) status, lectins[45] (for overviews on lectin history, please see Ref. [46]). However, the biochemical nature of what such a blood group is had been a mystery, and now we connect to α1,2-fucosylation as indicated above to solve it: when performing an inhibition assay as designed for serological reactions,[47] the agglutination of blood group H(0)-positive erythrocytes by eel (*Anguilla anguilla*) serum was most effectively blocked by a sugar derivative, i.e. α-methyl-L-fucopyranoside. Notably, its β-anomer was inactive.[48] The anti-H(0) agglutinin thus is a fucose-specific lectin with specificity for the α(1,2)-linkage that is shown in the bottom panel of Figure 1. This study confirmed and extended previous evidence that a haemagglutinin (in this case the glucose/mannose-specific concanavalin A) interacts with "some compounds present in the surface of such types of erythrocytes as it agglutinates – it is possible that this may be a carbohydrate group in a protein".[49] The recognition of distinct glycan epitopes on the erythrocyte surface and the *trans*-bridging by a phytohemagglutinin, e.g. the blood group A tetrasaccharide by *Dolichos biflorus* agglutinin, seen in such assays proves cell-cell adhesion to be established by glycan-protein (lectin) recognition, with implications far beyond the blood-group typing in transfusion medicine.[50] Owing to the pioneering detection and the purification of the first lectins from mammalian organs by affinity chromatography in 1974 and 1975,[51] it became clear that endogenous lectins are a link between the glycan-based vocabulary and its functional aspects, i.e. the entries into a dictionary of the sugar language. Glycan-based words of this vocabulary can receives their meaning(s) by pairing with lectin(s), experimentally detectable as read-out when measuring biochemical and cell biological post-binding effects (see below).

4. From the Vocabulary to Readers of the Sugar-Encoded Information

The structural unit of each lectin essential for glycan binding is the carbohydrate recognition domain (CRD). The assumption of a fundamental role of lectins in cell physiology by interplay with cellular glycans would be strongly supported, if not only a single type of CRD had developed in evolution. Instead, the diversity of the glycan vocabulary described above would much better be matched by a large pool size of CRDs. Respective analyses on lectin structure, indeed, disclosed that more than a dozen protein folds are able to generate a CRD. These folds are presented in our gallery of human and animal lectins (Supporting Information, Figure S4). The case of the multi-purpose use of the β-sandwich platform (adapted to make contacts to sugar ligands at different sites in the fold without/with the involvement of coordination bonds to protein-bound Ca^{2+}, which help to distinguish epimers at high-level accuracy; the Ca^{2+} of the laminin G-like domain (mostly no. 4 of the five linearly arrayed units) even reaches octahedral coordination with the carboxylate of GlcA and the 4-OH of xylose (Xyl) of the GlcAβ1,3Xyl disaccharide of matriglycan[52]) is a role model: it exemplifies the plasticity of a fold to serve as starting point for the development of different groups of fold-sharing lectins. Historically, this fold was the first to be detected in a lectin when solving the crystal structure of the already mentioned leguminous lectin

concanavalin A and later of the first animal lectin (i.e. galectin-1), in phytohemagglutinins allowing formation of di- and tetramers and the discovery of the importance of the quaternary structure for bioactivity.[53]

During phylogenesis, each CRD is then subject to sequence diversification after duplication events. In general, they can occur within the CRD (establishing more than one binding site per fold as for example seen in β-propeller lectins),[54] for the CRD within the gene for a modular protein (establishing a tandem-repeat arrangement of CRDs) and on the level of the entire gene. Naturally, individual preferences for ligand binding will hereby be shaped. As consequence of gene duplication, each ancestral CRD is the origin of a family of structurally homologous but distinct proteins. Particular sequence motifs, e.g. the Glu-Pro-Asn (or Gln-Pro-Asp) triad in the primary Ca^{2+}-binding site of C-type lectins for Man/Glc*N*Ac (or Gal/Gal*N*Ac) binding (programming coordination bonding to select distinct epimers) or the seven-amino-acid signature (with its Trp for C–H/π-interactions) for ga(lactose-binding)lectins (=galectins) govern direct contact building and are thus conserved (see below). Sequence variations in their local vicinity then implement grading of the fine- and subspecificities so that each member of a lectin family can select its set of binding partners among structurally related glycans. To visualize this point, Figure 5 provides a graphical account on natural β-galactosides to visualize diversity at branch ends of glycans. Of note, the entries of galectin names into the figure inform about preferences (for surveys on galectin contact sites and specificities, please see Ref. [55]).

On the cellular level, the actual features of glycan presentation, e.g. defined by branching, clustering or local vicinity among glycoconjugates in microdomains, will also have a major influence on its bioactivity. For example, multivalency of a glycoconjugate counterreceptor such as a glycoprotein separates the individual loading steps with lectins into a gradient of decreasing binding constants up to reaching full saturation so that the first binding process has the highest affinity; the emerging rule for fractional occupancy thus facilitates to initiate cross-linking (called lattice formation; this term originates from the "lattice" theory of serological reactions, in which an antibody precipitates antigens or agglutinates antigen-bearing cells[56]) of multivalent glycoconjugates by lectins at physiological concentrations despite the often low affinity of a sugar ligand when free in solution.[57] The fine-

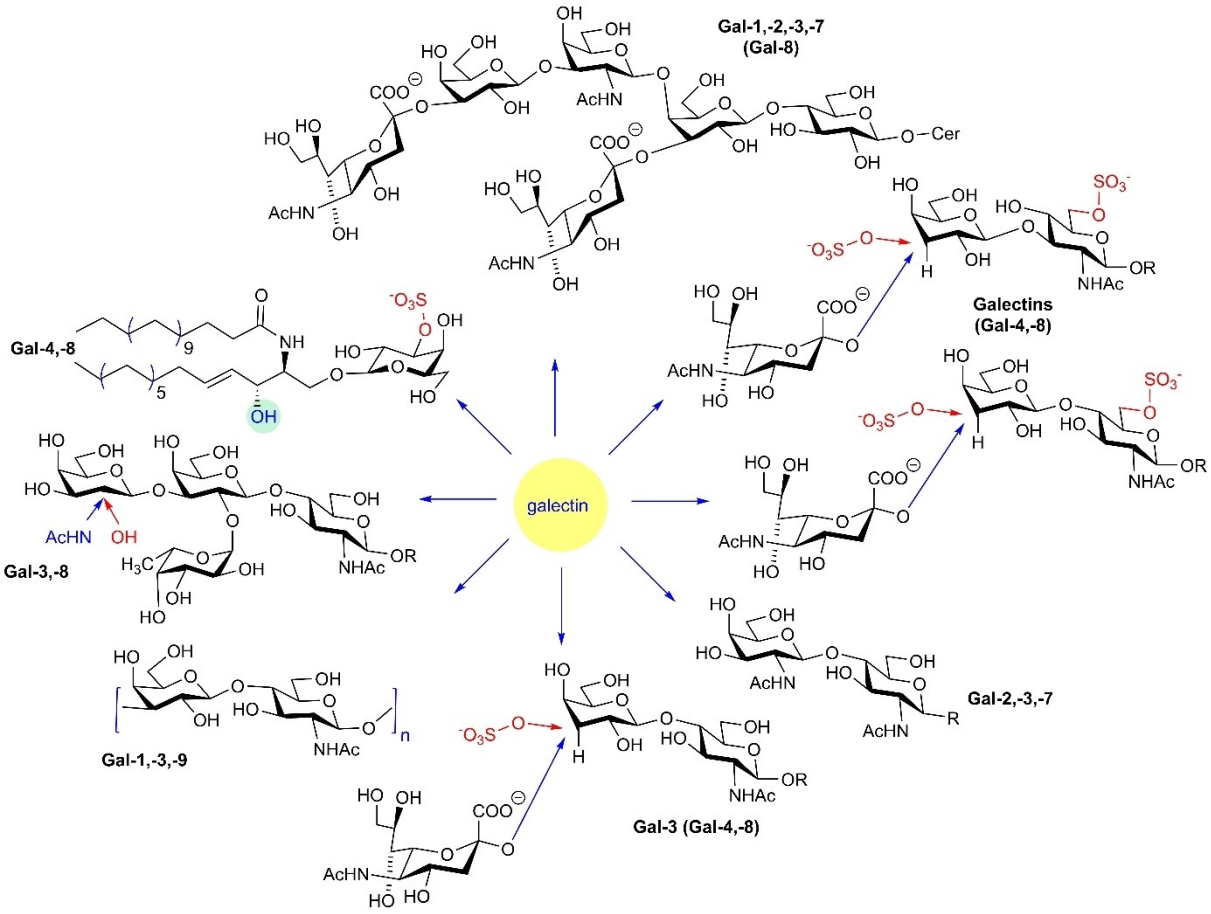

Figure 5. Illustrations of galectin binders from the class of natural β-galactosides and naming of examples of mammalian galectins with preference (galectins in parentheses bind with lesser affinity) for a glycan, for example galectin-8 (Gal-8) for 3′-sulfated Lac*N*Ac and the hexasaccharide of ganglioside GD1a or galectins-1, -2, -3 and -7 (Gal-1, -2, -3 and -7) for the pentasaccharide of ganglioside GM1 (please see Figure 11 for examples of bioactivity of GM1 binding by these adhesion/growth-regulatory galectins).

structural examination of each glycan will identify the contributors to the affinity and selectivity, and they also include non-glycan determinants (see below).

How different types of contact along the glycan structure can team up can be elucidated by using a strategic approach: the chemical synthesis of the binding partners for lectins makes their application in activity assays/structural analysis possible, as noted above when presenting the DiLac*N*Ac synthesis. Hereby, distinct structure-affinity relationships are traced. Sulfation at the 3′-position of Lac*N*Ac for instance adds ionic recognition to binding for two galectins.[58] The resulting interactions are illustrated in Figure 6 (top part): strong affinity of the CRDs of Gal-4 and -8 to 3′-sulfated Lac*N*Ac rests upon the combination of this ionic interaction with the typical hydrogen bonding

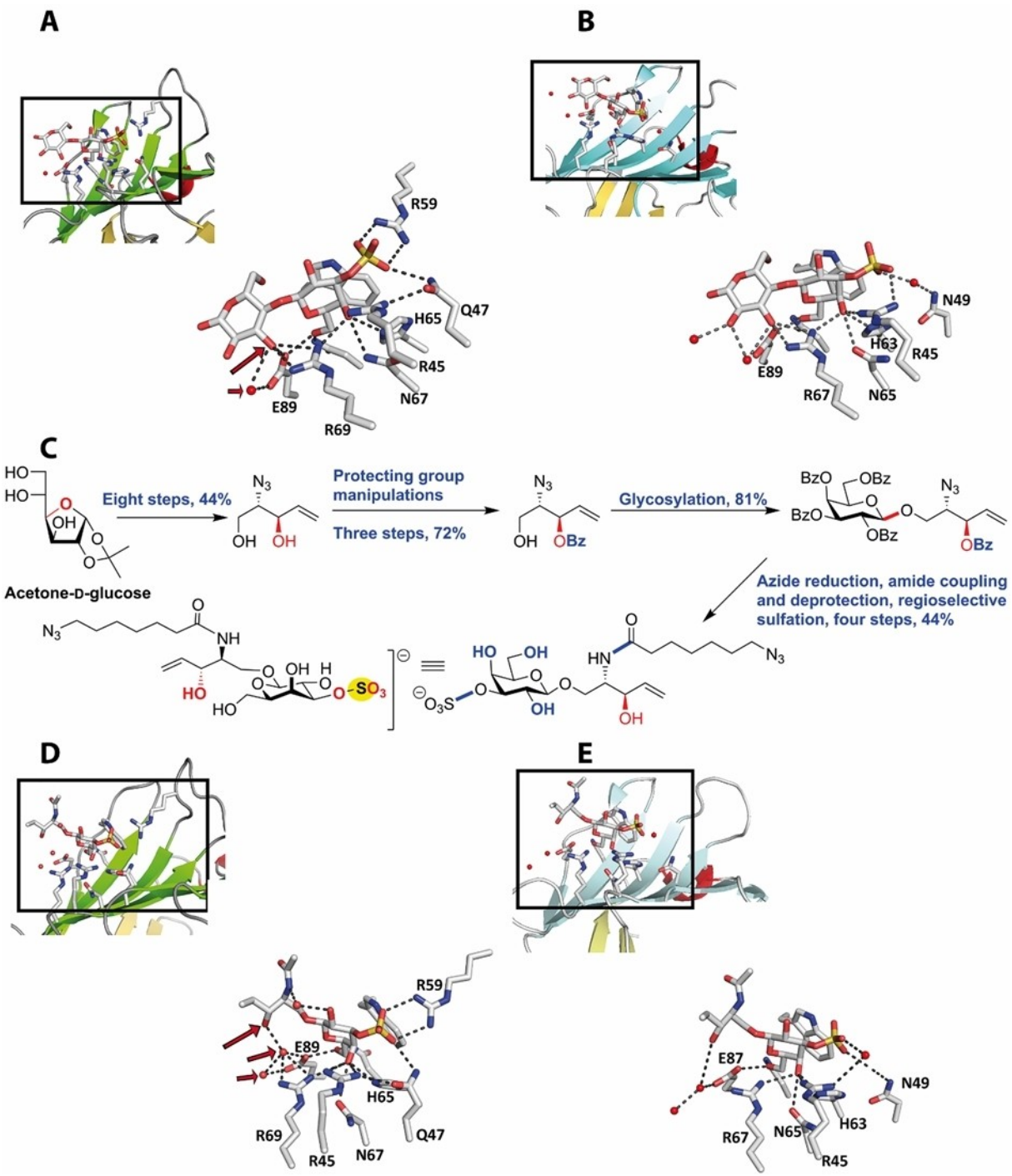

Figure 6. Illustrations of the contact pattern of 3′-sulfated Lac with the N-terminal CRD of human Gal-8 (PDB 3AP6) or Gal-4 (PDB 5DUW) (A, B), of the synthesis of a bioactive derivative of the sulfatide headgroup (for details, please see Supporting Information, Scheme 3) (C) and crystal/modeled structures of its binding profile with the two CRDs, the water-mediated contact to sphingosine's hydroxyl group highlighted by arrows (D, E). For details, please see Ref. [60].

(interestingly, an ionic interaction governs siglec recognition of the negatively charged (sialylated or sulfated) sugar, too, but there is an alternative to it described for the β-trefoil fold: a sulfate can be positioned (by stacking between the B-face of Gal and Trp's indole ring) to become acceptor for many H bonds; for details on these two cases, please see Supporting Information, Figure S5). Interestingly, negatively charged homogalacturonans bind to the Gal-3 CRD in an unconventional orientation with the reducing end GalA β-anomer taking the position of the non-reducing end galactose residue in lactose, and yet maintaining interactions with the conserved tryptophan and seven of the most crucial lactose-binding residues, albeit with different H-bonding interactions.[59] Why Gal-4 and Gal-8 bind the sulfatide headgroup despite the loss of the second sugar unit that contributes to the recognition had long been a riddle.[59]

This mystery, that is strong binding of the sulfatide headgroup despite its truncation to 3-sulfated Gal, has recently been solved by using a synthetic mimetic of the crucial part of the sulfatide, prepared as shown in Figure 6 (middle), and crystallography. Intriguingly, recruiting sphingosine's hydroxyl group and a water molecule to the interaction with galectin, hereby substituting for the 3-OH group of Glc*N*Ac of the disaccharide, brings about sufficient extent of bridging (Figure 6 (bottom part); for details on synthetic procedure of the sulfatide headgroup, please see Supporting Information, Scheme S3).[60] This case of molecular compensation of a loss of a carbohydrate, here Glc*N*Ac, by a non-glycan part documents the possibility for a broader ligand profile for lectins than exclusively binding carbohydrates. This is likewise seen in other classes of lectins such as C-type lectins, for example by detecting extended binding sites to accommodate phosphoglycolipids and especially the cord factor (trehalose-6,6′-dimycolate) of mycobacteria by the dendritic cell (immuno)activating receptor (DCAR) or the macrophage inducible C-type lectin (MINCLE), respectively (for details on post-binding outside-in signaling by MINCLE via an adaptor molecule, see below).[61] Obviously, a look at methods how to define ligand recognition is now warranted.

The analysis of glycan-lectin specificity is performed by a wide range of methods (for an overview including information on analyzed aspects and limitation, please see Table 1 in Ref. [62]). The strategic combination of carbohydrate chemistry with lectin assays has considerably fueled the progress on profiling glycan specificity. It also is a rich source of information on other aspects of the binding process such as the involvement of solvent rearrangement for affinity.

The solvent isotope effect measured by running calorimetric titrations comparatively in H_2O and D_2O first for leguminous lectins, recently initiated for a human C-type lectin and two galectins, indicates an altered solvation within the enthalpically driven thermodynamics best seen when using an oligosaccharide (for the example of the thermodynamics of DiLac*N*Ac-galectin binding in both solvents, please see Figure 2, bottom).[63] Applying diverse types of biophysical methods to study (ga)lectin structures has revealed that a broad-range impact on the protein can ensue from ligand association beyond the solvation of the contact site in certain proteins. Protein-type-dependent changes of surrounding loop regions or of global hydrodynamic properties up to a ligand-induced compaction and an increased internal protein dynamics have for example been detected in the cases of a collectin and human galectins (Supporting Information, Figure S6 shows an example).[64] Under such circumstances, crystallization of a complex with ligand will not be favored. Responses to ligand binding can even be transmitted to other modules beyond the CRD, to the neck domain of C-type lectin oligomers or the EGF-like domain of E- and P-selectins (see below).[65] The modules are therefore more than inert spacers between the cell surface and the CRD.

In addition to using natural glycans, the scope of experiments on interaction analysis can be extended when glycan derivatives with site-specific substitutions (reporter groups) are synthesized. Preparing deoxy- or fluoroderivatives has not only enabled the chemical mapping of sites of contact for hydroxyls of the ligand (for cases of applications on plant and animal lectins; please see Ref. [66]) but also opened the door to proceed from work with ^{13}C-labeled sugars to other isotopes. Adding ^{19}F (and also the ^{77}Se isotope of selenoglycosides) as NMR-spectroscopic sensor in interaction studies (for an example of a synthetic scheme to produce such a probe and of its recent application to analyze bound-state glycan structure(s), please see Figure 7 and Supporting Information, Scheme S4) is a means to map ligand-lectin contacts in solution quantitatively: combining short-range heteronuclear (^{1}H,^{19}F) relay to F (reF) with long-range homonuclear (^{1}H,^{1}H) TOCSY transfer enabled to determine that the dominant contact via one of the terminal residues of the shown trimannoside in the crystal of the lectin-ligand complex occurs at a 2:1 ratio between the α1,3- *vs* α1,6-linked moieties in solution.[63e,67] Evidently, binding modes in solution can be accurately dissected by the help of the ^{19}F sensor. Like the pieces of a puzzle, the data obtained from all such studies are further strengthening the postulated wide-scale ability of constituents of the cellular glycome to be lectin ligands with the noted key-to-lock-style conformer selection in the binding process.

To solidify this fundamental take-home message for tissue lectins, let us look at the first steps of a common route of glycan biosynthesis, i.e. mucin-type *O*-glycans (Figure 8). To make our point, we have inserted respective information: examples of cognate lectins, for example siglecs, are named along with the corresponding glycan for the products of core 1/2 synthesis of mucin-type *O*-glycans (Figure 8; for how outside-in signaling is elicited by such an interaction, see below). Names of lectins had likewise been added to the listing of glycans in the bottom panel of Figure 1 such as the binding of the core-fucosylated *N*-glycan stem by dectin-1, while binding of pauci- and oligomannosidic *N*-glycans by the macrophage mannose receptor and of Glc*N*Acβ1,2Man by the *l*iver and lymph node *s*inusoidal *e*ndothelial *c*ell C-type lec*tin* LSECtin prove this principle to be at work already for not fully mature *N*-glycans.[68] This relationship from words of the vocabulary to distinct 'readers', without and with involving the 'sulfation code', is also emerging for glycosaminoglycans.[69] The huge combinatorial potential offered by epimerase and sulfotransferase activities acting on

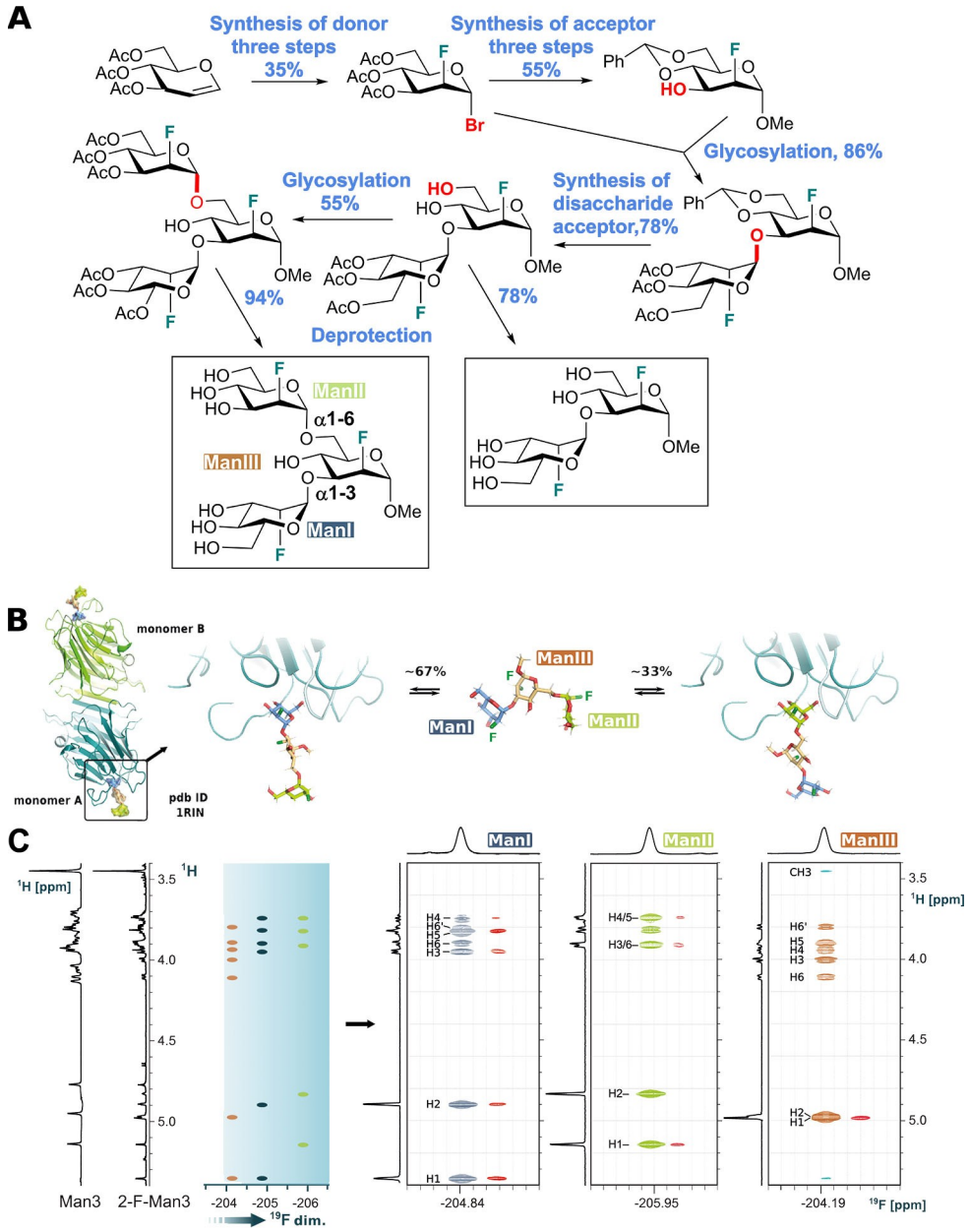

Figure 7. Overview of the synthesis of the trifluorinated *N*-glycan core trimannoside (for further information, please see Supporting Information, Scheme S4) (A), the crystallographic information on trimannoside binding in two modes (PDB 1RIN) (B) and NMR-spectroscopical information on binding of the trifluoro-trimannoside (2F-Man3) by *Pisum sativum* (pea) agglutinin (C); from left to right: 1D ^{1}H of Man3, 1D ^{1}H 2F-Man3; 2D ^{1}H,^{19}F TOCSYreF correlation spectrum; 2D ^{1}H,^{19}F STD TOCSYreF spectra (strips) of 2F-Man3 in the bound state revealing the 2:1 ratio of its two modes of docking via a terminal residue, i.e. the α1,3- or the α1,6-linked Man moiety, respectively (for details, please see Ref. [67d]).

their basic disaccharide units (see Supporting Information, Figure S3 for illustrations) is the basis for a large and further growing interactome with receptors (for recent compilations of 'readers', please see Refs. [69f,g]). These lines of evidence make clear that glycan-protein recognition is a frequently taken route for deciphering a glycan's functional meaning, and this also includes non-self glycans.

As alluded to above, the existence of a discriminatory glycan signature for bacterial surfaces offers the possibility that lectin recognition becomes a means to trace non-self: Gal*f* (and also bacterial ulosonic acids) as such signals are indeed efficiently detected by human intelectin via their common terminal exocyclic 1,2-diols as structural characteristic.[70] Similarly, animal and fungal six-bladed β-propeller-type tectonins such as tachylectin-1 have *O*-methylated glycans (Man/Fuc residues with an equatorial hydroxyl neighboring the methylated position) as conserved target in frontline defense against infection (bacteria) or predators (nematodes), therefore called a universal defense armor.[71] Considering this concept of recognition of an epitope and inhibition of antibody binding by

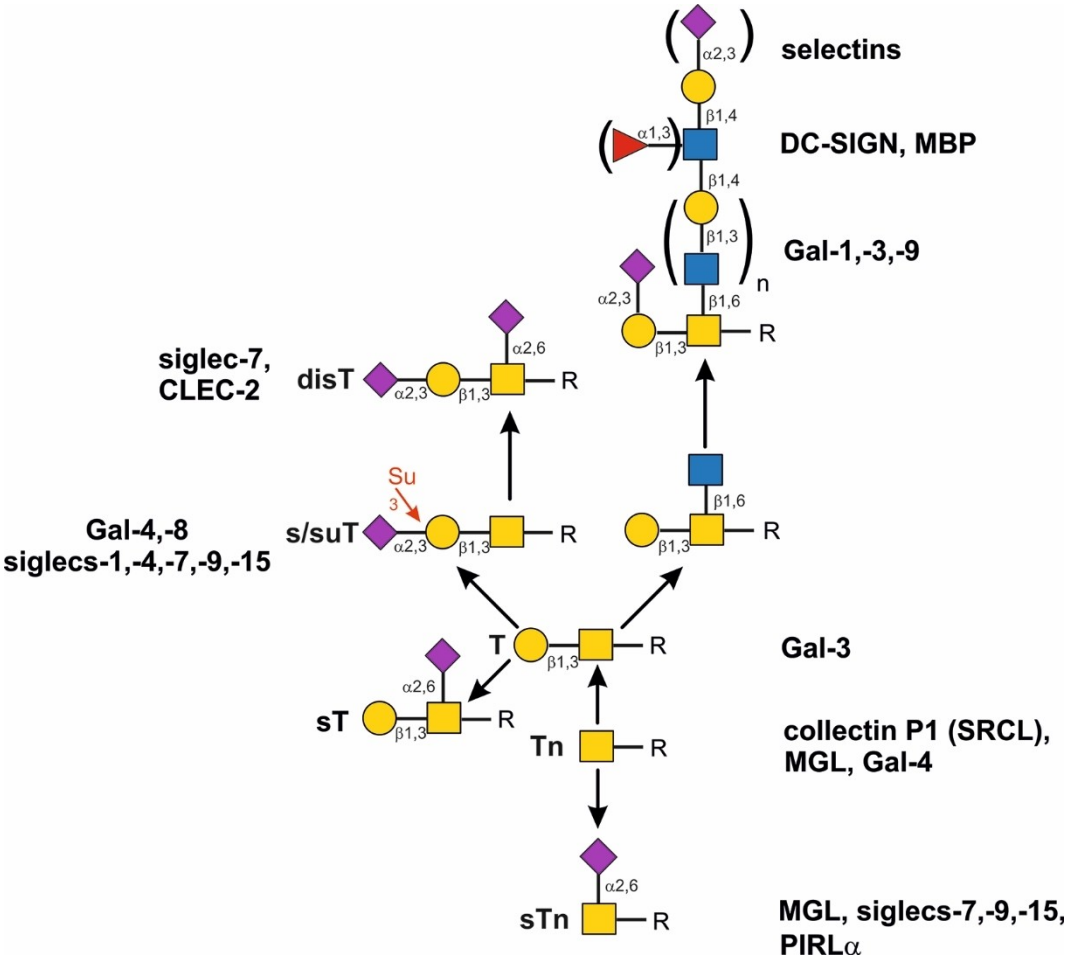

Figure 8. Illustration of routes within mucin-type core 1/2 *O*-glycan biosynthesis. The functional meaning of these words of the glycan vocabulary is indicated by naming of examples for mammalian lectins that bind respective glycans.

haptens, explained above when dealing with fucoside-inhibitable hemagglutination, also helps to solve the mentioned riddle of why the most common sugar, i.e. Glc, is a ligand exclusively for just three lectins in the closed environment of the endoplasmic reticulum, i.e. calnexin, calreticulin and malectin: the "curious absence" of Glc from mature glycoproteins is reasonable, because "the efficiency of a recognition surface based on D-glucosyl components would be impaired by free D-Glc much like haptens interfere with antigen-antibody interactions".[72]

That the ability of a lectin domain can well go beyond binding glycans has been substantiated above by illustrating in detail the case of sphingosine's hydroxyl group as part of a binding partner. Going even further, a separate second site used for molecular rendezvous can be presented by a lectin. The slime mold lectin discoidin I with its glycan-dependent externalization and the fibronectin-like Arg-Gly-Asp motif for cell-matrix interaction has provided a role model for developing and appreciating the concept of lectin bifunctionality.[73] In this case, the two sites are operative at different time points, that is first during the lectin's externalization and then extracellularly after the export. Are cases already known to see them cooperate at the same time?

Among mammalian lectins, the β-sandwich fold of the mentioned galectins with its F- and S-faces equips a lectin to bring two types of counterreceptors together, this by using the second site for specific protein binding, e.g. for the autophagy receptor NDP52 and other organizers of autophagy or mediators of endosomal membrane repair or for the chemokine CXCL12 (experimental data and a structural model of the Gal-3 CRD-chemokine pair are shown in Supporting Information, Figure S7).[74] By the way, this type of CRD also contains molecular switches such as oxidizable Cys or Trp residues or a prolyl peptide bond for *cis-trans* isomerization to swiftly regulate lectin activity or quaternary structure (an example of resonance splitting by such an isomerization process at the prolyl Pro4 peptide bond of human Gal-7 is shown in Supporting Information, Figure S8; the phenomenon of two conformational states and a shift to the *cis*-bond has been discovered for lectins in the case of concanavalin A and its Ca^{2+}-induced isomerization of the Ala207-Asp208 peptide bond, later seen for the two rat mannan-binding proteins (MBPs) and supposed to have a strong bearing on ligand binding, here at

the peptide bond preceding Pro186 (in serum MBP) or Pro191 (in liver MBP)).[75] Deserving particular attention, the lack of a signal peptide and thus cytoplasmic biosynthesis predestine galectins to this role in intracellular surveillance, because they detect otherwise absent *N*-glycans at this location after membrane damage (sensing danger); keeping galectins away from the classical route of secretion also precludes their *N*-glycosylation in the ER that has been shown to impair lattice formation in the case of the engineered version human Gal-1 that enters the ER by having been tagged by a signal peptide.[74e,76]

In summary, the type of the CRD is the common denominator of a lectin family. Having identified sequence signatures, searches for homology by scouring genomes accomplished to reach the full-scale description of lectin families. This sequence mapping disclosed such a wide range of diversity of a structural variable that it was puzzling at first. Now, it is increasingly making sense to give the vocabulary functional meaning by lectin recognition. Speaking of the modularity of lectin architecture, the spatial how of pair building between a lectin and its glycoconjugate counterreceptor is the salient factor toward triggering post-binding processes: they perform the actual translation of a 'word' of the glycome vocabulary into function(s). Thus, lectin design (together with glycan structure, multivalency and type of binding partner) contributes to shape functionality. Consequently, if glycans are functional counter-receptors of tissue lectins, then the number of ways to present a CRD must be large, exactly as we have seen this to be the case for the folds with glycan-binding capacity in our Gallery of Lectins. Figure 9 gives an impression that it is.

A CRD can stand alone or it becomes a part of a molecular puzzle by the association of modules to form homo- or hetero-oligomers, even coming together with other types of domains, covalently or non-covalently (Figure 9). Design diversity is most impressively illustrated by C-type lectins.[77] Bioactivity as anti-microbial protein has already been seen on the level of a C-type CRD in the case of murine RegIIIγ (the human ortholog is called *h*epato*i*ntestinal *p*ancreatic/*p*ancreatis-*a*ssociated *p*rotein (HIP/PAP)).[78] Joining different types of modules is ideal as means to create tools for many purposes. It is essential to allow aggregation for sensing glycans presented in clusters, the origin of the glycoside cluster effect (it is defined as "binding affinity enhancement exhibited by a multivalent carbohydrate ligand over and beyond that expected from the concentration increase resulting from its multivalency", in the case of the trimeric hepatic C-type lectin and mono-, bi- and trivalent oligosaccharides yielding a geometrical (logarithmic) increase in affinity from a numerical increase in valency[79]). The covalent conju-

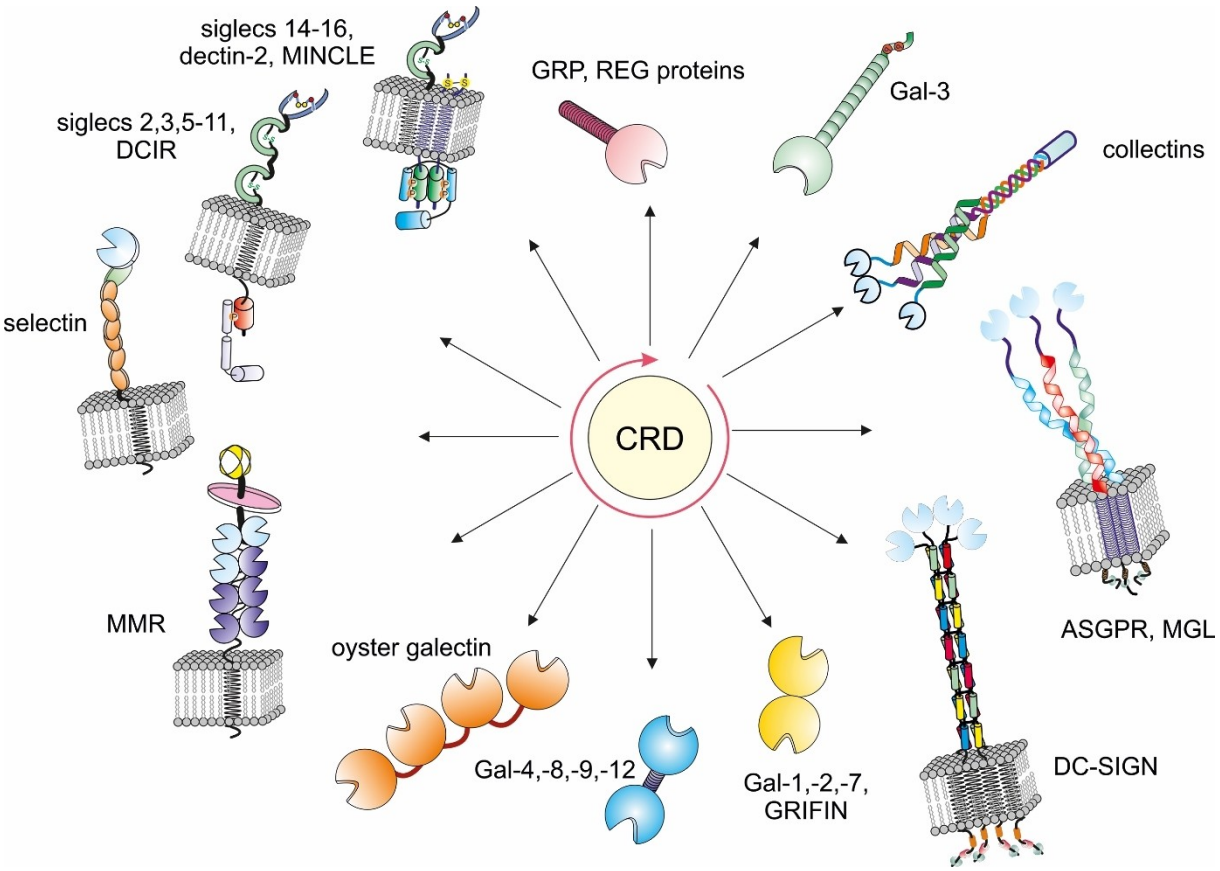

Figure 9. Illustration of examples of lectin design starting with a single CRD that can have a short or long tail (the latter for self-association). In clockwise manner, lectins with modules for covalent subunit association (via disulfide bonds), for non-covalent and linker-mediated modes of CRD associations and for building a puzzle-like architecture with intracellular domains for outside-in signaling are displayed. Abbreviations are given to define distinct lectins for each type of shown architecture (Reproduced with permission from Ref. [98c] Copyright 2015 Elsevier; for further information, please see Refs. [77b,g–k, 82f, 83c,e, 86i, 98]).

gation of CRDs fabricates the tandem-repeat design, it produces molecular tentacles when using other modules to present the CRD on their tips on the surface spatially readily accessible for making crucial contacts in cell-cell bridging, and it adds a place for site-specific phosphorylation/association of an intracellular adaptor in post-binding (outside-in) signaling (Figure 9). That the extracellular matrix adopts its highly ordered structural organization is in part made possible by lecticans, which are glue-like multipurpose tools with a C-type CRD.[80]

Historically, the enormous potential of permutations of CRD specificity with the diversity of modular design has first been realized in studies with plant lectins (agglutinins). Explicitly, the switch from natural tetra- to bivalency by chemical treatment (succinylation or acetylation) was shown to reduce cap formation on murine spleen cells: the type of quaternary structure matters.[81] So the take-home message is the modular design of the lectin is a factor that underlies the intriguing selectivity and specificity of pairing of a lectin with its counter-receptor(s). This process is intimately dependent on the context, giving lectins, glycans and glycoconjugates the fundamental ability to become multi-purpose tools *in vivo*. This raises our curiosity to learn about actual functionality of this interplay, its currently known spectrum and specific cases.

5. From the Vocabulary and the Readers to the Dictionary of the Sugar Code

A dictionary of the sugar language is supposed to correlate structural aspects of a glycan and of a lectin with a cellular function. The current status of knowledge on lectin functions is summarized in Supporting Information, Table S1 and listing general terms there calls for illustrating a specific case: Figure 10 presents a route from upregulated lectin expression to manifestation of a common disease with large socioeconomic impact, i.e. osteoarthritis.

The examples of how glycans and lectins cooperate in the already mentioned processes of leukocyte adhesion during inflammation (by selectins) or postal-code-like routing of distinct glycoproteins similarly supply information on the underlying intimate interplay and are thus outlined here. The mentioned high k_{on}-rates of the association of negatively charged sugars such as a sialyl Le^x epitope (most active with 6-sulfation of the Glc*N*Ac moiety) to a (selectin) CRD presented at the tip of the tentacle-like design that hereby reaches out into the bloodstream (see Figure 8) will make nearly immediate contact to this counterreceptor to slow down cells to a rolling on the endothelium to let integrins tighten the grip in the next

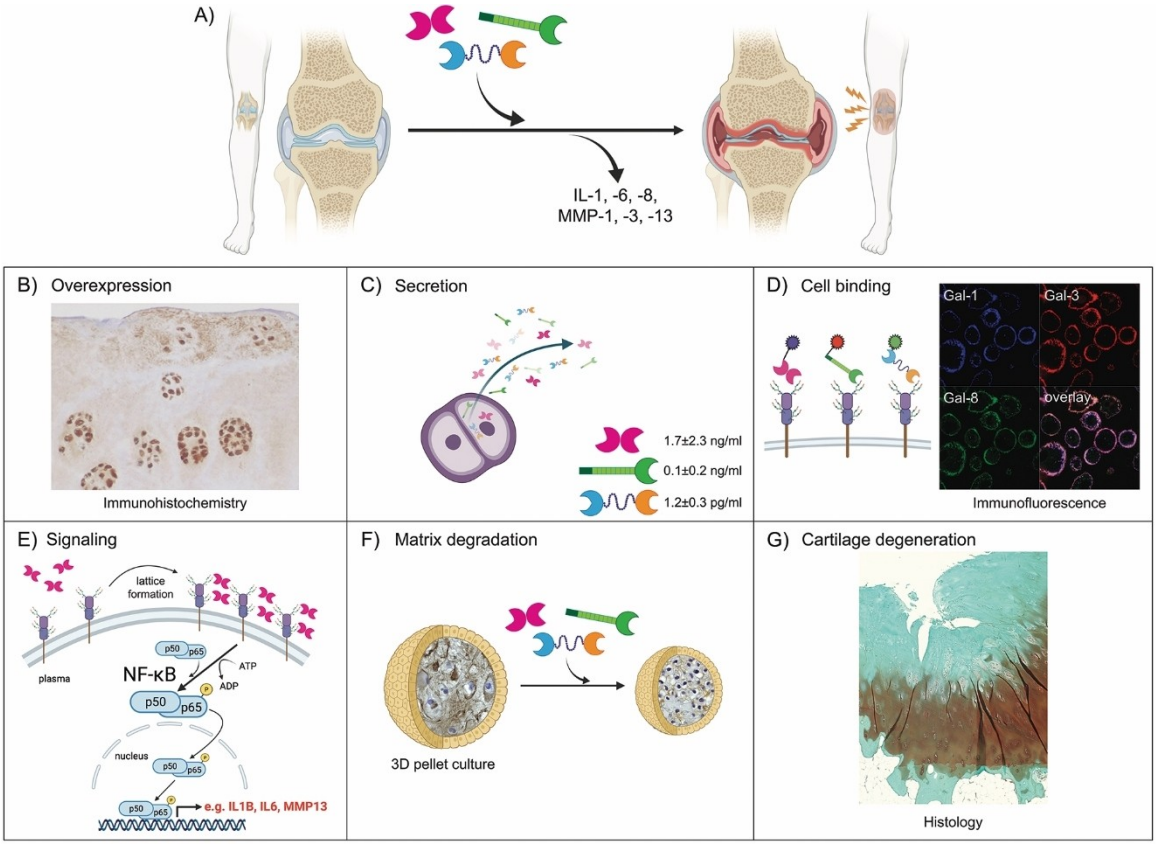

Figure 10. Illustration of the route of galectin-driven osteoarthritis pathogenesis by upregulation of pro-degradative/-inflammatory effectors such as interleukins (IL) and matrix metalloproteinases (MMPs) that starts with dysregulated galectin expression. Their secretion, cell surface binding and the triggered outside-in signaling to reprogram IL/MMP gene expression via a downstream effector, i.e. the transcription factor NK-κB, lead to matrix degradation *in vitro* and *in vivo* (for details, please see Ref. [99]).

step; triantennary *N*-glycans after α2,3-desialylation (for the asialoglycoprotein receptor), hybrid- or high-mannose-type *N*-glycans with Man-6-phosphate (for P-type lectins), the 4′-sulfated Gal*NAc*β1,4Glc*NAc*β1,R (Lacdi*NAc*) unit of *N*-glycans (for the contact site of the β-trefoil domain of the (macrophage) mannose receptor, see Supporting Information, Figure S5 (bottom)) and Lac*NAc*-terminated *N*-glycans of cargo glycoproteins together with the sulfatide headgroup (for the heterobivalent galectin-4) are postal codes written in glycans for lectin-specific routing and delivery.[50f,79c,82]

That post-binding signaling triggered by glycan-lectin pairing becomes pathophysiologically relevant has been docu-

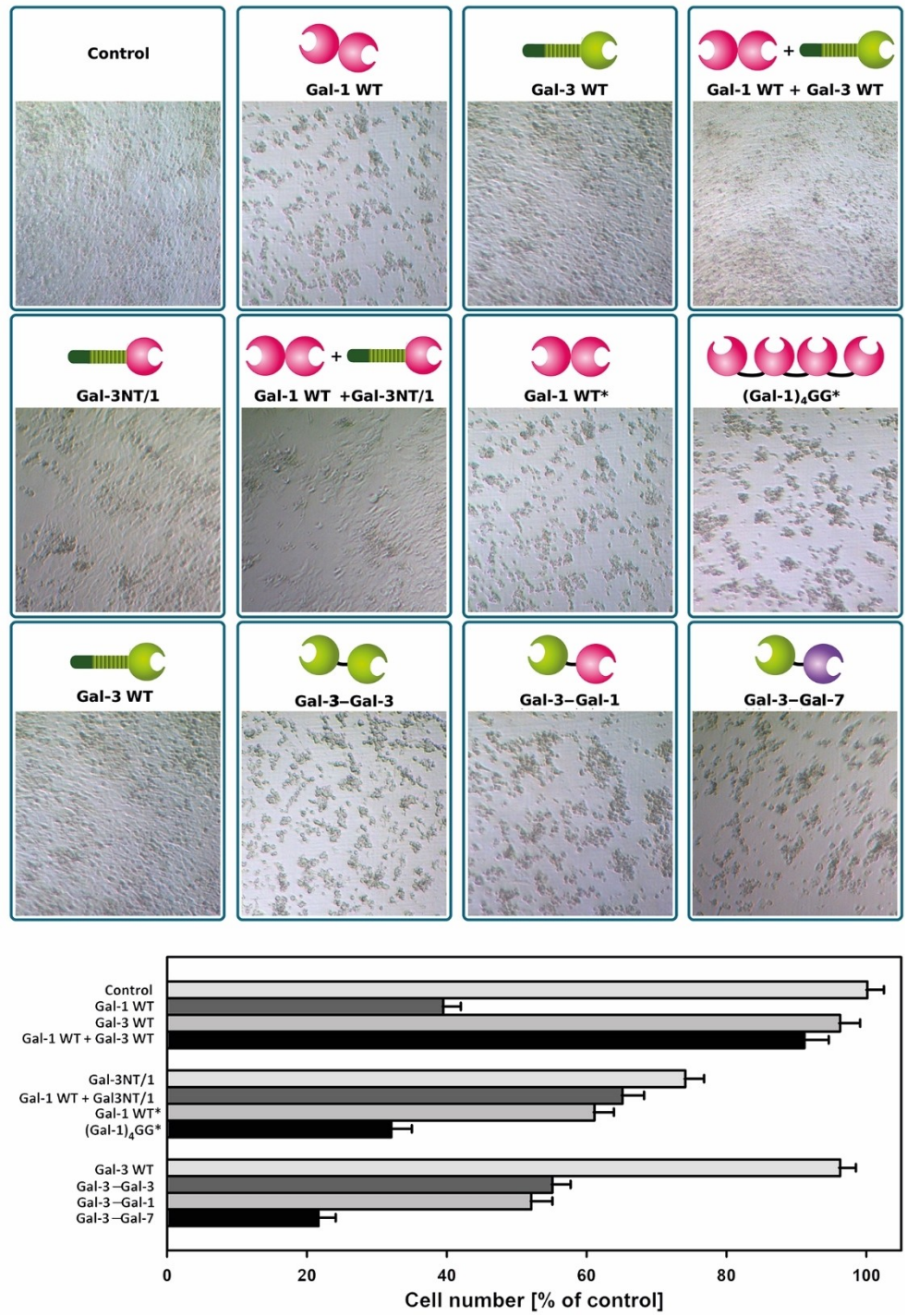

Figure 11. Effect of wild-type and of engineered human galectins on neuroblastoma cell (SK–N–MC) growth. Galectin architecture, microphotographs of representative cultures and a bar graph of cell numbers are shown. Galectins are tested at 100 μg/mL (*10 μg/mL), wild-type Gal-3 and its Gal-3NT/1 variant are used in 10fold excess in the mixtures with Gal-1 (for details on proteins, impact of architecture on lattice formation by testing synthetic glycoclusters and assay conditions, please see Refs. [95c,d,100]).

mented in various cases, and the processes leading to progression of osteoarthritis driven by galectins shown in Figure 10 give an instructive example. In order to avoid falling into the trap of thinking that lectin activities can simply be extrapolated, the context-dependent nature of response patterns definitely precludes extrapolations (we all know so well that the meaning of a word in any language can be contingent on the context of the sentence). Literally, the same lectin can hereby elicit opposite effects, for example pro- or anti-inflammatory or -tumoral activities. This inherent potential for duality warrants attention when considering pharmacological targeting of tissue lectins without reaching site-specific delivery. The recurring theme thus is that context matters.

Next, the following specific cases of *O*-glycans shown in Figure 8 underpin the principle that adding a symbol reprograms a glycan word's meaning profoundly: the addition of a sialic acid to mucin-type *O*-glycans can establish the basis for association to a member of the siglec family. Binding the sialylated core 1 disaccharide (also called the Thomsen-Friedenreich (TF) antigen or CD176) by siglec-7 converts monocytes into tumor-associated macrophages, binding the disialyl core 1 tetrasaccharide when presented clustered on leukemia cells by the glycoprotein CD43 primes the intracellular immunoreceptor tyrosine-based inhibitory motif of siglec-7 to convey negative signaling to reduce killing activity (on the other hand, the high-affinity interplay between 6′-sulfo sialyl Le^x and siglec-8 is an example for a self-glycan code-guided return to homeostasis after inflammation by depleting eosinophils from tissues), whereas sialyl T_n (CD175s) binding by siglec-15 (T_n, *O*-linked α-Gal*N*Ac (n = nouvelle), is not active) transmits a positive signal via association with *D*NAX-*a*ctivating *p*rotein of *12* kDa (DAP12) in the transmembrane region to increase spleen tyrosine kinase (Syk) activity and hereby transforming growth factor (TGF)-β secretion by tumor-associated macrophages.[83] A schematic drawing on the special route of signal transfer from the out- to the inside by association with adaptor molecules (DAP12 or the F_c receptor γ chain ($F_cR\gamma$)) that contain an *i*mmunoreceptor *t*yrosine *a*ctivation-like *m*otif (ITAM)) is shown in Figure 9. Not surprising, this mechanism is also operative for several C-type lectins like the already mentioned MINCLE (with $F_cR\gamma$).[84] Bidirectional (*cis* and *trans*) signaling between axon and myelin, to give a further example, is exerted by pairing of the sialyl core 1 trisaccharide present on gangliosides such as GD1a or GT1b with myelin-associated glycoprotein (siglec-4a), which then appears to favor lectin dimerization and to associate the cytoplasmic non-receptor tyrosine kinase Fyn as relay station when alone or as a heterotetramer with the dynein light chain.[85]

Research over decades has shown that cellular activation or differentiation, inflammation or the activity status of distinct genes such as oncogenes or tumor suppressor genes induce a reprogramming of aspects of glycosylation. For example, the extent of sialylation or of β1,6-branching and the occurrence of Lac*N*Ac repeats in *N*- and *O*-glycans, of sialyl $Lewis^x$ production or of conversion of ganglioside GD1a into the galectin counter-receptor GM1 are modulated. The concept of the sugar code predicts the possibility of an *in situ* interplay with tissue lectins, and, to reach a high level of an effect, lectin availability at the right place could be regulated in a coordinated manner. This has already been revealed to occur for selectins and galectins-1, -3 and -7.[33g,86] Serving as a proof-of-principle case, anoikis induction in pancreas cancer *in vitro* by the tumor suppressor $p16^{INK4a}$ is based on orchestrating a downregulation of sialic acid biosynthesis (at enzyme (NANS and GNE) expression level) and hereby of *N*-glycan α2,6-sialylation (that precludes galectin binding by occupying the OH group of the hydroxylmethyl of Gal, a major contact point) with an upregulation of both galectin-1 and its glycoprotein counterreceptor $\alpha_5\beta_1$-integrin, fully suited for integrin cross-linking without α2,6-sialylation, so that focal adhesion kinase and, further downstream, caspase-8 activation will drive these tumor cells into death.[86f] As noted above and worth to be emphasized, the glycan profile and structure, its local density and context-specific mode of presentation are the parameters for enabling a glycoconjugate to become the local counterreceptor for a lectin,[87] its nature such as an integrin and the type of lectin architecture then determining the post-binding effects. Already a perturbation of the glycan profile by a single *N*-glycan, as shown in the gain-of-glycan Thr160Asn mutant of the interferon-γ receptor 2 subunit, can have a tremendous impact, here it partitions the glycoprotein differently among compartments via galectin binding and thus impairs receptor functionality.[88]

Noteworthy in this context, adding a sugar to a certain acceptor can have a second consequence besides being a part of the region for lectin binding. Such a shift in the glycome can also make its presence felt by precluding synthesis of a glycan that is a lectin ligand. Figure 8 guides to this insight. The α2,6-sialylation of the T_n epitope, for example, also abrogates the generation of core 1/2 glycans, the presence of sialic acid preventing any glycosyltransferase from accepting sT_n as substrate (Figure 8). Since the core 1 disaccharide, is a Gal-3 binder and assumedly involved in cell contacts in the metastatic cascade,[89] the respective enzymatic activity (e.g. ST6Gal*N*Ac-II or IV) has been discussed as potential metastasis suppressor, by shifting product presence away from T(F) to T_n.[90] Alternatively, the sialic acid can occupy a crucial contact to block galectin binding: α2,6-sialylation does so at *N*-glycan termini (which is the case to avoid Gal-1-dependent induction of anoikis in tumor cells), in turn generating siglec binders (see above). Hereby, the rule for intimate correspondence between the glycan vocabulary and the meaning of its words is further demonstrated, reinforcing the case for a fundamental principle and the feasibility to set up a dictionary for the glycan vocabulary.

This principle also works wonders on a common acceptor in *N*-glycans, that is a Glc*N*Ac-terminated branch. Its alternative usage as substrate leads to words with separate meanings along the different routes. Briefly, when we look at the mentioned case of the generation of the Lacdi*N*Ac platform by Gal*N*Ac (not Gal) addition to Glc*N*Ac in an *N*-glycan branch end, the Pro-Leu-Arg-Ser-Lys-Lys recognition determinant of the glycoprotein in the vicinity of the *N*-glycosylation site accounts for already noted target specificity of this process and then 4′-sulfation follows to yield the mentioned routing (postal-code-like) signal, whereas α2,6-sialylation or α1,3-fucosylation of the

acceptor in other glycoproteins are possible.[91] Overall, interpreting glycome representation and shifts between usage of the vocabulary with a dictionary of the sugar language at hand will let more discoveries appear to be in store by respective investigations.

The implied relevance of the modular architecture of lectins shown in Figure 9 has already been revealed in diverse ways physiologically, as the entries in Table S1 attest. Clearly, CRDs alone would not be able to create such a large panel of bioactivities. The catch-bond phenomenon to let the strength of cell binding counterintuitively increase in sheer stress would be impossible, if ligand binding and the following change in orientation of the modular arrangement would not prepare a selectin (or a bacterial adhesin) to withstand even the influence of external force in its function as molecular anchor with clinical significance in defense (or in infection).[92] Notably, the way bacterial and fungal adhesins as well as viral haemagglutinins convert host glycans to docking sites for infection is a dark side of sugar coding.[93] Associations of uropathogenic *E. coli* or *Fusobacterium nucleatum* mediated by the *O*-glycan core 1 disaccharide and of *Helicobacter pylori* by *O*-glycan-presented Lewis epitopes give these 'words' further meanings to be added to the dictionary.[94]

What Figure 9 teaches us beyond documenting Nature's ingenuity in protein design is the large uncharted territory ahead of us to unveil the full significance of the known types of modular architecture and to unravel activity profiles of new types of design. The latter challenge is addressed by applying rational protein engineering to find answers: the merit of this approach is documented here with experimental data. By using galectins as proof-of-principle models in cell growth assays, first functional antagonism is seen between human Gal-1 and -3 (Figure 11; for details, please see legend). They compete for the same counterreceptor but differ in modular design so that the architecture of the lattice will look differently. Members of the same lectin family can thus interfere with each other in a certain cellular context, in contrast to the cooperation seen in osteoarthritis pathogenesis above (Figure 10). CRD switching by engineering demonstrates the importance of protein architecture (Figure 11).

Next, increased activity of an engineered tetramer relative to the homodimer is revealed (Figure 11). This result lets us wonder why no human galectin has adopted this type of modular design (the answer is that the tetramer's high affinity would sense already low-level ganglioside GM1 presence, making its assignment as molecular switch impossible). Finally, the potential for covalently linked heterodimer variants to exert higher activity than wild-type proteins is sketched, supporting physiological significance of heterodimers (Figure 11; an example for occurrence of galectin heterodimers in mixtures of Gal-7 together with the galectin-3 CRD by CRD switching is given in Supporting Information, Figure S9). These data embody the attractive perspectives for obtaining i) further understanding of structure-function relationships and ii) new reagents for biomedical applications by tinkering with a toolbox of human CRDs and other modules (for information on concept and on details, please see Ref. [95]).

6. Conclusions and Perspectives

A close inspection of the properties of sugars indeed proves that they are ideal symbols for a code. Cooperation by writers, editors and erasers establishes a large vocabulary by using these letters. Molecular complementarity by combinations of coordination, hydrogen and ionic bonding, C–H/π-interaction and stacking underlies the reading. Like glycans, sugar receptors (lectins) come in many forms, more than a dozen protein folds endowing sugar binding to the proteins of the lectin superfamily. The sheer size of sequence changes among CRDs as well as of the diversity of quaternary structures and of types of modular design equips the lectin toolbox with enormous possibilities for selectively interacting with cellular glycoconjugates and for eliciting meaningful post-binding events, the equivalent of the translation of a message. Hereby, the vocabulary is turned into a dictionary of the sugar code. Notably, a glycan 'word' can have different meanings depending on the context, as some ambiguity occurs in a language.

The emerging insights, to keep this part short and sweet, are sure to guide us to novel hypotheses and to a more thorough understanding of cellular systems. For example, powered by hypothesis-driven tinkering with glycan or lectin features, rational engineering can spawn new tools for applications, e.g. biomedically active lectin variants with non-natural architecture as platform for CRD presentation. These data also let us realize that and how the three alphabets of life are going hand in hand in the flow of biological information. Each is suited to meet special needs for life, each is a code system. Compelling evidence is thus available to let the term 'sugar code' reach common parlance. Turning back to the introductory statement by N. Sharon, he concluded his lecture by stating that it is his hope "that I have convinced you why this field is of such great importance, and why it is so exciting".[1]

Acknowledgements

All authors acknowledge the major contribution by Hans-Joachim Gabius to the preparation of this review, particularly in his leadership. Hans-Joachim, tragically, passed away unexpectedly on August 2nd 2021, after submission of the revised version of this manuscript – he was to have been the corresponding author. His passing has been a huge loss to the glycoscience field, to his many collaborators over many years, to his colleagues in Munich and to his family members. He was truly a believer in the sugar code. All of his independent scientific life, as attested by examples chosen in the review, was dedicated to deciphering this code. May he rest in peace.

Funding by the NIH grant CA242351 (to M.C.), the NSERC grant 707321 (to R.R.), the SFI Investigator Programme Awards 16/IA/4419 (to P.V.M.) and 13/IA/1959 & 16/RC/3889 (to S.O.), the AFOR (Association For Orthopaedic Research) foundation grant and the Johnson&Johnson Medical Products GmbH scholarship GMAFS20512 (both to S.T.), as well as by the grant BFU 2016-77835-R of the Spanish Ministry of Economy, Industry and

Competitiveness (to A.R.), the valuable recommendations by the reviewers and inspiring discussions with Drs. B. Friday, A. Leddoz and A. W. L. Nose are gratefully acknowledged. Open Access funding enabled and organized by Projekt DEAL.

Conflict of Interest

The authors declare no conflict of interest.

Keywords: adhesion · glycoproteins · glycosylation · lectins · proliferation

[1] N. Sharon, Complex Carbohydrates: Their Chemistry, Biosynthesis, and Functions, Addison-Wesley, Reading, **1975**.
[2] B. K. Brandley, R. L. Schnaar, *J. Leukocyte Biol.* **1986**, *40*, 97–111.
[3] N. Sharon, H. Lis, *Science* **1989**, *246*, 227–234.
[4] E. Eichwald, *Justus Liebigs Ann. Chem.(Eur. J. Org.Chem.)* **1865**, *134*, 177–211.
[5] a) G. Ledderhose, *Ber. Dtsch. Chem. Ges.* **1876**, *9*, 1200–1201; b) G. Ledderhose, *Z. Physiol. Chem.* **1879**, *2*, 213–227; c) G. Ledderhose, *Z. Physiol. Chem.* **1880**, *4*, 139–159; d) H. Merzendorfer, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 217–229.
[6] a) V. Ginsburg, E. F. Neufeld, *Annu. Rev. Biochem.* **1969**, *38*, 371–388; b) J. Montreuil, in *Glycoproteins* (Eds.: J. Montreuil, J. F. G. Vliegenthart, H. Schachter), Elsevier, Amsterdam, **1995**, pp. 1–12; c) H. Lis, N. Sharon, *Eur. J. Biochem.* **1993**, *218*, 1–27; d) M. Aebi, *Biochim. Biophys. Acta* **2013**, *1833*, 2430–2437; e) H. Nakagawa, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 71–83.
[7] H. S. Bennett, *J. Histochem. Cytochem.* **1963**, *11*, 14–23.
[8] a) S. Roseman, *Chem. Phys. Lipids* **1970**, *5*, 270–297; b) P. J. Winterburn, C. F. Phelps, *Nature* **1972**, *236*, 147–151; c) G. M. W. Cook, *J. Cell Sci.* **1986**, *Suppl. 4*, 45–70; d) O. P. Bahl, in *Glycoconjugates: Composition, Structure and Function* (Eds.: H. J. Allen, E. Kisailus), Marcel Dekker, New York, **1992**, pp. 1–12; e) G. W. Hart, *Curr. Opin. Cell Biol.* **1992**, *4*, 1017–1023; f) G. M. W. Cook, *Glycobiology* **1995**, *5*, 449–461; g) G. Reuter, H.-J. Gabius, *Cell. Mol. Life Sci.* **1999**, *55*, 368–422; h) R. G. Spiro, *Glycobiology* **2002**, *12*, 43R–56R; i) R. S. Haltiwanger, J. B. Lowe, *Annu. Rev. Biochem.* **2004**, *73*, 491–537; j) A. P. Corfield, M. Berry, *Trends Biochem. Sci.* **2015**, *40*, 351–359; k) H.-J. Gabius, J. Roth, *Histochem. Cell Biol.* **2017**, *147*, 111–117; l) A. Varki, *Glycobiology* **2017**, *27*, 3–49; m) K. Moriwaki, F. K. M. Chan, E. Miyoshi, *J. Biochem.* **2021**, *169*, 643–652; n) J. Y. Zhou, B. A. Cobb, *Annu. Rev. Immunol.* **2021**, *39*, 511–536.
[9] S. H. Barondes, in *The History of Neuroscience in Autobiography, Vol. 5* (Ed.: L. R. Squire), Academic Press, Orlando, **2006**, pp. 1–45.
[10] a) N. Sharon, *Acta Anat.* **1998**, *161*, 7–17; b) H.-J. Gabius, *Naturwissenschaften* **2000**, *87*, 108–121; c) H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero, H. Rüdiger, *ChemBioChem* **2004**, *5*, 740–764; d) K.-i. Kasai, *Trends Glycosci. Glycotechnol.* **2006**, *18*, 93–102; e) H.-J. Gabius, *Folia Biol. (Praha)* **2017**, *63*, 121–131; f) A. C. Broussard, M. Boyce, *Mol. Biol. Cell* **2019**, *30*, 525–529; g) H. Kaltner, J. Abad-Rodríguez, A. P. Corfield, J. Kopitz, H.-J. Gabius, *Biochem. J.* **2019**, *476*, 2623–2655; h) S. Dedola, M. D. Rugen, R. J. Young, R. A. Field, *ChemBioChem* **2020**, *21*, 423–427; i) F. A. Habermann, H. Kaltner, A. M. Higuero, G. García Caballero, A.-K. Ludwig, J. C. Manning, J. Abad-Rodríguez, H.-J. Gabius, *Acta Histochem. Cytochem.* **2021**, *54*, 31–48.
[11] H. Rüdiger, H.-J. Gabius, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 3–13.
[12] G. W. Hart, *J. Biol. Chem.* **2013**, *288*, 6903.
[13] S. I. Miller, *Cold Spring Harbor Symp. Quant. Biol.* **1987**, *52*, 14–27.
[14] J. Hirabayashi, *Q. Rev. Biol.* **1996**, *71*, 365–380.
[15] S. Oscarson, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 31–51.
[16] a) D. J. Becker, J. B. Lowe, *Biochim. Biophys. Acta* **1999**, *1455*, 193–204; b) M. K. Wild, K. Lühn, T. Marquardt, D. Vestweber, *Cells Tissues Organs* **2002**, *172*, 161–173; c) X. Wang, J. Gu, E. Miyoshi, K. Honke, N. Taniguchi, *Methods Enzymol.* **2006**, *417*, 11–22; d) T. Hennet, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 365–383; e) K. Honke, N. Taniguchi, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 385–401; f) A. Etzioni, *Clin. Rev. Allergy Immunol.* **2010**, *38*, 54–60; g) E. S. Harris, A. S. Weyrich, G. A. Zimmerman, *Curr. Opin. Hematol.* **2013**, *20*, 16–25; h) T. Hennet, J. Cabalzar, *Trends Biochem. Sci.* **2015**, *40*, 377–384.
[17] a) T. de Vries, R. M. Knegtel, E. H. Holmes, B. A. Macher, *Glycobiology* **2001**, *11*, 119R–128R; b) D. J. Becker, J. B. Lowe, *Glycobiology* **2003**, *13*, 41R–53R; c) B. Ma, J. L. Simala-Grant, D. E. Taylor, *Glycobiology* **2006**, *16*, 158R–184R; d) A. P. Moran, *Carbohydr. Res.* **2008**, *343*, 1952–1965; e) J. D. Aplin, C. J. Jones, *Glycobiology* **2012**, *22*, 470–478.
[18] a) M. Schneider, E. Al-Shareffi, R. S. Haltiwanger, *Glycobiology* **2017**, *27*, 601–618; b) B. C. Holdener, R. S. Haltiwanger, *Curr. Opin. Struct. Biol.* **2019**, *56*, 78–86.
[19] a) P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin, V. Ferrieres, *Carbohydr. Res.* **2008**, *343*, 1897–1923; b) M. R. Richards, T. L. Lowary, *ChemBioChem* **2009**, *10*, 1920–1938; c) B. Tefsen, A. F. Ram, I. van Die, F. H. Routier, *Glycobiology* **2012**, *22*, 456–469.
[20] E. Staudacher, *Biol. Chem.* **2012**, *393*, 675–685.
[21] R. A. Laine, in *Glycosciences: Status and Perspectives* (Eds.: H.-J. Gabius, S. Gabius), Chapman & Hall, London, **1997**, pp. 1–14.
[22] I. A. Roundtree, M. E. Evans, T. Pan, C. He, *Cell* **2017**, *169*, 1187–1200.
[23] a) S. Hickman, E. F. Neufeld, *Biochem. Biophys. Res. Commun.* **1972**, *49*, 992–999; b) M. L. Reitman, A. Varki, S. Kornfeld, *J. Clin. Invest.* **1981**, *67*, 1574–1579; c) A. Waheed, A. Hasilik, M. Cantz, K. von Figura, *Hoppe-Seyler's Z. Physiol. Chem.* **1982**, *363*, 169–178.
[24] a) R. Schauer, M. Wember, *Hoppe-Seyler's Z. Physiol. Chem.* **1971**, *352*, 1282–1290; b) L. Shaw, R. Schauer, *Biol. Chem. Hoppe-Seyler* **1988**, *369*, 477–486; c) G. Reuter, H.-J. Gabius, *Biol. Chem. Hoppe-Seyler* **1996**, *377*, 325–342; d) M. Münzel, D. Globisch, T. Carell, *Angew. Chem. Int. Ed.* **2011**, *50*, 6460–6468; *Angew. Chem.* **2011**, *123*, 6588–6596; e) R. Schauer, J. P. Kamerling, *Adv. Carbohydr. Chem. Biochem.* **2018**, *75*, 1–213.
[25] S. Roseman, *J. Biol. Chem.* **2001**, *276*, 41527–41542.
[26] R. L. Schnaar, Y. C. Lee, *Glycobiology* **2017**, *27*, 513–517.
[27] a) P. Stanley, *Annu. Rev. Genet.* **1984**, *18*, 525–552; b) S. K. Patnaik, P. Stanley, *Methods Enzymol.* **2006**, *416*, 159–182; c) T. K. Dam, C. F. Brewer, *Glycobiology* **2010**, *20*, 1061–1064; d) S. Takamatsu, A. Antonopoulos, K. Ohtsubo, D. Ditto, Y. Chiba, D. T. Le, H. R. Morris, S. M. Haslam, A. Dell, J. D. Marth, N. Taniguchi, *Glycobiology* **2010**, *20*, 485–497; e) M. N. Ismail, E. L. Stone, M. Panico, S. H. Lee, Y. Luu, K. Ramirez, S. B. Ho, M. Fukuda, J. D. Marth, S. M. Haslam, A. Dell, *Glycobiology* **2011**, *21*, 82–98; f) P. Stanley, *J. Mol. Biol.* **2016**, *428*, 3166–3182.
[28] a) J. F. G. Vliegenthart, J. Montreuil, in *Glycoproteins* (Eds.: J. Montreuil, J. F. G. Vliegenthart, H. Schachter), Elsevier, Amsterdam, **1995**, pp. 13–28; b) A. Dell, *Aust. J. Chem.* **2002**, *55*, 27–37; c) D. Zhou, *Curr. Protein Pept. Sci.* **2003**, *4*, 1–9; d) K. W. Moremen, M. Tiemeyer, A. V. Nairn, *Nat. Rev. Mol. Cell Biol.* **2012**, *13*, 448–462; e) A. Togayachi, H. Narimatsu, *Trends Glycosci. Glycotechnol.* **2012**, *24*, 95–111; f) A. P. Corfield, *Histochem. Cell Biol.* **2017**, *147*, 119–147; g) N. Suzuki, *Glycobiology* **2018**, *29*, 625–644.
[29] a) O. T. Avery, W. F. Göbel, *J. Exp. Med.* **1929**, *50*, 533–550; b) W. F. Göbel, O. T. Avery, *J. Exp. Med.* **1929**, *50*, 521–531; c) C. R. McBroom, C. H. Samanen, I. J. Goldstein, *Methods Enzymol.* **1972**, *28*, 212–219; d) C. P. Stowell, Y. C. Lee, *Adv. Carbohydr. Chem. Biochem.* **1980**, *37*, 225–281; e) J. D. Aplin, J. C. Wriston Jr., *Crit. Rev. Biochem.* **1981**, *10*, 259–306; f) H.-J. Gabius, *Angew. Chem. Int. Ed.* **1988**, *27*, 1267–1276; *Angew. Chem.* **1988**, *100*, 1321–1330; g) H.-J. Gabius, C. Schröter, S. Gabius, U. Brinck, L.-F. Tietze, *J. Histochem. Cytochem.* **1990**, *38*, 1625–1631; h) W. Kightlinger, K. F. Warfel, M. P. DeLisa, M. C. Jewett, *ACS Synth. Biol.* **2020**, *9*, 1534–1562; i) K. Vong, T. Yamamoto, K. Tanaka, *Small* **2020**, *16*, e1906890; j) J. M. Groenevelt, D. J. Corey, C. Fehl, *ChemBioChem* **2021**, *22*, 1854–1870.
[30] a) *Neoglycoconjugates. Preparation and Applications* (Eds.: Y. C. Lee, R. T. Lee), Academic Press, San Diego, **1994**; b) N. V. Bovin, H.-J. Gabius, *Chem. Soc. Rev.* **1995**, *24*, 413–421; c) Y. M. Chabre, R. Roy, *Adv. Carbohydr. Chem. Biochem.* **2010**, *63*, 165–393; d) Y. M. Chabre, R. Roy, *Chem. Soc. Rev.* **2013**, *42*, 4657–4708; e) C. S. Kwan, A. R. Cerullo, A. B. Braunschweig, *ChemPlusChem* **2020**, *85*, 2704–2721.
[31] a) V. Percec, P. Leowanawat, H. J. Sun, O. Kulikov, C. D. Nusbaum, T. M. Tran, A. Bertin, D. A. Wilson, M. Peterca, S. Zhang, N. P. Kamat, K. Vargo, D. Moock, E. D. Johnston, D. A. Hammer, D. J. Pochan, Y. Chen, Y. M. Chabre, T. C. Shiao, M. Bergeron-Brlek, S. André, R. Roy, H.-J. Gabius,

P. A. Heiney, *J. Am. Chem. Soc.* **2013**, *135*, 9055–9077; b) S. Zhang, R.-O. Moussodia, C. Murzeau, H. J. Sun, M. L. Klein, S. Vértesy, S. André, R. Roy, H.-J. Gabius, V. Percec, *Angew. Chem. Int. Ed.* **2015**, *54*, 4036–4040; *Angew. Chem.* **2015**, *127*, 4108–4112; c) Q. Xiao, Z. Wang, D. Williams, P. Leowanawat, M. Peterca, S. E. Sherman, S. Zhang, D. A. Hammer, P. A. Heiney, S. R. King, D. M. Markovitz, S. André, H.-J. Gabius, M. L. Klein, V. Percec, *ACS Cent. Sci.* **2016**, *2*, 943–953.

[32] a) J. D. Esko, S. B. Selleck, *Annu. Rev. Biochem.* **2002**, *71*, 435–471; b) E. Buddecke, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 199–216; c) N. K. Karamanos, Z. Piperigkou, A. D. Theocharis, H. Watanabe, M. Franchi, S. Baud, S. Brezillon, M. Gotte, A. Passi, D. Vigetti, S. Ricard-Blum, R. D. Sanderson, T. Neill, R. V. Iozzo, *Chem. Rev.* **2018**, *118*, 9152–9232; d) Y. Song, F. Zhang, R. J. Linhardt, *J. Histochem. Cytochem.* **2021**, *69*, 121–135.

[33] a) J. Kopitz, C. von Reitzenstein, C. Mühl, M. Cantz, *Biochem. Biophys. Res. Commun.* **1994**, *199*, 1188–1193; b) J. Kopitz, C. Mühl, V. Ehemann, C. Lehmann, M. Cantz, *Eur. J. Cell Biol.* **1997**, *73*, 1–9; c) C. Zuber, J. Roth, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 87–110; d) J. J. Caramelo, A. J. Parodi, *FEBS Lett.* **2015**, *589*, 3379–3387; e) A. Tannous, G. B. Pisoni, D. N. Hebert, M. Molinari, *Semin. Cell Dev. Biol.* **2015**, *41*, 79–89; f) J. Roth, C. Zuber, *Histochem. Cell Biol.* **2017**, *147*, 269–284; g) R. W. Ledeen, J. Kopitz, J. Abad-Rodríguez, H.-J. Gabius, *Prog. Mol. Biol. Transl. Sci.* **2018**, *156*, 289–324; h) M. Shenkman, G. Z. Lederkremer, *Trends Biochem. Sci.* **2019**, *44*, 827–836; i) Y. Ito, Y. Kajihara, Y. Takeda, *Chem. Eur. J.* **2020**, *26*, 15461–15470; j) S. Ninagawa, G. George, K. Mori, *Biochim. Biophys. Acta* **2021**, *1865*, 129812.

[34] a) T. Jenuwein, C. D. Allis, *Science* **2001**, *293*, 1074–1080; b) Z. Su, J. M. Denu, *ACS Chem. Biol.* **2016**, *11*, 564–574; c) C. E. Barnes, D. M. English, S. M. Cowley, *Essays Biochem.* **2019**, *63*, 97–107.

[35] a) Y. Narimatsu, H. J. Joshi, R. Nason, J. Van Coillie, R. Karlsson, L. Sun, Z. Ye, Y. H. Chen, K. T. Schjoldager, C. Steentoft, S. Furukawa, B. A. Bensing, P. M. Sullam, A. J. Thompson, J. C. Paulson, C. Bull, G. J. Adema, U. Mandel, L. Hansen, E. P. Bennett, A. Varki, S. Y. Vakhrushev, Z. Yang, H. Clausen, *Mol. Cell* **2019**, *75*, 394–407; b) K. T. Schjoldager, Y. Narimatsu, H. J. Joshi, H. Clausen, *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 729–749; c) Y. F. Huang, K. Aoki, S. Akase, M. Ishihara, Y. S. Liu, G. Yang, Y. Kizuka, S. Mizumoto, M. Tiemeyer, X. D. Gao, K. F. Aoki-Kinoshita, M. Fujita, *Dev. Cell* **2021**, *56*, 1195–1209, e1197; d) Y. Narimatsu, C. Bull, Y. H. Chen, H. H. Wandall, Z. Yang, H. Clausen, *J. Biol. Chem.* **2021**, 100448.

[36] S. Chen, R. Qin, L. K. Mahal, *Crit. Rev. Biochem. Mol. Biol.* **2021**, *56*, 301–320.

[37] a) S. V. Bennun, D. B. Hizal, K. Heffner, O. Can, H. Zhang, M. J. Betenbaugh, *J. Mol. Biol.* **2016**, *428*, 3337–3352; b) B. P. Kellman, N. E. Lewis, *Trends Biochem. Sci.* **2021**, *46*, 284–300; c) J. H. Prestegard, *J. Biol. Chem.* **2021**, *296*, 100556.

[38] J. Schrével, D. Gros, M. Monsigny, *Prog. Histochem. Cytochem.* **1981**, *14*, 1–269.

[39] a) J. Roth, *Biochim. Biophys. Acta* **1987**, *906*, 405–436; b) S. S. Spicer, B. A. Schulte, *J. Histochem. Cytochem.* **1992**, *40*, 1–38; c) F. Akif, H.-J. Gabius, A. Danguy, *Histol. Histopathol.* **1994**, *9*, 705–713; d) J. Roth, *Histochem. Cell Biol.* **2011**, *136*, 117–130; e) J. C. Manning, A. Romero, F. A. Habermann, G. García Caballero, H. Kaltner, H.-J. Gabius, *Histochem. Cell Biol.* **2017**, *147*, 199–222.

[40] a) F. Dall'Olio, *Glycoconjugate J.* **2000**, *17*, 669–676; b) K. A. Dorsett, M. P. Marciel, J. Hwang, K. E. Ankenbauer, N. Bhalerao, S. L. Bellis, *Glycobiology* **2021**, *31*, 530–539.

[41] a) P. Walter, D. Ron, *Science* **2011**, *334*, 1081–1086; b) M. Y. Wong, K. Chen, A. Antonopoulos, B. T. Kasper, M. B. Dewal, R. J. Taylor, C. A. Whittaker, P. P. Hein, A. Dell, J. C. Genereux, S. M. Haslam, L. K. Mahal, M. D. Shoulders, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E10089–E10098.

[42] a) E. Fischer, *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985–2993; b) J. P. Carver, *Pure Appl. Chem.* **1993**, *65*, 763–770; c) C.-W. von der Lieth, H.-C. Siebert, T. Kozár, M. Burchert, M. Frank, M. Gilleron, H. Kaltner, G. Kayser, E. Tajkhorshid, N. V. Bovin, J. F. G. Vliegenthart, H.-J. Gabius, *Acta Anat.* **1998**, *161*, 91–109; d) H.-J. Gabius, *Pharm. Res.* **1998**, *15*, 23–30; e) A. Imberty, S. Pérez, *Chem. Rev.* **2000**, *100*, 4567–4588.

[43] B. J. Hardy, *J. Mol. Struct.* **1997**, *395–396*, 187–200.

[44] P. Weiss, *Yale J. Biol. Med.* **1947**, *19*, 235–278.

[45] W. C. Boyd, E. Shapleigh, *Science* **1954**, *119*, 419.

[46] a) J. Kocourek, in *The Lectins. Properties, Functions and Applications in Biology and Medicine* (Eds.: I. E. Liener, N. Sharon, I. J. Goldstein), Academic Press, Orlando, **1986**, pp. 1–32; b) N. Sharon, H. Lis, *Glycobiology* **2004**, *14*, 53R–62R; c) H. Kaltner, G. García Caballero, A.-K. Ludwig, J. C. Manning, H.-J. Gabius, *Histochem. Cell Biol.* **2018**, *149*, 547–568.

[47] K. Landsteiner, J. van der Scheer, *J. Exp. Med.* **1931**, *54*, 295–305.

[48] W. M. Watkins, W. T. J. Morgan, *Nature* **1952**, *169*, 825–826.

[49] a) J. B. Sumner, S. F. Howell, *J. Immunol.* **1935**, *29*, 133–134; b) J. B. Sumner, S. F. Howell, *J. Bacteriol.* **1936**, *32*, 227–237.

[50] a) D. C. Kilpatrick, C. Green, in *Advances of Lectin Research*, Vol. 5 (Ed.: H. Franz) Ullstein Medical Verlags-, Berlin, **1992**, pp. 51–94; b) W. M. Watkins, *Trends Glycosci. Glycotechnol.* **1999**, *11*, 391–411; c) H. P. Schwarz, F. Dorner, *Br. J. Haematol.* **2003**, *121*, 556–565; d) W. C. Boyd, *Vox Sang.* **1963**, *8*, 1–32; e) W. C. Boyd, *Ann. N. Y. Acad. Sci.* **1970**, *169*, 168–190; f) H. Kaltner, H.-J. Gabius, *Thromb. Haemostasis* **2019**, *119*, 517–533.

[51] a) R. L. Hudgin, W. E. J. Pricer, G. Ashwell, R. J. Stockert, A. G. Morell, *J. Biol. Chem.* **1974**, *249*, 5536–5543; b) V. I. Teichberg, I. Silman, D. D. Beitsch, G. Resheff, *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1383–1387.

[52] a) E. Hohenester, D. Tisi, J. F. Talts, R. Timpl, *Mol. Cell* **1999**, *4*, 783–792; b) D. Tisi, J. F. Talts, R. Timpl, E. Hohenester, *EMBO J.* **2000**, *19*, 1432–1440; c) T. Yoshida-Moriguchi, K. P. Campbell, *Glycobiology* **2015**, *25*, 702–713; d) D. C. Briggs, T. Yoshida-Moriguchi, T. Zheng, D. Venzke, M. E. Anderson, A. Strazzulli, M. Moracci, L. Yu, E. Hohenester, K. P. Campbell, *Nat. Chem. Biol.* **2016**, *12*, 810–814.

[53] a) G. M. Edelman, B. A. Cunningham, G. N. Reeke, Jr., J. W. Becker, M. J. Waxdal, J. L. Wang, *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 2580–2584; b) K. D. Hardman, C. F. Ainsworth, *Biochemistry* **1972**, *11*, 4910–4919; c) Y. D. Lobsanov, M. A. Gitt, H. Leffler, S. H. Barondes, J. M. Rini, *J. Biol. Chem.* **1993**, *268*, 27034–27038; d) R. Loris, *Biochim. Biophys. Acta* **2002**, *1572*, 198–208.

[54] a) R. G. Smock, I. Yadid, O. Dym, J. Clarke, D. S. Tawfik, *Cell* **2016**, *164*, 476–486; b) F. Bonnardel, A. Kumar, M. Wimmerova, M. Lahmann, S. Pérez, A. Varrot, F. Lisacek, A. Imberty, *Structure* **2019**, *27*, 764–775.

[55] a) J. Iwaki, J. Hirabayashi, *Trends Glycosci. Glycotechnol.* **2018**, *30*, SE137–SE153; b) S. Kamitori, *Trends Glycosci. Glycotechnol.* **2018**, *30*, SE41–SE50; c) A. Romero, H.-J. Gabius, *Expert Opin. Ther. Targets* **2019**, *23*, 819–828.

[56] J. R. Marrack, *The Chemistry of Antigens and Antibodies*, H. M. Stationery Office, London, **1934**.

[57] T. K. Dam, H.-J. Gabius, S. André, H. Kaltner, M. Lensch, C. F. Brewer, *Biochemistry* **2005**, *44*, 12564–12571.

[58] a) Z. Tu, H.-W. Hsieh, C.-M. Tsai, C. W. Hsu, S. G. Wang, K. J. Wu, K. I. Lin, C. H. Lin, *Chem. Asian J.* **2013**, *8*, 1536–1550; b) R. Roy, Y. Cao, H. Kaltner, N. Kottari, T. C. Shiao, K. Belkhadem, S. André, J. C. Manning, P. V. Murphy, H.-J. Gabius, *Histochem. Cell Biol.* **2017**, *147*, 285–301; c) Q. Xiao, A.-K. Ludwig, C. Romanò, I. Buzzacchera, S. E. Sherman, M. Vetro, S. Vértesy, H. Kaltner, E. H. Reed, M. Möller, C. J. Wilson, D. A. Hammer, S. Oscarson, M. L. Klein, H.-J. Gabius, V. Percec, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E2509–E2518.

[59] a) H. Ideo, A. Seko, I. Ishizuka, K. Yamashita, *Glycobiology* **2003**, *13*, 713–723; b) H. Ideo, A. Seko, K. Yamashita, *J. Biol. Chem.* **2005**, *280*, 4730–4737; c) Y. Zheng, J. Su, M.C. Miller, T. Zhang, M. Mayzel, G. Tai, K.H. Mayo, Y. Zhou, *Glycobiology* **2021**, *31*, 341–350.

[60] P. V. Murphy, A. Romero, Q. Xiao, A.-K. Ludwig, S. Jogula, N. V. Shilova, T. Singh, A. Gabba, B. Javed, D. Zhang, F. J. Medrano, H. Kaltner, J. Kopitz, N. V. Bovin, A. M. Wu, M. L. Klein, V. Percec, H.-J. Gabius, *iScience* **2021**, *24*, 101919.

[61] a) E. Ishikawa, T. Ishikawa, Y. S. Morita, K. Toyonaga, H. Yamada, O. Takeuchi, T. Kinoshita, S. Akira, Y. Yoshikai, S. Yamasaki, *J. Exp. Med.* **2009**, *206*, 2879–2888; b) H. Feinberg, N. D. Rambaruth, S. A. Jegouzo, K. M. Jacobsen, R. Djurhuus, T. B. Poulsen, W. I. Weis, M. E. Taylor, K. Drickamer, *J. Biol. Chem.* **2016**, *291*, 21222–21233; c) K. Toyonaga, S. Torigoe, Y. Motomura, T. Kamichi, J. M. Hayashi, Y. S. Morita, N. Noguchi, Y. Chuma, H. Kiyohara, K. Matsuo, H. Tanaka, Y. Nakagawa, T. Sakuma, M. Ohmuraya, T. Yamamoto, M. Umemura, G. Matsuzaki, Y. Yoshikai, I. Yano, T. Miyamoto, S. Yamasaki, *Immunity* **2016**, *45*, 1245–1257; d) Z. Omahdi, Y. Horikawa, M. Nagae, K. Toyonaga, A. Imamura, K. Takato, T. Teramoto, H. Ishida, Y. Kakuta, S. Yamasaki, *J. Biol. Chem.* **2020**, *295*, 5807–5817.

[62] D. Solís, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Smetana Jr., H.-J. Gabius, *Biochim. Biophys. Acta* **2015**, *1850*, 186–235.

[63] a) M. C. Chervenak, E. J. Toone, *J. Am. Chem. Soc.* **1994**, *116*, 10533–10539; b) T. K. Dam, S. Oscarson, J. C. Sacchettini, C. F. Brewer, *J. Biol. Chem.* **1998**, *273*, 32826–32832; c) N. Ahmad, H.-J. Gabius, H. Kaltner, S. André, I. Kuwabara, F.-T. Liu, S. Oscarson, T. Norberg, C. F. Brewer, *Can. J. Chem.* **2002**, *80*, 1096–1104; d) D. M. Beckwith, F. G. Fitz Gerald, M. C.

Rodriguez Benavente, E. R. Mercer, A. K. Ludwig, M. Michalak, H. Kaltner, J. Kopitz, H. J. Gabius, M. Cudic, *Biochemistry* **2021**, *60*, 547–558; e) T. Diercks, F. J. Medrano, F. G. FitzGerald, D. Beckwith, M. J. Pedersen, M. Reihill, A.-K. Ludwig, A. Romero, S. Oscarson, M. Cudic, H.-J. Gabius, *Chem. Eur. J.* **2021**, *27*, 316–325.

[64] a) L. He, S. André, H.-C. Siebert, H. Helmholz, B. Niemeyer, H.-J. Gabius, *Biophys. J.* **2003**, *85*, 511–524; b) C. Diehl, S. Genheden, K. Modig, U. Ryde, M. Akke, *J. Biomol. NMR* **2009**, *45*, 157–169; c) C. Diehl, O. Engstrom, T. Delaine, M. Hakansson, S. Genheden, K. Modig, H. Leffler, U. Ryde, U. J. Nilsson, M. Akke, *J. Am. Chem. Soc.* **2010**, *132*, 14577–14589; d) A. Göhler, S. André, H. Kaltner, M. Sauer, H.-J. Gabius, S. Doose, *Biophys. J.* **2010**, *98*, 3044–3053; e) F. Shang, M. J. Rynkiewicz, F. X. McCormack, H. Wu, T. M. Cafarella, J. F. Head, B. A. Seaton, *J. Biol. Chem.* **2011**, *286*, 757–765; f) E. Ermakova, M. C. Miller, I. V. Nesmelova, L. Lopez-Merino, M. A. Berbís, Y. Nesmelov, Y. V. Tkachev, L. Lagartera, V. A. Daragan, S. André, F. J. Cañada, J. Jiménez-Barbero, D. Solís, H.-J. Gabius, K. H. Mayo, *Glycobiology* **2013**, *23*, 508–523; g) A. Schedlbauer, U. Gilles, A.-K. Ludwig, A. Adler, H. Kaltner, I. Lindner, K. H. Mayo, T. Diercks, D. Reusch, H.-J. Gabius, *Biochimie* **2021**, *187*, 48–56.

[65] B. G. Keller, C. Rademacher, *Curr. Opin. Struct. Biol.* **2020**, *62*, 31–38.

[66] a) I. J. Goldstein, R. D. Poretz, in *The Lectins: Properties, Functions, and Applications in Biology and Medicine* (Eds.: I.E. Liener, N. Sharon, I. J. Goldstein), Academic Press, Orlando, **1986**, pp. 33–247; b) D. Gupta, S. Oscarson, T. S. Raju, P. Stanley, E. J. Toone, C. F. Brewer, *Eur. J. Biochem.* **1996**, *242*, 320–326; c) R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373–380; d) D. Solís, A. Romero, H. Kaltner, H.-J. Gabius, T. Díaz-Mauriño, *J. Biol. Chem.* **1996**, *271*, 12744–12748; e) T. K. Dam, S. Oscarson, C. F. Brewer, *J. Biol. Chem.* **1998**, *273*, 32812–32817.

[67] a) C. F. Brewer, H. Sternlicht, D. M. Marcus, A. P. Grollman, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 1007–1011; b) G. M. Alter, J. A. Magnuson, *Biochemistry* **1974**, *13*, 4038–4045; c) E. Matei, S. André, A. Glinschert, A. S. Infantino, S. Oscarson, H.-J. Gabius, A. M. Gronenborn, *Chem. Eur. J.* **2013**, *19*, 5364–5374; d) T. Diercks, A. S. Infantino, L. Unione, J. Jiménez-Barbero, S. Oscarson, H.-J. Gabius, *Chem. Eur. J.* **2018**, *24*, 15761–15765; e) G. Nestor, T. Anderson, S. Oscarson, A. M. Gronenborn, *J. Am. Chem. Soc.* **2017**, *139*, 6210–6216; f) M. Raics, I. Timári, T. Diercks, L. Szilágyi, H.-J. Gabius, K. E. Kövér, *ChemBioChem* **2019**, *20*, 1688–1692; g) G. Fittolani, E. Shanina, M. Guberman, P. H. Seeberger, C. Rademacher, M. Delbianco, *Angew. Chem. Int. Ed.* **2021**, *60*, 13302–13309.

[68] a) A. S. Powlesland, T. Fisch, M. E. Taylor, D. F. Smith, B. Tissot, A. Dell, S. Pohlmann, K. Drickamer, *J. Biol. Chem.* **2008**, *283*, 593–602; b) R. Yabe, H. Tateno, J. Hirabayashi, *FEBS J.* **2010**, *277*, 4010–4026; c) H.-J. Gabius, J. C. Manning, J. Kopitz, S. André, H. Kaltner, *Cell. Mol. Life Sci.* **2016**, *73*, 1989–2016; d) Y. Manabe, R. Marchetti, Y. Takakura, M. Nagasaki, W. Nihei, T. Takebe, K. Tanaka, K. Kabayama, F. Chiodo, S. Hanashima, Y. Kamada, E. Miyoshi, H. P. Dulal, Y. Yamaguchi, Y. Adachi, N. Ohno, H. Tanaka, A. Silipo, K. Fukase, A. Molinaro, *Angew. Chem. Int. Ed.* **2019**, *58*, 18697–18702; *Angew. Chem.* **2019**, *131*, 18870–18875; e) K. Stavenhagen, L. C. Laan, C. Gao, A. Y. Mehta, J. Heimburg-Molinaro, J. N. Glickman, I. van Die, R. D. Cummings, *Cell. Mol. Life Sci.* **2021**, *78*, 5569–5585.

[69] a) *Proteoglycans. Structure, Biology and Molecular Interactions* (Ed.: R. V. Iozzo), Marcel Dekker, New York, **2000**; b) I. Capila, R. J. Linhardt, *Angew. Chem. Int. Ed.* **2002**, *41*, 390–412; *Angew. Chem.* **2002**, *114*, 426–450; c) C. I. Gama, L. C. Hsieh-Wilson, *Curr. Opin. Chem. Biol.* **2005**, *9*, 609–619; d) B. Caterson, J. Melrose, *Glycobiology* **2018**, *28*, 182–206; e) M. C. Miller, C. Cai, K. Wichapong, S. Bhaduri, N. L. B. Pohl, R. J. Linhardt, H.-J. Gabius, K. H. Mayo, *Sci. Rep.* **2020**, *10*, 15708; f) A. Gómez Toledo, J. T. Sorrentino, D. R. Sandoval, J. Malmstrom, N. E. Lewis, J. D. Esko, *J. Histochem. Cytochem.* **2021**, *69*, 105–119; g) S. D. Vallet, O. Clerc, S. Ricard-Blum, *J. Histochem. Cytochem.* **2021**, *69*, 93–104.

[70] a) D. A. Wesener, K. Wangkanont, R. McBride, X. Song, M. B. Kraft, H. L. Hodges, L. C. Zarling, R. A. Splain, D. F. Smith, R. D. Cummings, J. C. Paulson, K. T. Forest, L. L. Kiessling, *Nat. Struct. Mol. Biol.* **2015**, *22*, 603–610; b) D. A. Wesener, A. Dugan, L. L. Kiessling, *Curr. Opin. Struct. Biol.* **2017**, *44*, 168–178; c) C. M. McMahon, C. R. Isabella, I. W. Windsor, P. Kosma, R. T. Raines, L. L. Kiessling, *J. Am. Chem. Soc.* **2020**, *142*, 2386–2395.

[71] a) T. Saito, S. Kawabata, M. Hirata, S. Iwanaga, *J. Biol. Chem.* **1995**, *270*, 14493–14499; b) C.-G. Huh, J. Aldrich, J. Mottahedeh, H. Kwon, C. Johnson, R. Marsh, *J. Biol. Chem.* **1998**, *273*, 6565–6574; c) D. H. Low, Z. Ang, Q. Yuan, V. Frecer, B. Ho, J. Chen, J. L. Ding, *PLoS One* **2009**, *4*, e6260; d) T. Wohlschlager, A. Butschi, P. Grassi, G. Sutov, R. Gauss, D. Hauck, S. S. Schmieder, M. Knobel, A. Titz, A. Dell, S. M. Haslam, M. O. Hengartner, M. Aebi, M. Künzler, *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E2787–2796; e) R. Sommer, O. N. Makshakova, T. Wohlschlager, S. Hutin, M. Marsh, A. Titz, M. Künzler, A. Varrot, *Structure* **2018**, *26*, 391–402, e394.

[72] B. M. Gesner, V. Ginsburg, *Proc. Natl. Acad. Sci. USA* **1964**, *52*, 750–755.

[73] a) W. R. Springer, D. N. W. Cooper, S. H. Barondes, *Cell* **1984**, *39*, 557–564; b) S. H. Barondes, P. L. Haywood-Reid, D. N. W. Cooper, *J. Cell Biol.* **1985**, *100*, 1825–1833; c) H.-J. Gabius, W. R. Springer, S. H. Barondes, *Cell* **1985**, *42*, 449–456; d) S. H. Barondes, *Trends Biochem. Sci.* **1988**, *13*, 480–482; e) H.-J. Gabius, *Int. J. Biochem.* **1994**, *26*, 469–477.

[74] a) B.-W. Kim, S. B. Hong, J. H. Kim, H. Kwon do, H. K. Song, *Nat. Commun.* **2013**, *4*, 1613; b) S. Li, M. P. Wandel, F. D. Li, Z. H. Liu, C. He, J. H. Wu, Y. Y. Shi, F. Randow, *Sci. Signaling* **2013**, *6*, ra9; c) M.-H. Hong, I.-C. Weng, F.-T. Liu, *Trends Glycosci. Glycotechnol.* **2018**, *30*, SE179–SE184; d) V. Eckardt, M. C. Miller, X. Blanchet, R. Duan, J. Leberzammer, J. Duchene, O. Soehnlein, R. T. Megens, A.-K. Ludwig, A. Dregni, A. Faussner, K. Wichapong, H. Ippel, I. Dijkgraaf, H. Kaltner, Y. Doring, K. Bidzhekov, T. M. Hackeng, C. Weber, H.-J. Gabius, P. von Hundelshausen, K. H. Mayo, *EMBO Rep.* **2020**, *21*, e47852; e) M. H. Hong, I. C. Weng, F. Y. Li, W. H. Lin, F.-T. Liu, *J. Biomed. Sci.* **2021**, *28*, 16; f) S. L. Bell, K. L. Lopez, J. S. Cox, K. L. Patrick, R. O. Watson, *mBio* **2021**, *12*, e01871–e01820; g) F. Reggiori, H.-J. Gabius, M. Aureli, W. Römer, S. Sonnino, E.-L. Eskelinen, *Glycoconjugate J.* **2021**, DOI https://doi.org/10.1007/s10719-021-10007-x.

[75] a) S. H. Koenig, C. F. Brewer, R. D. Brown, 3rd, *Biochemistry* **1978**, *17*, 4251–4260; b) C. F. Brewer, R. D. Brown, 3rd, S. H. Koenig, *J. Biomol. Struct. Dyn.* **1983**, *1*, 961–997; c) P. L. Whitney, J. T. Powell, G. L. Sanford, *Biochem. J.* **1986**, *238*, 683–689; d) V. Wells, L. Mallucci, *Cell* **1991**, *64*, 91–97; e) J. Bouckaert, R. Loris, F. Poortmans, L. Wyns, *Proteins* **1995**, *23*, 510–524; f) K. K.-S. Ng, S. Park-Snyder, W. I. Weis, *Biochemistry* **1998**, *37*, 17965–17976; g) K. K.-S. Ng, W. I. Weis, *Biochemistry* **1998**, *37*, 17977–17989; h) T. Kadoya, H. Horie, *Curr. Drug Targets* **2005**, *6*, 375–383; i) M. F. López-Lucendo, D. Solís, J. L. Sáiz, H. Kaltner, R. Russwurm, S. André, H.-J. Gabius, A. Romero, *J. Mol. Biol.* **2009**, *386*, 366–378; j) M. C. Miller, A.-K. Ludwig, K. Wichapong, H. Kaltner, J. Kopitz, H.-J. Gabius, K. H. Mayo, *Biochem. J.* **2018**, *475*, 1003–1018; k) M. Sakakura, M. Tamura, N. Fujii, T. Takeuchi, T. Hatanaka, S. Kishimoto, Y. Arata, H. Takahashi, *FEBS J.* **2018**, *285*, 1129–1145; l) M. C. Miller, I. V. Nesmelova, V. A. Daragan, H. Ippel, M. Michalak, A. Dregni, H. Kaltner, J. Kopitz, H.-J. Gabius, K. H. Mayo, *Biochem. J.* **2020**, *477*, 3147–3165; m) D. Sakthivel, S. Preston, R. B. Gasser, T. Costa, J. N. Hernandez, A. Shahine, M. D. Shakif-Azam, P. Lock, J. Rossjohn, M. A. Perugini, J. F. Gonzalez, E. Meeusen, D. Piedrafita, T. Beddoe, *Commun. Biol.* **2020**, *3*, 464; n) H.-J. Gabius, *Glycoforum* **2021**, *24*, A1.

[76] a) T. J. Wilson, M. N. Firth, J. T. Powell, F. L. Harrison, *Biochem. J.* **1989**, *261*, 847–852; b) R. C. Hughes, *Biochim. Biophys. Acta* **1999**, *1473*, 172–185; c) D. N. W. Cooper, *Biochim. Biophys. Acta* **2002**, *1572*, 209–231; d) S. Sato, *Trends Glycosci. Glycotechnol.* **2018**, *30*, SE199–SE209; e) G. García Caballero, H. Kaltner, T. J. Kutzner, A.-K. Ludwig, J. C. Manning, S. Schmidt, F. Sinowatz, H.-J. Gabius, *Histol. Histopathol.* **2020**, *35*, 509–539; f) T. J. Kutzner, A. M. Higuero, M. Süßmair, J. Kopitz, M. Hingar, N. Díez-Revuelta, G. García Caballero, H. Kaltner, I. Lindner, J. Abad-Rodríguez, D. Reusch, H.-J. Gabius, *Biochim. Biophys. Acta* **2020**, *1864*, 129449.

[77] a) S. H. Barondes, *Annu. Rev. Biochem.* **1981**, *50*, 207–231; b) H.-J. Gabius, *Eur. J. Biochem.* **1997**, *243*, 543–576; c) H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674; d) W. I. Weis, M. E. Taylor, K. Drickamer, *Immunol. Rev.* **1998**, *163*, 19–34; e) J. N. Gready, A. N. Zelensky, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 329–346; f) D. C. Kilpatrick, *Biochim. Biophys. Acta* **2002**, *1572*, 187–197; g) G. P. Bhide, K. J. Colley, *Histochem. Cell Biol.* **2017**, *147*, 149–174; h) S. Mayer, M. K. Raulf, B. Lepenies, *Histochem. Cell Biol.* **2017**, *147*, 223–237; i) M. E. Taylor, K. Drickamer, *FEBS J.* **2019**, *286*, 1800–1814; j) K. Li, D. M. Underhill, *Curr. Top. Microbiol. Immunol.* **2020**, *429*, 1–18; k) M. Nagae, Y. Yamaguchi, *Curr. Top. Microbiol. Immunol.* **2020**, *429*, 147–176.

[78] a) H. L. Cash, C. V. Whitham, C. L. Behrendt, L. V. Hooper, *Science* **2006**, *313*, 1126–1130; b) S. Mukherjee, C. L. Partch, R. E. Lehotzky, C. V. Whitham, H. Chu, C. L. Bevins, K. H. Gardner, L. V. Hooper, *J. Biol. Chem.* **2009**, *284*, 4881–4888; c) R. E. Lehotzky, C. L. Partch, S. Mukherjee, H. L. Cash, W. E. Goldman, K. H. Gardner, L. V. Hooper, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 7722–7727; d) T. Miki, O. Holst, W. D. Hardt, *J. Biol. Chem.* **2012**, *287*, 34844–34855; e) S. Mukherjee, H. Zheng, M. G.

Derebe, K. M. Callenberg, C. L. Partch, D. Rollins, D. C. Propheter, J. Rizo, M. Grabe, Q. X. Jiang, L. V. Hooper, *Nature* **2014**, *505*, 103–107.

[79] a) Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.* **1983**, *258*, 199–202; b) R. T. Lee, P. Lin, Y. C. Lee, *Biochemistry* **1984**, *23*, 4255–4261; c) Y. C. Lee, *FASEB J.* **1992**, *6*, 3193–3200; d) R. T. Lee, Y. C. Lee, in *Neoglycoconjugates. Preparation and Applications* (Eds.: Y. C. Lee, R. T. Lee), Academic Press, San Diego, **1994**, pp. 23–50; e) R. T. Lee, Y. C. Lee, in *Glycosciences: Status and Perspectives* (Eds.: H.-J. Gabius, S. Gabius), Chapman & Hall, London, **1997**, pp. 55–77.

[80] a) E. Ruoslahti, *Glycobiology* **1996**, *6*, 489–492; b) Y. Yamaguchi, *Cell. Mol. Life Sci.* **2000**, *57*, 276–289; c) A. Aspberg, *J. Histochem. Cytochem.* **2012**, *60*, 987–996; d) R. V. Iozzo, L. Schaefer, *Matrix Biol.* **2015**, *42*, 11–55.

[81] G. R. Gunther, J. L. Wang, I. Yahara, B. A. Cunningham, G. M. Edelman, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 1012–1016.

[82] a) G. Ashwell, J. Harford, *Annu. Rev. Biochem.* **1982**, *51*, 531–554; b) H.-J. Gabius, *Biochim. Biophys. Acta* **1991**, *1071*, 1–18; c) D. Fiete, M. C. Beranek, J. U. Baenziger, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 11256–11261; d) E. Y. Adashi, M. Fukuda, *J. Clin. Invest.* **2002**, *109*, 169–170; e) N. M. Dahms, M. K. Hancock, *Biochim. Biophys. Acta* **2002**, *1572*, 317–340; f) J. Lu, C. Teh, U. Kishore, K. B. M. Reid, *Biochim. Biophys. Acta* **2002**, *1572*, 387–400; g) A. L. Sorensen, H. Clausen, H. H. Wandall, *Biochim. Biophys. Acta* **2012**, *1820*, 1797–1808; h) S. Velasco, N. Díez-Revuelta, T. Hernández-Iglesias, H. Kaltner, S. André, H.-J. Gabius, J. Abad-Rodríguez, *J. Neurochem.* **2013**, *125*, 49–62; i) S. Kornfeld, *Annu. Rev. Biochem.* **2018**, *87*, 1–21.

[83] a) R. Takamiya, K. Ohtsubo, S. Takamatsu, N. Taniguchi, T. Angata, *Glycobiology* **2013**, *23*, 178–187; b) J. M. Pröpster, F. Yang, S. Rabbani, B. Ernst, F. H.-T. Allain, M. Schubert, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, e4170–e4179; c) T. Angata, *Adv. Exp. Med. Biol.* **2020**, *1204*, 215–230; d) R. Beatson, R. Graham, F. Grundland Freile, D. Cozzetto, S. Kannambath, E. Pfeifer, N. Woodman, J. Owen, R. Nuamah, U. Mandel, S. Pinder, C. Gillett, T. Noll, I. Bouybayoune, J. Taylor-Papadimitriou, J. M. Burchell, *Commun. Biol.* **2020**, *3*, 644; e) H. Läubli, A. Varki, *Cell. Mol. Life Sci.* **2020**, *77*, 593–605; f) S. Wisnovsky, L. Mockl, S. A. Malaker, K. Pedram, G. T. Hess, N. M. Riley, M. A. Gray, B. A. H. Smith, M. C. Bassik, W. E. Moerner, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2015024118.

[84] a) S. Yamasaki, E. Ishikawa, M. Sakuma, H. Hara, K. Ogata, T. Saito, *Nat. Immunol.* **2008**, *9*, 1179–1188; b) I. M. Dambuza, G. D. Brown, *Curr. Opin. Immunol.* **2015**, *32*, 21–27; c) D. Zhao, X. Han, X. Zheng, H. Wang, Z. Yang, D. Liu, K. Han, J. Liu, X. Wang, W. Yang, Q. Dong, S. Yang, X. Xia, L. Tang, F. He, *PLoS Pathog.* **2016**, *12*, e1005487; d) J. Ostrop, R. Lang, *J. Immunol.* **2017**, *198*, 1403–1414; e) P. S. Sung, W. C. Chang, S. L. Hsieh, *Adv. Exp. Med. Biol.* **2020**, *1204*, 57–73.

[85] a) L. J. S. Yang, C. B. Zeller, N. L. Shaper, M. Kiso, A. Hasegawa, R. E. Shapiro, R. L. Schnaar, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 814–818; b) R. L. Schnaar, B. E. Collins, L. P. Wright, M. Kiso, M. B. Tropak, J. C. Roder, P. R. Crocker, *Ann. N. Y. Acad. Sci.* **1998**, *845*, 92–105; c) P. H. H. Lopez, *Adv. Neurobiol.* **2014**, *9*, 245–262; d) M. F. Pronker, S. Lemstra, J. Snijder, A. J. R. Heck, D. M. Thies-Weesie, R. J. Pasterkamp, B. J. Janssen, *Nat. Commun.* **2016**, *7*, 13584; e) M. Myllykoski, M. A. Eichel, R. B. Jung, S. Kelm, H. B. Werner, P. Kursula, *J. Neurochem.* **2018**, *147*, 764–783.

[86] a) L. Warren, C. A. Buck, G. P. Tuszynski, *Biochim. Biophys. Acta* **1978**, *516*, 97–127; b) M. Pierce, J. Arango, *J. Biol. Chem.* **1986**, *261*, 10772–10777; c) K. Polyak, Y. Xia, J. L. Zweier, K. M. Kinzler, B. Vogelstein, *Nature* **1997**, *389*, 300–305; d) M. Sperandio, *FEBS J.* **2006**, *273*, 4377–4389; e) S. Bi, L. G. Baum, *Biochim. Biophys. Acta* **2009**, *1790*, 1599–1610; f) M. Amano, H. Eriksson, J. C. Manning, K. M. Detjen, S. André, S.-I. Nishimura, J. Lehtiö, H.-J. Gabius, *FEBS J.* **2012**, *279*, 4062–4080; g) S. R. Stowell, T. Ju, R. D. Cummings, *Annu. Rev. Pathol. Mech. Dis.* **2015**, *10*, 473–510; h) S. Thiemann, L. G. Baum, *Annu. Rev. Immunol.* **2016**, *34*, 243–264; i) H. Kaltner, S. Toegel, G. García Caballero, J. C. Manning, R. W. Ledeen, H.-J. Gabius, *Histochem. Cell Biol.* **2017**, *147*, 239–256; j) C. G. H. M. de Jong, H.-J. Gabius, W. Baron, *Cell. Mol. Life Sci.* **2020**, *77*, 1289–1317; k) S. Duan, J. C. Paulson, *Annu. Rev. Immunol.* **2020**, *38*, 365–395.

[87] a) A. Varki, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 7390–7397; b) A. Varki, *J. Clin. Invest.* **1997**, *99*, 158–162; c) T. K. Dam, C. F. Brewer, *Glycobiology* **2010**, *20*, 270–279; d) H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero, D. Solís, *Trends Biochem. Sci.* **2011**, *36*, 298–313; e) J. W. Dennis, C. F. Brewer, *Mol. Cell. Proteomics* **2013**, *12*, 913–920.

[88] C. M. Blouin, Y. Hamon, P. Gonnord, C. Boularan, J. Kagan, C. Viaris de Lesegno, R. Ruez, S. Mailfert, N. Bertaux, D. Loew, C. Wunder, L. Johannes, G. Vogt, F. X. Contreras, D. Marguet, J. L. Casanova, C. Gales, H. T. He, C. Lamaze, *Cell* **2016**, *166*, 920–934.

[89] a) C. Sparrow, H. Leffler, S. H. Barondes, *J. Biol. Chem.* **1987**, *262*, 7383–7390; b) L. G. Yu, N. Andrews, Q. Zhao, D. McKean, J. F. Williams, L. J. Connor, O. V. Gerasimenko, J. Hilkens, J. Hirabayashi, K. Kasai, J. M. Rhodes, *J. Biol. Chem.* **2007**, *282*, 773–781; c) M. Krzeminski, T. Singh, S. André, M. Lensch, A. M. Wu, A. M. J. J. Bonvin, H.-J. Gabius, *Biochim. Biophys. Acta* **2011**, *1810*, 150–161; d) P. Guha, E. Kaptan, G. Bandyopadhyaya, S. Kaczanowska, E. Davila, K. Thompson, S. S. Martin, D. V. Kalvakolanu, G. R. Vasta, H. Ahmed, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5052–5057.

[90] a) G. Hübner, *Z. Immunitaetsforsch.* **1926**, *45*, 223–248; b) O. Thomsen, *Z. Immunitaetsforsch.* **1927**, *52*, 85–107; c) V. Friedenreich, *The Thomsen Hemagglutination Phenomenon, Vol. 1*, Levin and Munksgaard, Copenhagen, **1930**; d) R. Moreau, J. Dausset, J. Bernard, J. Moullec, *Bull. Mem. Soc. Med. Hop. Paris* **1957**, *73*, 569–587; e) G. Uhlenbruck, *Immunol. Invest.* **1981**, *10*, 251–264; f) I. Brockhausen, J. Schutzbach, W. Kuhns, *Acta Anat.* **1998**, *161*, 36–78; g) E. G. Berger, *Biochim. Biophys. Acta* **1999**, *1455*, 255–268; h) I. Brockhausen, *EMBO Rep.* **2006**, *7*, 599–604; i) T. Ju, V. I. Otto, R. D. Cummings, *Angew. Chem. Int. Ed.* **2011**, *50*, 1770–1791; *Angew. Chem.* **2011**, *123*, 1808–1830; j) N. Murugaesu, M. Iravani, A. van Weverwijk, A. Ivetic, D. A. Johnson, A. Antonopoulos, A. Fearns, M. Jamal-Hanjani, D. Sims, K. Fenwick, C. Mitsopoulos, Q. Gao, N. Orr, M. Zvelebil, S. M. Haslam, A. Dell, H. Yarwood, C. J. Lord, A. Ashworth, C. M. Isacke, *Cancer Dis.* **2014**, *4*, 304–317; k) A. P. Corfield, *Biochim. Biophys. Acta* **2015**, *1850*, 236–252; l) C. J. Dimitroff, *Cancer Res.* **2015**, *75*, 3195–3202; m) N. E. Reticker-Flynn, S. N. Bhatia, *Cancer Dis.* **2015**, *5*, 168–181; n) J. M. Burchell, R. Beatson, R. Graham, J. Taylor-Papadimitriou, V. Tajadura-Ortega, *Biochem. Soc. Trans.* **2018**, *46*, 779–788.

[91] a) P. L. Smith, J. U. Baenziger, *Science* **1988**, *242*, 930–933; b) P. L. Smith, J. U. Baenziger, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 329–333; c) L. V. Hooper, S. M. Manzella, J. U. Baenziger, in *Glycosciences: Status and Perspectives* (Eds.: H.-J. Gabius, S. Gabius), Chapman & Hall, London, **1997**, pp. 261–276.

[92] a) W. Thomas, *Annu. Rev. Biomed. Eng.* **2008**, *10*, 39–57; b) E. V. Sokurenko, V. Vogel, W. E. Thomas, *Cell Host Microbe* **2008**, *4*, 314–323; c) P. Sundd, M. K. Pospieszalska, K. Ley, *Mol. Immunol.* **2013**, *55*, 59–69; d) R. C. Preston, R. P. Jakob, F. P. Binder, C. P. Sager, B. Ernst, T. Maier, *J. Mol. Cell. Biol.* **2016**, *8*, 62–72; e) P. Mehta-D'souza, A. G. Klopocki, V. Oganesyan, S. Terzyan, T. Mather, Z. Li, S. R. Panicker, C. Zhu, R. P. McEver, *J. Biol. Chem.* **2017**, *292*, 2510–2518.

[93] a) I. Ofek, N. Sharon, *Curr. Top. Microbiol. Immunol.* **1990**, *151*, 91–114; b) S. Holgersson, A. Gustafsson, S. Gaunitz, in *The Sugar Code: Fundamentals of Glycosciences* (Ed.: H.-J. Gabius), Wiley-VCH, Weinheim, **2009**, pp. 279–300; c) P. W. J. de Groot, O. Bader, A. D. de Boer, M. Weig, N. Chauhan, *Eukaryotic Cell* **2013**, *12*, 470–481; d) W. van Breedam, S. Pohlmann, H. W. Favoreel, R. J. de Groot, H. J. Nauwynck, *FEMS Microbiol. Rev.* **2014**, *38*, 598–632; e) K. Moonens, H. Remaut, *Curr. Opin. Struct. Biol.* **2017**, *44*, 48–58; f) J. Poole, C. J. Day, M. von Itzstein, J. C. Paton, M. P. Jennings, *Nat. Rev. Microbiol.* **2018**, *16*, 440–452; g) S. L. Taylor, M. A. McGuckin, S. Wesselingh, G. B. Rogers, *Trends Microbiol.* **2018**, *26*, 92–101; h) A. J. Thompson, R. P. de Vries, J. C. Paulson, *Curr. Opin. Virol.* **2019**, *34*, 117–129; i) M. L. Klein, A. Romero, H. Kaltner, V. Percec, H.-J. Gabius, *Biophys. J.* **2021**, *120*, 1031–1039.

[94] a) T. Borén, S. Normark, P. Falk, *Trends Microbiol.* **1994**, *2*, 221–228; b) K. A. Karlsson, *Glycobiology* **2000**, *10*, 761–771; c) M. Kobayashi, H. Lee, J. Nakayama, M. Fukuda, *Glycobiology* **2009**, *19*, 453–461; d) J. Abed, J. E. Emgard, G. Zamir, M. Faroja, G. Almogy, A. Grenov, A. Sol, R. Naor, E. Pikarsky, K. A. Atlan, A. Mellul, S. Chaushu, A. L. Manson, A. M. Earl, N. Ou, C. A. Brennan, W. S. Garrett, G. Bachrach, *Cell Host Microbe* **2016**, *20*, 215–225; e) M. S. Conover, S. Ruer, J. Taganna, V. Kalas, H. De Greve, J. S. Pinkner, K. W. Dodson, H. Remaut, S. J. Hultgren, *Cell Host Microbe* **2016**, *20*, 482–492; f) D. Keilberg, K. M. Ottemann, *Environ. Microbiol.* **2016**, *18*, 791–806.

[95] a) J. Hirabayashi, R. Arai, *Interface Focus* **2019**, *9*, 20180068; b) A.-K. Ludwig, H. Kaltner, J. Kopitz, H.-J. Gabius, *Biochim. Biophys. Acta* **2019**, *1863*, 935–940; c) A.-K. Ludwig, M. Michalak, Q. Xiao, U. Gilles, F. J. Medrano, H. Ma, F. G. FitzGerald, W. D. Hasley, A. Melendez-Davila, M. Liu, K. Rahimi, N. Y. Kostina, C. Rodriguez-Emmenegger, M. Möller, I. Lindner, H. Kaltner, M. Cudic, D. Reusch, J. Kopitz, A. Romero, S. Oscarson, M. L. Klein, H.-J. Gabius, V. Percec, *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 2837–2842; d) G. García Caballero, D. Beckwith, N. V. Shilova, A. Gabba, T. J. Kutzner, A.-K. Ludwig, J. C. Manning, H. Kaltner,

F. Sinowatz, M. Cudic, N. V. Bovin, P. V. Murphy, H.-J. Gabius, *Histochem. Cell Biol.* **2020**, *154*, 135–153.
[96] a) S. André, B. Liu, H.-J. Gabius, R. Roy, *Org. Biomol. Chem.* **2003**, *1*, 3909–3916; b) S. André, P. J. Cejas Ortega, M. Alamino Perez, R. Roy, H.-J. Gabius, *Glycobiology* **1999**, *9*, 1253–1261.
[97] a) H.-J. Gabius, B. Wosgien, M. Hendrys, A. Bardosi, *Histochemistry* **1991**, *95*, 269–277; b) J. C. Manning, G. García Caballero, C. Knospe, H. Kaltner, H.-J. Gabius, *J. Anat.* **2017**, *231*, 23–37; c) J. C. Manning, G. García Caballero, C. Knospe, H. Kaltner, H.-J. Gabius, *Ann. Anat.* **2018**, *217*, 66–81; d) J. C. Manning, G. Garcia Caballero, A.-K. Ludwig, H. Kaltner, F. Sinowatz, H.-J. Gabius, *Anat. Rec.* **2021**, *304*, 1597–1628.
[98] a) T. Angata, E. C. M. Brinkman van der Linden, *Biochim. Biophys. Acta* **2002**, *1572*, 294–316; b) L. East, C. M. Isacke, *Biochim. Biophys. Acta* **2002**, *1572*, 364–386; c) H.-J. Gabius, H. Kaltner, J. Kopitz, S. André, *Trends Biochem. Sci.* **2015**, *40*, 360–376.
[99] a) S. Toegel, D. Bieder, S. André, K. Kayser, S. M. Walzer, G. Hobusch, R. Windhager, H.-J. Gabius, *Histochem. Cell Biol.* **2014**, *142*, 373–388; b) S. Toegel, D. Weinmann, S. André, S. M. Walzer, M. Bilban, S. Schmidt, C. Chiari, R. Windhager, C. Krall, I. M. Bennani-Baiti, H.-J. Gabius, *J. Immunol.* **2016**, *196*, 1910–1921; c) D. Weinmann, K. Schlangen, S. André, S. Schmidt, S. M. Walzer, B. Kubista, R. Windhager, S. Toegel, H.-J. Gabius, *Sci. Rep.* **2016**, *6*, 39112; d) D. Weinmann, M. Kenn, S. Schmidt, K. Schmidt, S. M. Walzer, B. Kubista, R. Windhager, W. Schreiner, S. Toegel, H.-J. Gabius, *Cell. Mol. Life Sci.* **2018**, *75*, 4187–4205; e) K. M. Pichler, D. Weinmann, S. Schmidt, B. Kubista, R. Lass, L. Martelanz, J. Alphonsus, R. Windhager, H.-J. Gabius, S. Toegel, *Calcif. Tissue Int.* **2021**, *108*, 377–390.
[100] a) J. Kopitz, C. von Reitzenstein, M. Burchert, M. Cantz, H.-J. Gabius, *J. Biol. Chem.* **1998**, *273*, 11205–11211; b) J. Kopitz, C. von Reitzenstein, S. André, H. Kaltner, J. Uhl, V. Ehemann, M. Cantz, H.-J. Gabius, *J. Biol. Chem.* **2001**, *276*, 35917–35923; c) N. Ahmad, H.-J. Gabius, S. André, H. Kaltner, S. Sabesan, R. Roy, B. Liu, F. Macaluso, C. F. Brewer, *J. Biol. Chem.* **2004**, *279*, 10841–10847; d) J. Kopitz, S. Vértesy, S. André, S. Fiedler, M. Schnölzer, H.-J. Gabius, *Biochimie* **2014**, *104*, 90–99; e) J. Kopitz, Q. Xiao, A.-K. Ludwig, A. Romero, M. Michalak, S. E. Sherman, X. Zhou, C. Dazen, S. Vértesy, H. Kaltner, M. L. Klein, H.-J. Gabius, V. Percec, *Angew. Chem. Int. Ed.* **2017**, *56*, 14677–14681; *Angew. Chem.* **2017**, *129*, 14869–14873; f) A.-K. Ludwig, M. Michalak, A. Gabba, T. J. Kutzner, D. Beckwith, F. G. FitzGerald, G. García Caballero, J. C. Manning, M. Kriegsmann, H. Kaltner, P. V. Murphy, M. Cudic, J. Kopitz, H.-J. Gabius, *Histochem. Cell Biol.* **2021**, DOI https://doi.org/10.1007/s00418-021-02004-w.

Manuscript received: July 2, 2021
Revised manuscript received: September 7, 2021
Accepted manuscript online: September 8, 2021
Version of record online: September 22, 2021