Biophysical Journal
Article

Biophysical Society

From examining the relationship between (corona) viral adhesins and galectins to glyco-perspectives

Michael L. Klein,[1,*] Antonio Romero,[2] Herbert Kaltner,[3] Virgil Percec,[4] and Hans-Joachim Gabius[3,*]
[1]Institute of Computational Molecular Science, Temple University, Philadelphia, Pennsylvania; [2]Department of Structural and Chemical Biology, CIB Margarita Salas, CSIC, Madrid, Spain; [3]Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University Munich, Munich, Germany; and [4]Roy & Diana Vagelos Laboratories, Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania

ABSTRACT Glycan-lectin recognition is vital to processes that impact human health, including viral infections. Proceeding from crystallographical evidence of case studies on adeno-, corona-, and rotaviral spike proteins, the relationship of these adhesins to mammalian galectins was examined by computational similarity assessments. Intrafamily diversity among human galectins was in the range of that to these viral surface proteins. Our findings are offered to inspire the consideration of lectin-based approaches to thwart infection by present and future viral threats, also mentioning possible implications for vaccine development.

SIGNIFICANCE The coating of biological surfaces by glycans is important for recognition processes. Lectins, a ubiquitous superfamily of glycan-binding proteins classified by the type (fold) of the carbohydrate recognition domain(s), enable bridging of biological surfaces. Ga(lactose-binding)lectins are present in all animals. The discovery of fold similarity between virus attachment proteins (adhesins) and mammalian galectins prompted us to calculate similarity scores systematically, leading to describing perspectives to counter viral threats: one could develop therapeutics that inhibit contact between viral adhesins and their cellular counterreceptors, as inspired by experience gleaned from studying (ga) lectins, and these similarities may well be relevant for vaccine development.

INTRODUCTION

Biological surfaces contain a "sugary coating" (1) of complex carbohydrates or heterosaccharides (2–4). The unsurpassed level of structural diversity of glycan chains, which are attached to proteins and/or lipid anchors, is the molecular basis for the popular concept of the sugar code (5,6). While the "letters" of this third alphabet of life (i.e., carbohydrates), like nucleotides or amino acids (aas), form the molecular messages (i.e., glycans), "readers" and "translators" of glycan-encoded information have evolved to make functional coupling possible (4,7,8). Indeed, their ability to "select" (or to "read") is the origin of the term "lectin" (from the Latin word legere) (9,10). Communication via lectins and their counterreceptors is emerging as a general theme that underlies many processes in pathology and physiology, also of relevance for pandemic threats, so that its understanding may even guide us to new (glycobiological) ways to fight infections.

More than a dozen folds have acquired capacity to bind carbohydrate ligands, and structural diversification has let families grow in numbers (11). This toolbox serves nature's demand for multipurpose effectors also involved in the first step of bacterial and viral infections. Strategically, the tips of bacterial pili and viral spikes equip them with the needed molecular tentacles for docking. Evolutionarily, instead of an independent development of such adhesins, the alternative is to add such genes from the host to viral genomes, where they will then undergo often rapid sequence changes. Irrespective of their origin, the detection of shared sequence and/or folding characteristics is the starting point to calculate respective scores, and this relationship can help guide efforts to find biopharmaceutical ways to block adhesion by analogy with tissue lectins.

Starting this route, sequence alignments of respective virus glycoproteins that initiate the infection process had disclosed similarity to carbohydrate recognition domains (CRDs) of mammalian lectins, specifically for C-type

Submitted August 25, 2020, and accepted for publication November 17, 2020.

*Correspondence: mlklein@temple.edu or gabius@tiph.vetmed.uni-muenchen.de

Editor: Tamar Schlick.
https://doi.org/10.1016/j.bpj.2020.11.020

lectin-like domains (12,13). Of relevance for this coronavirus pandemic, the rotavirus outer capsid spike protein (VP4), which is responsible for cell attachment and membrane penetration, was proposed to have gained receptor capacity by the insertion of another type of host-derived CRD into an ancestral membrane interaction protein, i.e., the galectin (Gal) (β-sandwich-type) domain followed by sequence alterations relevant for viral tropism and pathogenicity (14). Intriguingly, sequence analysis of the putative fiber protein of porcine adenovirus isolate NDAC-1 also detected significant homology to the galectins (15). They are a class of multifunctional proteins present in all animals (16–20). Emerging evidence from computational and panel testing (e.g., by molecular dynamics simulations (21) or by frontal chromatography (22)) as well as from counterreceptor identification suggests that each galectin has its own set of context-dependent binding partners for a special outcome like the induction of anoikis or apoptosis such as a glycoprotein (e.g., an integrin as the $\alpha_5\beta_1$-integrin or CD7) or a glycosphingolipid such as ganglioside GM1 (for details, please see (17); for explanations of lectin selectivity for binding partners on six levels, please see (23)). These data prompted us to examine the matter of the adhesin-galectin relationship in more detail by systematic similarity calculations, starting by defining the level of intrafamily diversification in the host, to set a standard.

Herein, we illustrate the extent of natural diversity among galectins using the human (and also bat as the assumed severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) host) proteins as an instructive example. These intrafamily comparisons affirm that sequence deviations in the structural platform of the galectin CRD led to a panel of homologous glycan- and peptide-binding proteins. Next and importantly, we document the results of sequence comparisons and secondary-structure similarity calculations to β-sandwich-type adhesins using a human galectin (hGal) as standard. Notably, scores for similarity among certain galectin-adhesin pairs do not differ much from those between certain galectin pairs, giving us reason to proceed to envisioning biopharmaceutical perspectives. In this part (in the Discussion), we argue that systematic consideration of the general glycobiology of virus-host interplay could likely prove helpful to detect a possible weakness, i.e., an Achilles heel, in the virus attachment strategy. This, in turn, might give a direction to find ways to thwart viral infection.

MATERIALS AND METHODS

Alignment of sequences and calculations of similarity

Multiple alignments of aa sequences (human and bat galectins and viral spike proteins) were performed using the European Molecular Biology – European Bioinformatics Institute (EMBL-EBI) Clustal Omega software (24) (www.ebi.ac.uk/Tools/msa/clustalo/), which is included in the Align tool (www.uniprot.org/align/) of the UniProt database. Calculations (please see Figs. 2, 4, and S4, A–C, and also the respective figure legends) were done using the BLAST (www.uniprot.org/blast/) and Align tools. Scoring matrix was PAM (point accepted mutation) 250, the term identity indicates positions that have a single, fully conserved residue, the term similarity indicates conservation between groups of strongly (scoring $>$ 0.5), and/or weakly (scoring $\leq$ 0.5) similar properties. The following sequences were based on ID entries in the UniProt Knowledgebase, UniProtKB, ExPASy Proteomics Server (www.expasy.org) were used (Figs. 2, S2, and S4, *A–C*; please see also the figure legends of Figs. 1 and 2, plus Fig. S2): hGal-1 (UniProt: P09382, aa 1–135), hGal-3 CRD (UniProt: P17931: aa 17–250) and hGal-8C (UniProt: O00214: aa 185–317), hGal-2 (UniProt: P05162: aa 1–132), hGal-4N (UniProt: P56470: aa 1–148), hGal-4C (UniProt: P56470: aa 206–323), hGal-7 (UniProt: P47929: aa 1–136), hGal-8N (UniProt: O00214: aa 1–152), hGal-9N (UniProt: O00182: aa 1–146), hGal-9C (Uniprot: O00182: aa 224–355), hGal-10 (Uniprot: Q05315: aa 1–142), hGal-12N (UniProt: Q96DT0: aa 1–181), hGal-12C (UniProt: Q96DT0: aa 209–336), hGal-13 (UniProt: Q9UHV8: aa 1–139), h*Gal-r*elated inter*fiber* prote*in* (GRIFIN; UniProt: A4D1Z8: aa 1–144), h*Gal-r*elated *p*rotein (GRP; UniProt: Q3ZCW2: aa 1–172); *Rhinolophus ferrumequinum* (bat) Gal-3 (UniProt: A0A671DPG4: aa 1–283), human severe

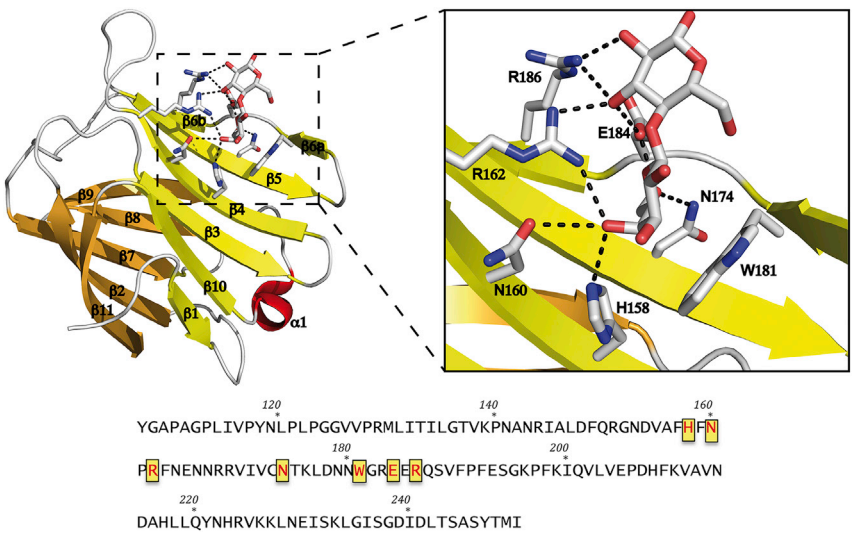

FIGURE 1 Illustration of the crystal structure (*top*; PDB: 6FOF) and aa sequence (*bottom*; UniProt: P17931) of the human galectin (hGal)-3 CRD (aa 107–250). The ribbon diagram is shown top left. The galectin signature sequence of highly conserved aas, which are involved in binding cognate β-galactosides, is highlighted with red letters on a yellow background. The boxed image on the top right shows the CRD with higher-level magnification. To see this figure in color, go online.

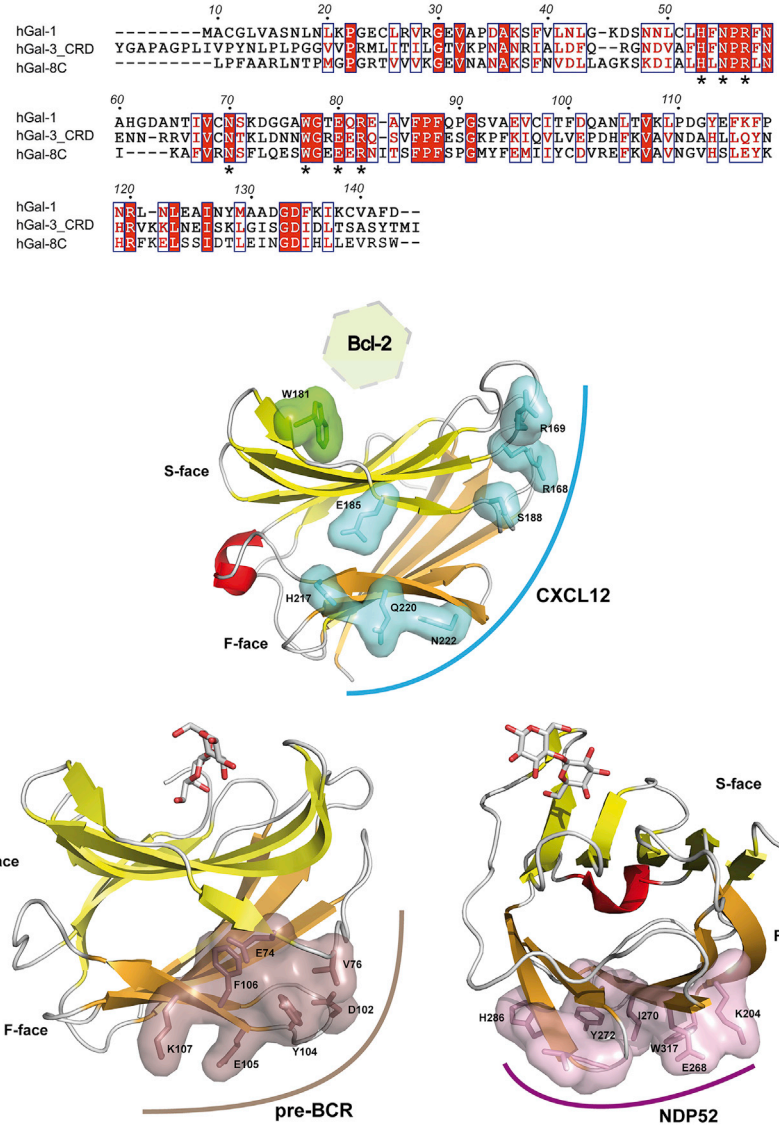

FIGURE 2 Multiple aa sequence alignment (Clustal Omega software; *top*) of hGal-1 (UniProt: P09382), hGal-3 CRD (UniProt: P17931), and hGal-8C (UniProt: O00214). Strictly conserved (*red background*) and homologous residues (>70% conservation; *boxed red letters*) are highlighted by coloring. The aas in contact with the canonical ligand are marked by asterisks. Illustrated in the middle and bottom, respectively, are hGal-3 CRD in interaction with Bcl-2 (based on suggestions for the protein-protein contact via the NWGR motif (25)) or CXCL12 (26) at different sites, hGal-1 with pre-BCR (27), and the recognition site of hGal-8C for the autophagic receptor NDP52 (PDB: 4HAN). The location of carbohydrate-binding sites of hGal-1 and of hGal-8C in the proteins' S-face are given by lactose. The coloring of the protein-binding region (for a distinct binding partner by a *line* and by including selected aas in *coloring*) highlight these galectins' inherent bifunctionality. To see this figure in color, go online.

acute respiratory syndrome coronavirus (SARS-CoV) spike glycoprotein (UniProt: P59594, SARS-CoV N-terminal domain (NTD): aa 14–292 and C-terminal domain (CTD): aa 317–507), and human SARS-CoV-2 (2019-nCoV) spike glycoprotein (UniProt: P0DTC2, SARS-CoV-2 NTD: aa 16–305 and CTD: aa 330–521), Middle-East-respiratory-syndrome-related coronavirus spike glycoprotein (UniProt: A0A0D3MU71, Middle East respiratory syndrome coronavirus (MERS-CoV) NTD: aa 18–353 and CTD: aa 382–503), bovine coronavirus spike glycoprotein NTD (UniProt: Q1HLC5, bovine coronavirus NTD: aa 8–284), murine hepatitis virus spike glycoprotein NTD (UniProt: Q9J3E7, mouse hepatitis CoV NTD: aa 16–290), porcine adenovirus type 4 NADC-1 isolate fiber galectin domain (UniProt: Q83467, NTD CRD, NADC-1 NTD: aa 394–523), rhesus rotavirus VP4 sialic-acid-binding domain (UniProt: P12473; rhesus rotavirus VP8*: aa 64–224), and human rotavirus VP4 sialic acid-binding domain (UniProt: Q86169; human rotavirus VP8*: aa 64–224).

Construction of the phylogenetic tree

The following sequences of hGals from identifier entries in the UniProt Knowledgebase, ExPASy Proteomics Server (www.expasy.org), were used to construct the phylogenetic tree (Fig. S2; please see also legend to Fig. S2): hGal-1 (UniProt: P09382: aa 1–135), hGal-3 CRD (UniProt: P17931: aa 17–250) and hGal-8C (UniProt: O00214: aa 185–317), hGal-2 (UniProt: P05162: aa 1–132), hGal-4N (UniProt: P56470: aa 1–148), hGal-4C (UniProt: P56470: aa 206–323), hGal-7 (UniProt: P47929: aa 1–136), hGal-8N (UniProt: O00214: aa 1–152), hGal-9N (UniProt: O00182: aa 1–146), hGal-9C (UniProt: O00182: aa 224–355), hGal-10 (UniProt: Q05315: aa 1–142), hGal-12N (UniProt: Q96DT0: aa 1–181), hGal-12C (UniProt: Q96DT0: aa 209–336), hGal-13 (UniProt: Q9UHV8: aa 1–139), hGRIFIN (GRIFIN; UniProt: A4D1Z8: aa 1–144), and hGRP (GRP; UniProt: Q3ZCW2: aa 1–172). Analysis of evolutionary relationships and construction of the phylogenetic tree for the members of the galectin family was done using the maximal likelihood method implemented in the MEGA X software package (28,29). The initial tree(s) for the heuristic search were obtained by applying the neighbor-joining method to a matrix of pairwise distances estimated using the Jones-Taylor-Thornton model as substitution. The test of phylogeny was performed applying bootstrap analysis (with 1000 replicates) and calculating a bootstrap consensus tree. The percentage of tree(s), in which the associated sequences clustered together, is presented next to the branch nodes.

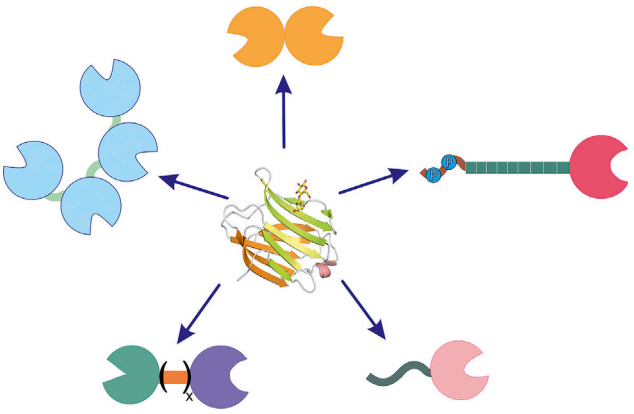

FIGURE 3 Modular architecture of galectins. The galectin CRD (β-sandwich fold) in the center, here hGal-3 CRD (PDB: 6FOF) chosen as an example, is presented in five types of architectural display. Shown clockwise are the following: a noncovalently associated homodimer (*top*), a chimera with collagen-like repeats mediating self-association (nine in humans) and an N-terminal peptide with two sites of serine phosphorylation, an N-tailed proto-type-like (monomeric) protein, a linker-connected (hetero)dimer of two different domains the length of the linker being variable among such proteins symbolized by "x," and the heterotetrameric design found in oysters. To see this figure in color, go online.

Secondary-structure comparison and visualization

Coordinates of the different vertebrate galectins and viral proteins were taken from the Protein Data Bank (PDB) (30); for PDB entries, see legends of Figs. 1, 2, 3, and 4, plus Figs. S1 and S4 *D*. Analyses of the protein-ligand interactions were performed using the Protein Interfaces, Surfaces and Assemblies (PISA) web server (31). The structures were visually inspected using the Crystallographic Object-Oriented Toolkit (COOT) (32) and PyMOL Molecular Graphics System programs (33). Structural superpositions were performed with the secondary-structure-matching (SSM) routine (34) incorporated in COOT. To calculate structural similarity, several score parameters for different scoring functions were applied. The Z-score measures the statistical significance of a pairwise match in terms of Gaussian statistics so that significant similarities reach a high level of Z-score values. The root mean-square deviation measures the deviation between C_α-atoms of matched residues between both structures. The secondary structure shows similarity using the SSM tool (34). Figs. 1, 2, 3, and 4, plus Figs. S1 and S4 *D* were generated applying the PyMOL Molecular Graphics System.

RESULTS

Versatility of the galectin fold

The canonical glycan specificity of galectins to lactose and its derivatives is based on a sequence signature. Its hallmark is a central Trp for C-H and π-bonding with the B-face of galactose in the core of cognate glycans (Fig. 1). Sequence diversification at neighboring sites or within this set of aas lets family members acquire individual glycan specificity profiles (Fig. 1; Fig. S1). This process can proceed to affect several positions of the essential sequence signature, hereby in consequence harming the canonical binding capacity to β-galactosides so that the phylogenetic tree for galectins includes more family members than lactose-binding proteins (Fig. S2). In mammalian galectins, affinity to distinct (glyco)proteins as counterreceptors via protein-protein contacts has also developed at the canonical site and at sites different from the canonical contact region (Fig. 2). hGal-1, -3, and -8 (Gal-1, -3, and -8) and also other members of this family, such as Gal-7 (35) or the GRIFIN (36,37), demonstrate that the galectin fold exhibits a plasticity for choice of binding partner and—explicitly—a capacity to recognize proteins, a point to note with respect to coronaviral adhesins, especially SARS-CoV-2. In addition to sequence variations, the galectin architecture, i.e., the modularity of CRD presentation, as illustrated in Fig. 3, is a key factor for specific effector activity.

CRD-counterreceptor complementarity, galectin architecture, and binding-site presentation, e.g., in microdomains, seem to be relevant for homing in on certain glycoconjugates and then triggering distinct postbinding effects. Examples include Gal-1-specific binding to complex-type N-glycans of the $\alpha_5\beta_1$-integrin for induction of anoikis in tumor cells and, of note for coronaviral counterreceptor selection, Gal-4 association with the glycoproteins carcinoembryonic antigen (CEA), dipeptidyl peptidase-IV (DPP-IV), or aminopeptidase-N (APN) and sulfatide for their apical transport and delivery in enterocytes (Fig. S3; see also below). Facilitating firm and specific contacts, reaching up to nanomolar affinity, and letting evolution create new binding sites on a structurally stable fold make the galectin CRD an experienced and attractive platform for use in viral attachment to host cells.

Galectin-like viral proteins

Discovery of the galectin fold in the N-terminal part of the rhesus rotavirus outer capsid spike protein (14) was followed by further detection of such a domain in viral surface proteins. Notably, for the rotaviral spike protein, a spatial shift of the sugar-binding site from its location in a galectin to a cleft region was found, along with accommodation of sialic acids there, plus adaptation to allow docking of a potent ligand of hGal-3, namely the histo-blood group A antigen (Fig. 4). Corresponding sequence comparisons including galectin-pair calculations are given in Fig. S4, *A–C*; notably, Gal-12C is closest in similarity score, with a secondary-structure similarity of 44% (rhesus) and 55% (human) not much different from that between Gal-1 and -13 of 62%.

The fiber head of porcine adenovirus (type 4) presents two galectin-like domains separated by a linker of 23 residues (38). This display mimics the host tandem-repeat-type galectin architecture shown in Fig. 3. In this case, *N*-acetyllactosamine (LacNAc) and its oligomers bind to the N-terminal CRD at the canonical site (Fig. 4); fittingly, a conserved signature is seen in sequence alignment, and the score is highest in its pairing with Gal-3 (Figs. 1 and 2; Fig. S4 *A*). Because the adenovirus fiber is trimeric, as with endocytic

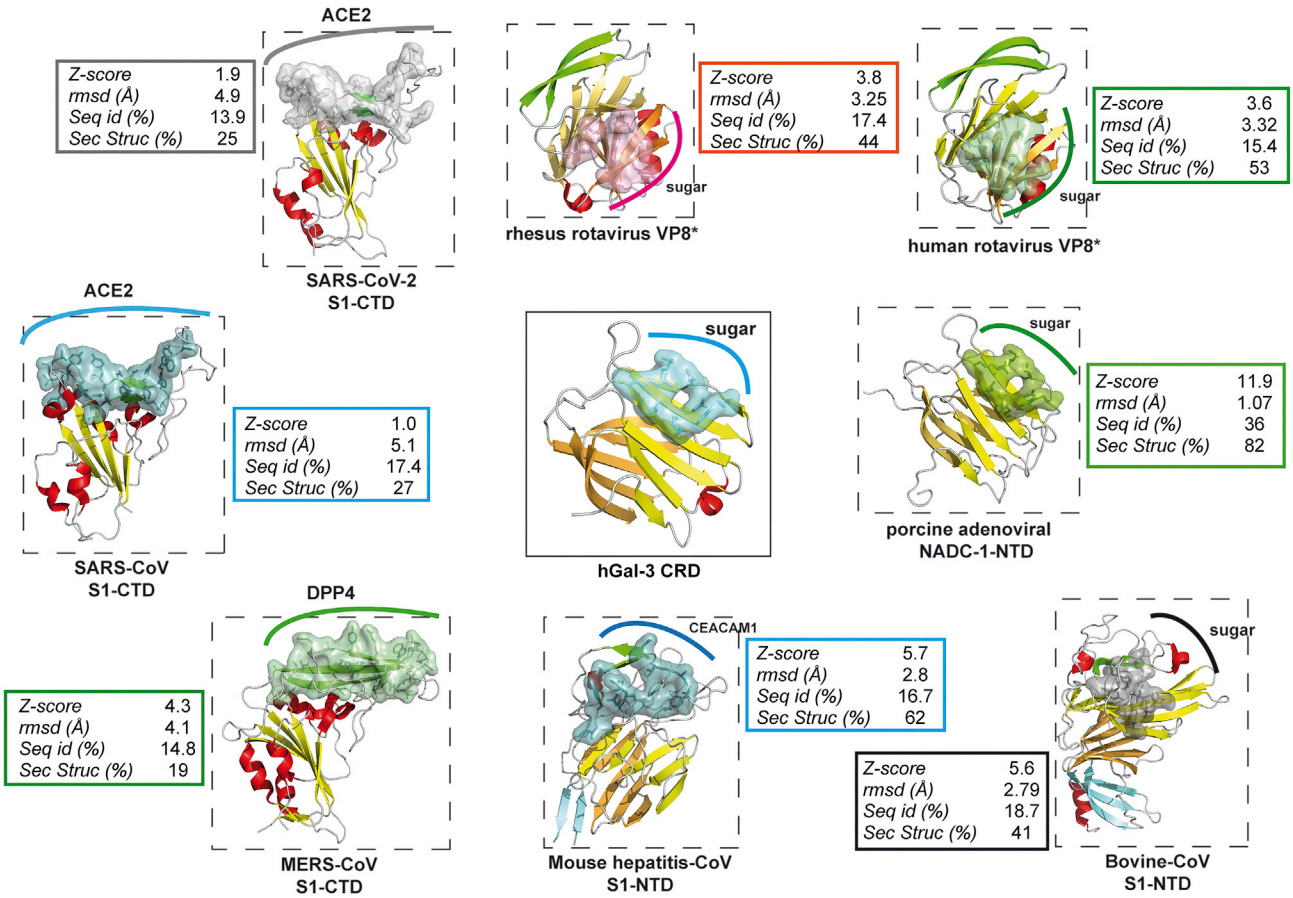

FIGURE 4 Carbohydrate recognition domains (CRDs) and receptor-binding domains (RBDs) (sites for binding drawn in semitransparent surface presentation with crucial aas presented as sticks; contact region between RBD and its (sugar or protein) binding partner highlighted as *colored curved line*). hGal-3 (PDB: 6FOF), in a complex with lactose, is in the center of the figure. Viral proteins are arranged clockwise around the hGal, starting from the top: rhesus rotavirus VP4 sialic acid-binding domain (rhesus rotavirus VP8*, PDB: 1KQR) in complex with 2-*O*-methyl-α-D-*N*-acetylneuraminic-acid, human Wa rotavirus VP8* carbohydrate-recognizing domain (human rotavirus VP8*, PDB: 2DWR) bound to methyl-α-D-N-acetylneuraminic acid, porcine adenoviral NADC-1 (NTD, PDB: 2WSV) isolate fiber in complex with lactose, the lectin domain of bovine coronavirus spike protein (bovine coronavirus S1-NTD, PDB: 4H14), mouse coronavirus RBD (mouse hepatitis coronavirus S1-NTD, PDB: 3R4D) complexed with its receptor mCEACAM1a, Middle-East-respiratory-syndrome-related MERS-CoV spike RBD complexed with its receptor DPP-IV (MERS-CoV S1-CTD, PDB: 4KR0), SARS-CoV spike RBD (SARS-CoV S1-CTD, PDB: 2AJF) complexed with its receptor ACE2 and 2019-nCoV RBD (SARS-CoV-2 S1-CTD, PDB: 6M0J) in complex with ACE2. Given in boxes are the quantitative measures: Z-score, the statistical metric parameter for extent of fold similarity; rmsd, the root mean-square deviation; Seq id, the sequence identity calculation from the Basic Local Alignment Search Tool BLAST (www.uniprot.org/blast/) and Align (www.uniprot.org/align/) tools of the UniProt database; Sec Struc, the secondary-structure similarity using the SSM tool (23). The scoring matrix was PAM (point accepted mutation) 250, identity indicates positions, which have a single fully conserved residue, and similarity indicates conservation between groups of strongly (scoring > 0.5) and/or weakly (scoring ≤ 0.5) similar properties. To see this figure in color, go online.

tissue lectins, e.g., the hepatic asialoglycoprotein receptor, and 12 fibers populate the surface, there are multiple sticky docking options to attach to a host cell surface.

In analogy to mammalian selectins and siglecs (39), the receptor-binding domain (RBD) of the trimeric coronaviral prefusion spike is located at the tip. As is the case for the aforementioned adenovirus, it consists of two structural units referred to as NTDs (most distal) and CTDs, like the tandem-repeat-type galectin design shown in Fig. 3 (*bottom right*). Strategically, both can act alone or in concert and acquire individual features by dynamic diversification. The α- and β-coronaviral NTD is most similar in scores to the galectin fold, binding glycans in a species-specific manner, i.e., 9-*O*-acetylated sialic acid and α2,6-sialylated mucin-type core 1 (Thomsen-Friedenreich antigen, CD176) *O*-glycan, and also a peptide motif of a glycoprotein, i.e., CEA-related cell adhesion molecule 1 (Fig. 4; Fig. S4, *A* and *B*). The second CTD, more distant to galectins, as shown in Fig. S4, *A* and *B* and also by secondary-structure illustration in Fig. S4 *D*, harbors the infamous affinity to glycoproteins, namely APN, DPP-IV, and angiotensin-converting enzyme 2 (ACE2) (Fig. 4). Importantly, the high degree of homology among mammalian galectins and the availability of crystallographical structures for the human proteins allow an hGal reference point instead of a galectin of the true host of recent pandemic malefactors

(e.g., bat). To illustrate this point, hGal-3 CRD has 86.7% sequence identity and an additional 10.4% sequence similarity to its homolog in the greater horseshoe bat (*R. ferrumequinum*), which is in the same genus as the coronavirus hosts *R. macrotis* and *R. sinicus*. Nonetheless, obtaining crystallographic information on bat galectins would seem very timely and relevant.

At this point, the question arises on the degree of sequence similarity between hGals relative to that between pairs of an hGal and a viral adhesin. When, therefore, comparing sequences among members of the family of hGals, cases are revealed with an extent of deviation seen between hGals and viral adhesins (Fig. S4 *C*), as noted above for secondary-structure elements. Specifically, Gal-1 and the GRP, a family member with deviations from the sequence signature (40,41), share less identity and similarity positions (47.1%) than Gal-1 and SARS-CoV-2 NTD (50.3%) or CTD (52.6%) (Fig. S4, *A–C*). The Gal-12C CRD is a similar case, with scores in this range in comparison with the tested set of galectins (see Fig. S4 *C*).

DISCUSSION

CRDs of cell surface lectins are frequently presented on protein stalks to ensure spatial accessibility as viral spikes do. Functionally, the galectin CRD is a means to build bridging molecules, ideal for recognition between surfaces (42,43). Similarly docking on glycans, viral adhesins thus harbor lectin capacity (44) and, as the coronavirus case teaches, have developed protein binding on the platform of the galectin fold. Recruitment of galectin-like domains to the initial step of coronavirus infection provides critical determinants of the host range and tissue tropism, having given reason to propose an origin of the clinical consequences by a gene capture (14,45,46), and qualifies to become a target for vaccination by inducing blocking antibodies. In that case, it is mandatory to have checked the similarity score to tissue proteins to be sufficiently low. Intriguingly, the presence of anti-galectin autoantibodies is known in association with certain disorders, e.g., autoimmune diseases, providing a likely source of information on antigenic sites on the galectin CRD platform (47,48). Moreover, this insight invites a systematic consideration of fatal bridging by galectin-like domains as well as of glycosylation of virus and host proteins as sites to be able to interfere with undesirable contacts. After all, endogenous lectins are among the effectors of innate immunity, protecting us against microbial challenges by recognizing pathogen-associated molecular patterns (49–54). Tactically, an invader should be confronted with restrictions to its entry or lured into binding to smart decoy receptors.

A possible in vitro strategy is to saturate the viral receptor domain by a specific high-affinity compound, thereby precluding docking. Conceptually, this antiadhesion approach is inspired by the classical hemagglutination inhibition assay (55,56) and is discussed in the context of clinically harmful activities of endogenous galectins (57): Sialic acid, for rhesus rotavirus VP8* (58); histo-blood group A tetrasaccharide, for the human rotaviral receptor (59,60); LacNAc oligomers, for the porcine adenovirus type 4 fiberhead (38); *N*-acetyl(glycolyl)neuraminic acid and its 9-*O*-acetylated derivative, for β- and likely some α- and γ-genus coronaviruses (61–65), also for a δ-coronavirus NTD binding mucin (66) as well as sialyl T_n antigen (CD175s) for the feline infectious peritonitis α-coronavirus (67) have all emerged as respective candidates and starting points for structural optimization to minimize cross-reactivity to tissue lectins.

If the binding partner is a protein as shown in Fig. 2 or Fig. 4, then this principle can swiftly be adapted: conformationally restricted peptides, instead of the sugar, may fulfill the same purpose by occupying sites of contact with cellular protein counterreceptors, e.g., either APN, CEACAM1, DPP-IV, or ACE2. The characterization of contact points and generators of affinity, e.g., in the case of ACE2 recognition by SARS-CoV-2 (68), is a source for gaining optimal complementarity, and this principle can be readily adapted to cases of cell adhesion molecules as viral entry site (69). Supramolecular chemistry also offers an attractive concept, i.e., intercepting the virus on its way to a cell by a high-density display of cognate compounds, presented either as soluble dendrimer constructs or suitably tailored nanoscale dendrimersomes (70,71).

Based on the nature of the mutual contact sites, as illustrated in Fig. 4, epitope masking of the cellular determinants deserves attention. Toward this end, either the viral receptor domain, which may even evoke an immune response, or minimally sized peptides with receptor properties, as identified for hGal-1 and -3 by proteolytic on-site excision and mass spectrometric mapping (72), could serve this purpose. Making the cellular binding partner invisible for the viral receptor can alternatively be accomplished by a tissue lectin that is resident before a virus moves in. Candidates are hGal-1, -3, or -9 for Lac*N*Ac oligomers or the CRD of the macrophage galactose-type lectin (named MGL, CD301, or CLEC10A) (73) and siglec-2, -3, -5, and -6 (74) for the sialyl T_n antigen. In combination with defense molecules working at a different level of protection such as a θ-defensin (retrocyclin 2), which is building a protective barricade of immobilized surface proteins (75), these resident tissue lectins may result in additive effects.

In an indirect manner, the nature of the cellular glycoprotein glycosylation warrants attention. In principle, glycan chains can let lectins dock, and their association can then automatically impede other contacts in their vicinity. Hereby exploiting glycans as markers together with the target specificity of tissue lectins, *N*- and *O*-glycans of cellular glycoprotein counterreceptors for viral adhesins may direct receptor proteins to strategic sites at the cellular surface, as consequence blocking viral attachment by this homing in on

their preferred target, the mentioned Gal-4 with its binding to APN, CEA, and DPP-IV (76–78) being a candidate. Thus, it is of interest to map the glycan profiles of cellular counter-receptors of coronaviruses, ideally obtained for in situ material to exclude source-dependent variations of glycan structures, with such a perspective in mind. Conversely, lectins from either the host's defense system (49,51–54) or from external sources (79) can become guardians against viruses by connecting to viral glycans.

To identify suitable lectins, it is necessary to determine viral glycan signatures, often an essential shield against immune surveillance. Interestingly, in the special case of a cross-reactive antibody for SARS-CoV and SARS-CoV-2, affinity was found to be increased by an N-glycan at N370 (80). Because the glycan chains of viral surface glycoproteins are assembled by the host's enzymatic machinery, they are much less likely to serve as vaccines than bacterial glycocompounds. Lectins, however, may find an Achilles heel on the virus surface. Stepwise refinements of capacity for tracing viral surface features through tissue lectins by mutational tuning, architecture engineering, and chimera design has potential to create variants with favorable properties, for example for collectins, defensins, or galectins (81–85).

CONCLUSION

Examining the functional analogy and the similarity scores between (corona)viral adhesins and mammalian galectins has led us to view the initial step of viral infection from the perspective of glycobiology. Doing so suggests the possibility of exploring lectins or lectin-like domains, glycans, or peptides as both targets and tools in efforts toward inhibiting coronaviral infection.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.11.020.

AUTHOR CONTRIBUTIONS

M.L.K. and H.-J.G. designed the research. A.R. and H.K. performed the research. A.R. and H.K. contributed analytic tools. A.R., H.K., V.P., M.L.K., and H.-J.G. analyzed the data. M.L.K. and H.-J.G. wrote the article.

ACKNOWLEDGMENTS

We gratefully acknowledge inspiring discussions with Drs. B. Friday, A. Leddoz, and A. W. L. Nose, and the valuable input by the three reviewers.

We acknowledge support from European Cooperation in Science and Technology Action CA18103 (InnoGly). V.P. and M.L.K. are thankful for the generous support of the P. Roy Vagelos Chair at University of Pennsylvania and the Sheikh Saqr Foundation, RAK, respectively.

REFERENCES

1. Bennett, H. S. 1963. Morphological aspects of extracellular polysaccharides. *J. Histochem. Cytochem.* 11:14–23.
2. Ginsburg, V., and E. F. Neufeld. 1969. Complex heterosaccharides of animals. *Annu. Rev. Biochem.* 38:371–388.
3. Sharon, N. 1975. Complex Carbohydrates. Their Chemistry, Biosynthesis, and Functions. Addison-Wesley Publ. Co., Reading, MA.
4. Kaltner, H., J. Abad-Rodríguez, …, H.-J. Gabius. 2019. The sugar code: letters and vocabulary, writers, editors and readers and biosignificance of functional glycan-lectin pairing. *Biochem. J.* 476:2623–2655.
5. Roseman, S. 2001. Reflections on glycobiology. *J. Biol. Chem.* 276:41527–41542.
6. Gabius, H.-J., and J. Roth. 2017. An introduction to the sugar code. *Histochem. Cell Biol.* 147:111–117.
7. Gabius, H.-J. 2018. The sugar code: why glycans are so important. *BioSystems.* 164:102–111.
8. Dedola, S., M. D. Rugen, …, R. A. Field. 2020. Revisiting the language of glycoscience: readers, writers and erasers in carbohydrate biochemistry. *ChemBioChem.* 21:423–427.
9. Boyd, W. C., and E. Shapleigh. 1954. Specific precipitating activity of plant agglutinins (lectins). *Science.* 119:419.
10. Kaltner, H., G. García Caballero, …, H.-J. Gabius. 2018. From glycophenotyping by (plant) lectin histochemistry to defining functionality of glycans by pairing with endogenous lectins. *Histochem. Cell Biol.* 149:547–568.
11. Manning, J. C., A. Romero, …, H.-J. Gabius. 2017. Lectins: a primer for histochemists and cell biologists. *Histochem. Cell Biol.* 147:199–222.
12. Zelensky, A. N., and J. E. Gready. 2005. The C-type lectin-like domain superfamily. *FEBS J.* 272:6179–6217.
13. Gready, J. N., and A. N. Zelensky. 2009. Routes in lectin evolution: case study on the C-type lectin-like domains. *In* The Sugar Code. Fundamentals of Glycosciences. H.-J. Gabius, ed. Wiley-VCH, pp. 329–346.
14. Dormitzer, P. R., Z. Y. Sun, …, S. C. Harrison. 2002. The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J.* 21:885–897.
15. Kleiboeker, S. B. 1995. Sequence analysis of the fiber genomic region of a porcine adenovirus predicts a novel fiber protein. *Virus Res.* 39:299–309.
16. Barondes, S. H. 1984. Soluble lectins: a new class of extracellular proteins. *Science.* 223:1259–1264.
17. Kaltner, H., S. Toegel, …, H.-J. Gabius. 2017. Galectins: their network and roles in immunity/tumor growth control. *Histochem. Cell Biol.* 147:239–256.
18. Hirabayashi, J., (ed) 2018. Special issue on galectins. *Trends Glycosci. Glycotechnol.* 30:SE1–SE223.
19. Godula, K. 2018. Following sugar patterns in search of galectin function. *Proc. Natl. Acad. Sci. USA.* 115:2548–2550.
20. García Caballero, G., H. Kaltner, …, H.-J. Gabius. 2020. How galectins have become multifunctional proteins. *Histol. Histopathol.* 35:509–539.
21. Ford, M. G., T. Weimar, …, R. J. Woods. 2003. Molecular dynamics simulations of galectin-1-oligosaccharide complexes reveal the molecular basis for ligand diversity. *Proteins.* 53:229–240.
22. Iwaki, J., and J. Hirabayashi. 2018. Carbohydrate-binding specificity of human galectins: an overview by frontal affinity chromatography. *Trends Glycosci. Glycotechnol.* 30:SE137–SE153.
23. Gabius, H.-J., S. André, …, D. Solís. 2011. From lectin structure to functional glycomics: principles of the sugar code. *Trends Biochem. Sci.* 36:298–313.

24. Sievers, F., A. Wilm, …, D. G. Higgins. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539.

25. Akahani, S., P. Nangia-Makker, …, A. Raz. 1997. Galectin-3: a novel antiapoptotic molecule with a functional BH1 (NWGR) domain of Bcl-2 family. *Cancer Res.* 57:5272–5276.

26. Eckardt, V., M. C. Miller, …, K. H. Mayo. 2020. Chemokines and galectins form heterodimers to modulate inflammation. *EMBO Rep.* 21:e47852.

27. Elantak, L., M. Espeli, …, C. Schiff. 2012. Structural basis for galectin-1-dependent pre-B cell receptor (pre-BCR) activation. *J. Biol. Chem.* 287:44703–44713.

28. García Caballero, G., A. Flores-Ibarra, …, H.-J. Gabius. 2016. Galectin-related protein: an integral member of the network of chicken galectins. 1. From strong sequence conservation of the gene confined to vertebrates to biochemical characteristics of the chicken protein and its crystal structure. *Biochim. Biophys. Acta.* 1860:2285–2297.

29. Kumar, S., G. Stecher, …, K. Tamura. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol. Biol. Evol.* 35:1547–1549.

30. PDBe-KB Consortium. 2020. PDBe-KB: a community-driven resource for structural and functional annotations. *Nucleic Acids Res.* 48:D344–D353.

31. Krissinel, E., and K. Henrick. 2007. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372:774–797.

32. Emsley, P., B. Lohkamp, …, K. Cowtan. 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66:486–501.

33. DeLano, W. 2002. The PyMOL molecular graphics system. Delano Scientifics, Palo Alto, CA.

34. Krissinel, E., and K. Henrick. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr. D Biol. Crystallogr.* 60:2256–2268.

35. Villeneuve, C., L. Baricault, …, F. Larminat. 2011. Mitochondrial proteomic approach reveals galectin-7 as a novel BCL-2 binding protein in human cells. *Mol. Biol. Cell.* 22:999–1013.

36. Ruiz, F. M., U. Gilles, …, H.-J. Gabius. 2018. Chicken GRIFIN: structural characterization in crystals and in solution. *Biochimie.* 146:127–138.

37. García Caballero, G., S. Schmidt, …, H.-J. Gabius. 2020. Chicken lens development: complete signature of expression of galectins during embryogenesis and evidence for their complex formation with α-, β-, δ-, and τ-crystallins, N-CAM, and N-cadherin obtained by affinity chromatography. *Cell Tissue Res.* 379:13–35.

38. Guardado-Calvo, P., E. M. Muñoz, …, M. J. van Raaij. 2010. Crystallographic structure of porcine adenovirus type 4 fiber head and galectin domains. *J. Virol.* 84:10558–10568.

39. Kaltner, H., and H.-J. Gabius. 2019. Sensing glycans as biochemical messages by tissue lectins: the sugar code at work in vascular biology. *Thromb. Haemost.* 119:517–533.

40. Cooper, D. N. W. 2002. Galectinomics: finding themes in complexity. *Biochim. Biophys. Acta.* 1572:209–231.

41. Manning, J. C., G. García Caballero, …, H.-J. Gabius. 2018. Members of the galectin network with deviations from the canonical sequence signature. 2. *G*alectin-*r*elated *p*rotein (GRP). *Trends Glycosci. Glycotechnol.* 30:SE11–SE20.

42. Harrison, F. L., and C. J. Chesterton. 1980. Factors mediating cell–cell recognition and adhesion. Galaptins, a recently discovered class of bridging molecules. *FEBS Lett.* 122:157–165.

43. Barondes, S. H. 1997. Galectins: a personal overview. *Trends Glycosci. Glycotechnol.* 9:1–7.

44. Thompson, A. J., R. P. de Vries, and J. C. Paulson. 2019. Virus recognition of glycan receptors. *Curr. Opin. Virol.* 34:117–129.

45. Li, F. 2015. Receptor recognition mechanisms of coronaviruses: a decade of structural studies. *J. Virol.* 89:1954–1964.

46. Li, F. 2016. Structure, function, and evolution of coronavirus spike proteins. *Annu. Rev. Virol.* 3:237–261.

47. Lutomski, D., R. Joubert-Caron, …, M. Caron. 1997. Anti-galectin-1 autoantibodies in serum of patients with neurological diseases. *Clin. Chim. Acta.* 262:131–138.

48. Sarter, K., C. Janko, …, H.-J. Gabius. 2013. Autoantibodies against galectins are associated with antiphospholipid syndrome in patients with systemic lupus erythematosus. *Glycobiology.* 23:12–22.

49. van de Wetering, J. K., L. M. G. van Golde, and J. J. Batenburg. 2004. Collectins: players of the innate immune system. *Eur. J. Biochem.* 271:1229–1249.

50. Lehrer, R. I., G. Jung, …, W. Lu. 2009. Multivalent binding of carbohydrates by the human α-defensin, HD5. *J. Immunol.* 183:480–490.

51. Sato, S., C. St-Pierre, …, J. Nieminen. 2009. Galectins in innate immunity: dual functions of host soluble β-galactoside-binding lectins as damage-associated molecular patterns (DAMPs) and as receptors for pathogen-associated molecular patterns (PAMPs). *Immunol. Rev.* 230:172–187.

52. Van Breedam, W., S. Pöhlmann, …, H. J. Nauwynck. 2014. Bittersweet symphony: glycan-lectin interactions in virus biology. *FEMS Microbiol. Rev.* 38:598–632.

53. Endo, Y., M. Matsushita, and T. Fujita. 2015. New insights into the role of ficolins in the lectin pathway of innate immunity. *Int. Rev. Cell Mol. Biol.* 316:49–110.

54. Wesener, D. A., A. Dugan, and L. L. Kiessling. 2017. Recognition of microbial glycans by soluble human lectins. *Curr. Opin. Struct. Biol.* 44:168–178.

55. Sharon, N. 1987. Bacterial lectins, cell-cell recognition and infectious disease. *FEBS Lett.* 217:145–157.

56. Sharon, N. 2006. Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta.* 1760:527–537.

57. Romero, A., and H.-J. Gabius. 2019. Galectin-3: is this member of a large family of multifunctional lectins (already) a therapeutic target? *Expert Opin. Ther. Targets.* 23:819–828.

58. Dormitzer, P. R., Z. Y. Sun, …, S. C. Harrison. 2002. Specificity and affinity of sialic acid binding by the rhesus rotavirus VP8* core. *J. Virol.* 76:10512–10517.

59. Hu, L., S. E. Crawford, …, B. V. Prasad. 2012. Cell attachment protein VP8* of a human rotavirus specifically interacts with A-type histo-blood group antigen. *Nature.* 485:256–259.

60. Sun, X., L. Wang, …, Z. Duan. 2018. Human group C rotavirus VP8*s recognize type A histo-blood group antigens as ligands. *J. Virol.* 92:e00442-18.

61. Vlasak, R., W. Luytjes, …, P. Palese. 1988. Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc. Natl. Acad. Sci. USA.* 85:4526–4529.

62. Peng, G., L. Xu, …, F. Li. 2012. Crystal structure of bovine coronavirus spike protein lectin domain. *J. Biol. Chem.* 287:41931–41938.

63. Li, W., R. J. G. Hulswit, …, B. J. Bosch. 2017. Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *Proc. Natl. Acad. Sci. USA.* 114:E8508–E8517.

64. Hulswit, R. J. G., Y. Lang, …, R. J. de Groot. 2019. Human coronaviruses OC43 and HKU1 bind to 9-*O*-acetylated sialic acids via a conserved receptor-binding site in spike protein domain A. *Proc. Natl. Acad. Sci. USA.* 116:2681–2690.

65. Tortorici, M. A., A. C. Walls, …, D. Veesler. 2019. Structural basis for human coronavirus attachment to sialic acid receptors. *Nat. Struct. Mol. Biol.* 26:481–489.

66. Shang, J., Y. Zheng, …, F. Li. 2018. Cryo-electron microscopy structure of porcine δ-coronavirus spike protein in the prefusion state. *J. Virol.* 92:e01556-17.

67. Yang, T. J., Y. C. Chang, …, S. D. Hsu. 2020. Cryo-EM analysis of a feline coronavirus spike protein reveals a unique structure and camouflaging glycans. *Proc. Natl. Acad. Sci. USA.* 117:1438–1446.

68. Shang, J., G. Ye, …, F. Li. 2020. Structural basis of receptor recognition by SARS-CoV-2. *Nature.* 581:221–224.

69. Bhella, D. 2015. The role of cellular adhesion molecules in virus attachment and entry. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140035.

70. Percec, V., P. Leowanawat, …, P. A. Heiney. 2013. Modular synthesis of amphiphilic Janus glycodendrimers and their self-assembly into glycodendrimersomes and other complex architectures with bioactivity to biomedically relevant lectins. *J. Am. Chem. Soc.* 135:9055–9077.

71. Zhang, S., Q. Xiao, …, V. Percec. 2015. Glycodendrimersomes from sequence-defined Janus glycodendrimers reveal high activity and sensor capacity for the agglutination by natural variants of human lectins. *J. Am. Chem. Soc.* 137:13334–13344.

72. Moise, A., S. André, …, H.-J. Gabius. 2011. Toward bioinspired galectin mimetics: identification of ligand-contacting peptides by proteolytic-excision mass spectrometry. *J. Am. Chem. Soc.* 133:14844–14847.

73. Mortezai, N., H. N. Behnken, …, C. Wagener. 2013. Tumor-associated Neu5Ac-T_n and Neu5Gc-T_n antigens bind to C-type lectin CLEC10A (CD301, MGL). *Glycobiology.* 23:844–852.

74. Brinkman-Van der Linden, E. C. M., and A. Varki. 2000. New aspects of siglec binding specificities, including the significance of fucosylation and of the sialyl-T_n epitope. Sialic acid-binding immunoglobulin superfamily lectins. *J. Biol. Chem.* 275:8625–8632.

75. Leikina, E., H. Delanoe-Ayari, …, L. V. Chernomordik. 2005. Carbohydrate-binding molecules inhibit viral fusion and entry by crosslinking membrane glycoproteins. *Nat. Immunol.* 6:995–1001.

76. Danielsen, E. M., and B. van Deurs. 1997. Galectin-4 and small intestinal brush border enzymes form clusters. *Mol. Biol. Cell.* 8:2241–2251.

77. Delacour, D., V. Gouyer, …, G. Huet. 2005. Galectin-4 and sulfatides in apical membrane trafficking in enterocyte-like cells. *J. Cell Biol.* 169:491–501.

78. Ideo, H., A. Seko, and K. Yamashita. 2005. Galectin-4 binds to sulfated glycosphingolipids and carcinoembryonic antigen in patches on the cell surface of human colon adenocarcinoma cells. *J. Biol. Chem.* 280:4730–4737.

79. Mitchell, C. A., K. Ramessar, and B. R. O'Keefe. 2017. Antiviral lectins: selective inhibitors of viral entry. *Antiviral Res.* 142:37–54.

80. Yuan, M., N. C. Wu, …, I. A. Wilson. 2020. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science.* 368:630–633.

81. White, M. R., E. Crouch, …, K. L. Hartshorn. 2001. Increased antiviral and opsonic activity of a highly multimerized collectin chimera. *Biochem. Biophys. Res. Commun.* 286:206–213.

82. Kutzner, T. J., A. Gabba, …, H.-J. Gabius. 2019. How altering the modular architecture affects aspects of lectin activity: case study on human galectin-1. *Glycobiology.* 29:593–607.

83. Ludwig, A.-K., H. Kaltner, …, H.-J. Gabius. 2019. Lectinology 4.0: altering modular (ga)lectin display for functional analysis and biomedical applications. *Biochim. Biophys. Acta, Gen. Subj.* 1863:935–940.

84. García Caballero, G., D. Beckwith, …, H.-J. Gabius. 2020. Influence of protein (human galectin-3) design on aspects of lectin activity. *Histochem. Cell Biol.* 154:135–153.

85. Wang, C., S. Wang, …, J. Wang. 2020. Human intestinal defensin 5 inhibits SARS-CoV-2 invasion by cloaking ACE2. *Gastroenterology.* 159:1145–1147.e4.