

Structural basis for sulfation-dependent self-glycan recognition by the human immune-inhibitory receptor Siglec-8

Johannes M. Pröpster^a, Fan Yang^b, Said Rabbani^b, Beat Ernst^b, Frédéric H.-T. Allain^{a,1}, and Mario Schubert^{a,1,2}

^aInstitute of Molecular Biology and Biophysics, ETH Zurich, 8093 Zurich, Switzerland; and ^bInstitute of Molecular Pharmacy, Pharmazentrum, University of Basel, 4056 Basel, Switzerland

Edited by Carolyn R. Bertozzi, Stanford University, Stanford, CA, and approved May 20, 2016 (received for review February 10, 2016)

Siglec-8 is a human immune-inhibitory receptor that, when engaged by specific self-glycans, triggers eosinophil apoptosis and inhibits mast cell degranulation, providing an endogenous mechanism to down-regulate immune responses of these central inflammatory effector cells. Here we used solution NMR spectroscopy to dissect the fine specificity of Siglec-8 toward different sialylated and sulfated carbohydrate ligands and determined the structure of the Siglec-8 lectin domain in complex with its prime glycan target 6'-sulfo sialyl Lewis^x. A canonical motif for sialic acid recognition, extended by a secondary motif formed by unique loop regions, recognizing 6-O-sulfated galactose dictates tight specificity distinct from other Siglec family members and any other endogenous glycan recognition receptors. Structure-guided mutagenesis revealed key contacts of both interfaces to be equally essential for binding. Our work provides critical structural and mechanistic insights into how Siglec-8 selectively recognizes its glycan target, rationalizes the functional impact of site-specific glycan sulfation in modulating this lectin-glycan interaction, and will enable the rational design of Siglec-8-targeted agonists to treat eosinophil- and mast cell-related allergic and inflammatory diseases, such as asthma.

NMR spectroscopy | protein-carbohydrate recognition | glycan sulfation | immune regulation | glycoimmunology

Discrimination between self and nonself constitutes one of the most challenging tasks for the immune system. Innate immunity relies on an array of germ-line-encoded receptors that detect conserved features on invaders, providing immediate defense against infection, whereas adaptive immunity ensures to keep pace with the vast diversity of constantly evolving pathogenic threats. To prevent self-destructive immune responses, both systems are counterbalanced by inhibitory receptors that down-regulate immune activation upon sensing molecular markers of self. Glycans decorating the surfaces of all living cells are at the forefront of such recognition processes, providing with their unique and complex structures highly discriminative signatures of cellular identity. Terminal sialic acids are an immunological hallmark of vertebrate self-glycans (1–3), and recent evidence suggests glycan sulfation to be another self-marker specific to mammals (4).

Sialic acid-binding Ig-like lectins (Siglecs) represent the largest family of mammalian innate-immune cell surface receptors that recognize self-associated glycans and convert these extracellular recognition events into inhibition of immune cell function (2, 3, 5). To date, 14 human Siglecs have been identified, mostly expressed on various leukocyte populations. Each Siglec displays extracellularly a unique N-terminal lectin domain that binds distinct sialic acid-containing glycan (sialoglycan) ligands, and most members contain conserved immunoreceptor tyrosine-based inhibitory motifs (ITIMs) (6) in their cytoplasmic tails. Ligand-induced activation results in ITIM-phosphorylation by Src family tyrosine kinases and recruitment of SH2-containing phosphatases (SHPs), which interfere with activation pathways of the underlying cell (2). Although Siglecs are generally known to control and balance immune cellular responses by acting as inhibitory receptors, recent

studies intriguingly demonstrated that signaling of specific members is implicated in the regulation of the life span of immune cells, expanding their function as immune modulators to potential mediators of immune homeostasis (7).

Most prominently, Siglec-8, originally identified from a cDNA library derived from a patient suffering from hypereosinophilic syndrome (8, 9), has emerged as a critical negative regulator of inflammatory response during allergic airway inflammation. It is highly and exclusively expressed on human eosinophils and mast cells and weakly on basophils and is conserved only among primates, but it lacks clear orthologs in any other mammalian species. Cross-linking of Siglec-8 on eosinophils with antibodies or a synthetic glycan ligand-coated polymer in vitro induces their rapid apoptosis (10, 11), whereas on mast cells, Siglec-8 ligation results in the inhibition of IgE/FcεRI-mediated inflammatory mediator release, without affecting their survival (12). Preactivation of eosinophils with survival-promoting, proinflammatory cytokines, a key process in allergic inflammation, not only fails to counteract Siglec-8-triggered cell death but instead potently enhances their sensitivity to undergoing apoptosis in response to Siglec-8 ligation

Significance

Siglec-8 downregulates eosinophil- and mast cell-mediated inflammatory responses upon engagement by specific self-glycans. We used solution NMR spectroscopy to determine the structure of the N-terminal lectin domain of human Siglec-8 in complex with its preferred glycan target 6'-sulfo sialyl Lewis^x. Quantitative binding studies with differently sulfated glycans and structure-based mutants demonstrate that Siglec-8 simultaneously recognizes a terminal N-acetylneuraminic acid (sialic acid) and an underlying 6-O-sulfated galactose, yielding a tight and unique specificity. We offer direct structural and mechanistic insights into how the self-glycan code is deciphered by Siglec-8, emphasize the crucial role of glycan sulfation in immunological control of inflammation, and provide a rational framework for designing Siglec-8 agonists to harness its signaling pathway in allergic and inflammatory disorders.

Author contributions: J.M.P., B.E., and M.S. designed research; J.M.P., F.Y., S.R., and M.S. performed research; J.M.P., F.Y., S.R., and M.S. analyzed data; and J.M.P., F.Y., B.E., F.H.-T.A., and M.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates, NMR chemical shifts, and restraints of the herein reported structures have been deposited in the Protein Data Bank (PDB), www.rcsb.org, and the Biological Magnetic Resonance Data Bank (BMRB), www.bmrb.wisc.edu, with the following accession codes: ligand-free Siglec-8 (PDB, [2N7A](https://doi.org/10.1073/pnas.1602214113); BMRB, [25798](https://doi.org/10.1002/bmrb.12579)) and Siglec-8-6'S sLe^x complex (PDB, [2N7B](https://doi.org/10.1073/pnas.1602214113); BMRB, [25799](https://doi.org/10.1002/bmrb.12579)).

¹To whom correspondence may be addressed. Email: mario.schubert@sbg.ac.at or allain@mol.biol.ethz.ch.

²Present address: Department of Molecular Biology, University of Salzburg, 5020 Salzburg, Austria.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1602214113/-DCSupplemental.

(13–15), suggesting a particular role for Siglec-8-mediated immune suppression under inflammatory conditions. Supporting this notion, an *ex vivo* study showed that eosinophils isolated from the bronchoalveolar fluid of allergen-challenged patients have increased susceptibility to Siglec-8-mediated apoptosis (16). Immunohistochemical analysis consistently revealed that endogenous Siglec-8 ligands are markedly up-regulated in inflamed compared with normal human airway tissues (17). Collectively, these findings led to the proposal that Siglec-8 may provide a safeguard mechanism for the immune system to selectively deplete eosinophils from inflamed tissues and simultaneously diminish mast cell inflammatory responses and may thus constitute a basic immunoregulatory pathway to trigger the active resolution of inflammation and return to homeostasis (7, 18) (Fig. 1A). There is increasing awareness that dysregulated Siglec-8 function might be critically involved in the pathobiology of allergic and chronic inflammatory disorders, including asthma, where accumulation and delayed apoptosis of activated eosinophils and mast cells in the airways are among the leading causes of persistent inflammation and tissue damage (19, 20). Indeed, Siglec-8 gene polymorphisms were identified to correlate with increased asthma risk (21), albeit the detailed molecular pathways linking Siglec-8 to disease have yet to be clarified. On account of its potent eosinophil proapoptotic and mast cell-inhibitory activities, combined with its selective expression on these key inflammatory effector cells, Siglec-8 is considered a promising target for novel antiinflammatory, proresolving treatment strategies for asthma and other disease conditions in which inappropriate and/or prolonged inflammatory responses of these cell types contributes to pathology (20, 22–25).

Understanding the structural basis of how Siglec-8 interacts with its carbohydrate ligand is an essential prerequisite for deciphering its molecular mechanism of action, under both physiological and pathological conditions, and will facilitate the rational design of highly specific agents to exploit its antiinflammatory signaling pathway for therapeutic purposes. Although the exact biochemical identity of its natural tissue ligands still remains unknown, glycan microarray analyses have revealed tight specificity for a unique sulfated and sialylated tetrasaccharide glycan epitope, termed 6'-sulfo sialyl Lewis^x (6'S sLe^x) or Neu5Ac α 2-3[6S]Gal β 1-4[Fuc α 1-3]GlcNAc (26, 27), whereas no binding was detected to the closely related 6S sLe^x (Neu5Ac α 2-3Gal β 1-4[Fuc α 1-3][6S]GlcNAc) or the nonsulfated sLe^x, which are known ligands for E-, P-, and L-selectins (28).

Here we present the solution structures of the human Siglec-8 lectin domain in its ligand-free form and in complex with its

preferred glycan epitope 6'S sLe^x. Our structural analysis, combined with quantitative binding studies and site-directed mutagenesis data, provides detailed insight into how Siglec-8 selectively recognizes its glycan target and demonstrates the critical influence that site-specific glycan sulfation can exert in the modulation of lectin-glycan recognition.

Results

Fine Specificity Toward Differently Sulfated sLe^x Glycan Epitopes Assessed by Solution NMR

Previous glycan microarray analyses concordantly identified 6'S sLe^x as the best candidate ligand for Siglec-8 (26, 27), whereas a subsequent study indicated strongest binding to the disulfated 6,6'S sLe^x (29), which contains an additional sulfate group at the C6 position of the GlcNAc moiety. To investigate the importance of individual sulfate modifications on sLe^x for Siglec-8 glycan recognition, we chemically synthesized a series of monosulfated (6S sLe^x and 6'S sLe^x) and disulfated (6,6'S sLe^x) sialyl Lewis^x variants (Figs. 1B and 2A and *SI Appendix*) to compare their binding with nonsulfated sLe^x toward a soluble recombinant ¹⁵N-labeled form of the isolated human Siglec-8 lectin domain (30). Notably, because the lectin domain of Siglec-8 (unlike those of several other Siglec family members) does not contain any potential *N*-linked glycosylation site in its native sequence, the recombinant Siglec-8 closely resembles the natural protein domain. We used 2D NMR spectroscopy titration experiments to monitor changes in Siglec-8 induced upon addition of increasing amounts of ligands. In the 2D ¹H,¹⁵N-HSQC spectrum each NH group is represented by a cross-peak whose spectral position directly reflects its individual local chemical environment. Ligand-binding leads either to gradual changes in cross-peak positions (fast exchange) or to the appearance of new cross-peaks (slow exchange), as ligand concentration is increased. This experimental strategy allowed not only detection and mapping of carbohydrate binding directly on the protein at residue-specific precision but also quantification of monovalent interactions under nearly physiological solution conditions, which is feasible even for weak-affinity interactions (K_d in the μ M to mM range). Binding of each of the four tetrasaccharides induced substantial chemical shift changes in a number of Siglec-8 amide resonances and was characterized by fast exchange kinetics on the NMR time scale (Fig. S1). Perturbed residues clustered to the same discrete sequence regions of Siglec-8, indicating that each of the four sLe^x variants occupies the same binding surface, whereas the magnitudes of the chemical shift changes differed markedly between the four titrations (Fig. 2B and Fig. S1). Examination of the overlaid

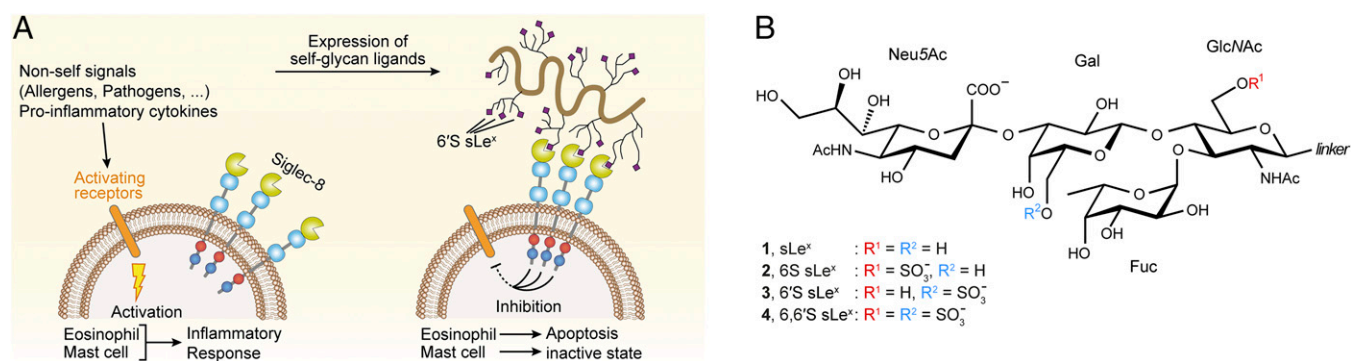


Fig. 1. Proposed mechanism of Siglec-8 function in modulating eosinophil- and mast cell-mediated immune responses and potential glycan ligands for Siglec-8. (A) Ongoing inflammation results in up-regulated expression of specific endogenous Siglec-8 sialoglycan ligands (presumably a secreted, high-molecular weight mucin, carrying multiple 6'S sLe^x glycan epitopes) in airway tissues. Ligand binding to the N-terminal lectin domain of Siglec-8 initiates intracellular signaling cascades, which, through phosphorylation of cytoplasmic ITIM/ITIM-like motifs and recruitment of downstream effector proteins, ultimately lead to apoptosis of eosinophils and inhibition of mast cell degranulation. (B) Chemical structures of glycan ligands used in this study: sLe^x (sialyl Lewis^x: Neu5Ac α 2-3Gal β 1-4[Fuc α 1-3]GlcNAc); 6S sLe^x (6-sulfo sLe^x); 6'S sLe^x (6'-sulfo sLe^x); and 6,6'S sLe^x (6,6'-disulfo sLe^x). The chemical linker for oligosaccharide 1 was β OCH₃ and for 2–4 β O(CH₂)₃NH₃⁺.

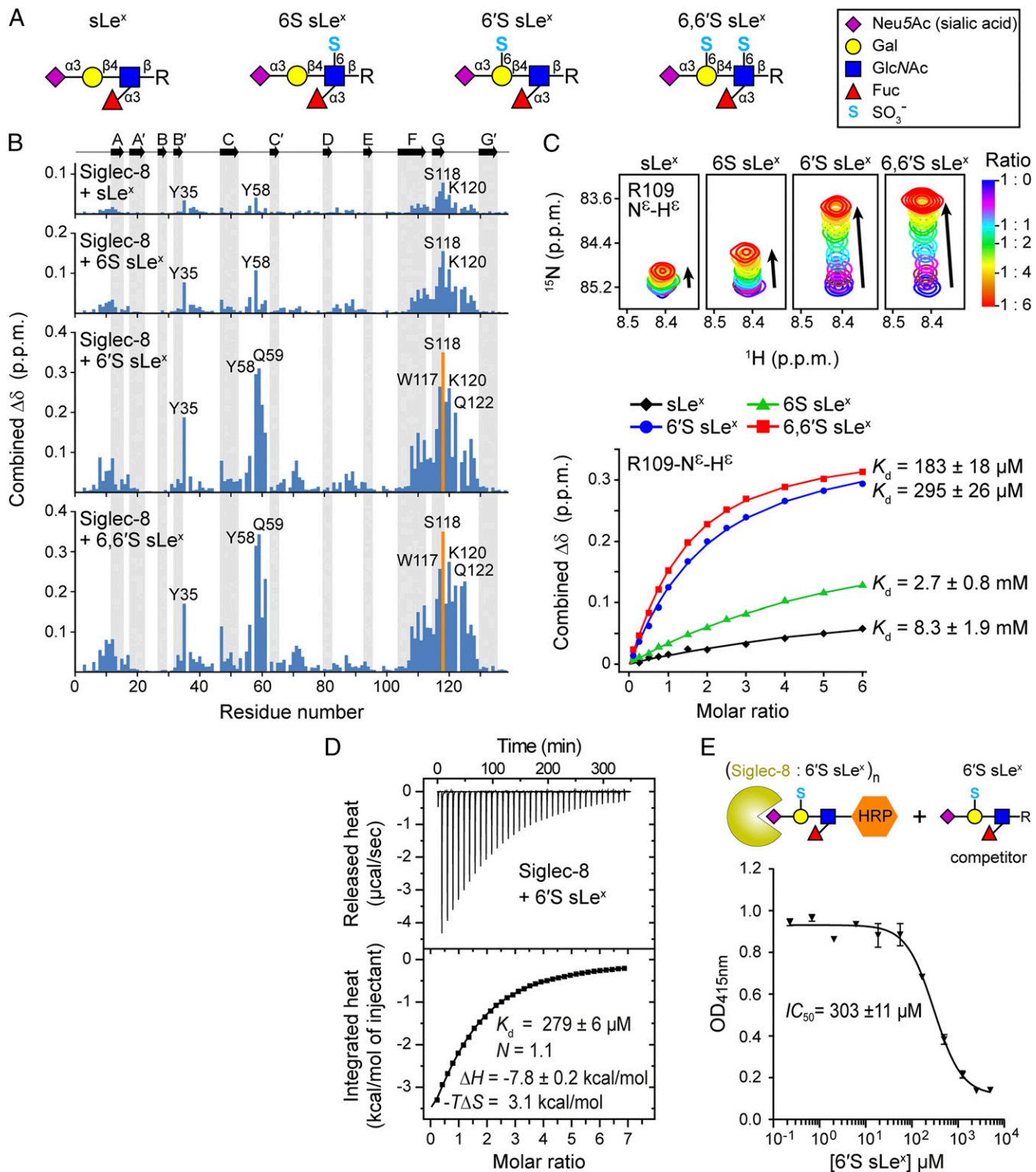


Fig. 2. Binding of Siglec-8 to nonsulfated and differently sulfated sLe^x variants. (A) Schematic representations of the glycan structures shown in Fig. 1B. Monosaccharide symbols are indicated on the right. (B) Combined ¹H-¹⁵N chemical shift changes ($\Delta\delta$) of the ¹⁵N-labeled Siglec-8 lectin domain observed upon NMR titration with the different sLe^x variants (to sixfold molar excess), plotted versus residue number. Secondary-structure elements derived from subsequent structural analysis are shown at the top. Orange bars represent residues whose ¹H-¹⁵N cross-peak intensities progressively decreased during NMR titration, indicating intermediate exchange on the NMR time scale. (C) Overlaid sections of ¹H-¹⁵N-HSQC spectra of Siglec-8 showing the representative Arg109 side chain N^εH^ε cross-peak, acquired before and after stepwise addition of indicated carbohydrate ligands (*Top*), and the corresponding NMR-binding isotherms (*Bottom*). Reported K_d values are the mean \pm SD from separate fitting of binding curves of 9–21 individual residues. The full set of NMR titration curves used for K_d determination is shown in Fig. S2. (D) ITC data obtained by injecting 6'S sLe^x into a solution of Siglec-8. N , stoichiometry (carbohydrate/protein); ΔH , change in enthalpy; $-T\Delta S$, change in entropy. (E) Polyacrylamide-based competitive binding assay for determination of IC_{50} for Siglec-8 binding to 6'S sLe^x. Immobilized lectin domains were incubated with the carbohydrate ligand together with a streptavidin-horseradish peroxidase (HRP) conjugated polyacrylamide glycopolymer whose concentration was assessed by colorimetric detection of HRP activity.

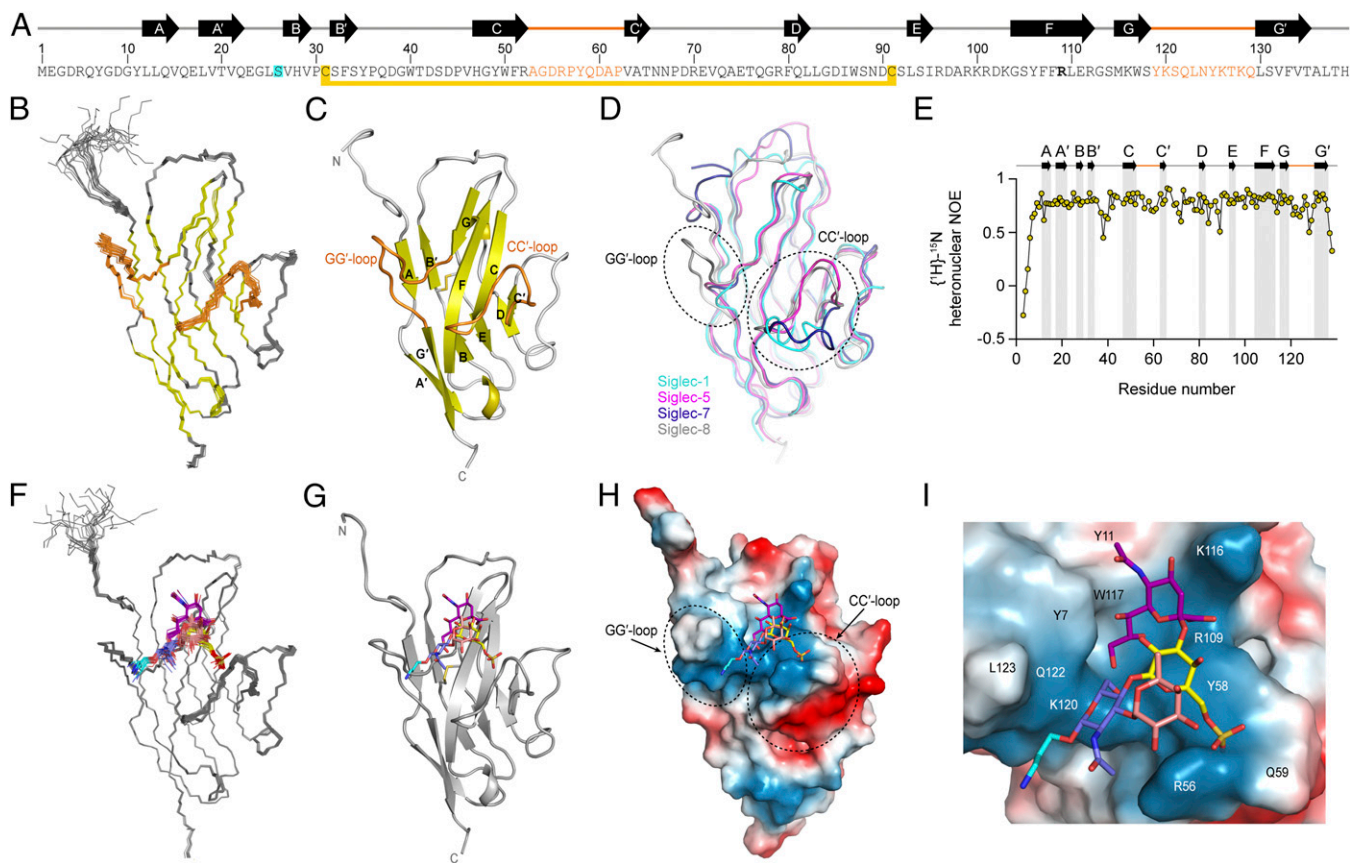


Fig. 3. Solution structures of the human Siglec-8 lectin domain unliganded and in complex with 6'S sLe^x. (A) Primary sequence of the human Siglec-8 lectin domain. Sequence numbers correspond to the mature protein. Secondary-structure elements are shown at the top. The intrasheet disulfide bond between Cys31 and Cys91 is indicated by a yellow bar, the C265 mutation (30) and the essential arginine (Arg109) is in bold. The Siglec-8-unique CC' and GG' loops are highlighted in orange. (B) Ensemble of the 20 lowest-energy structures of Siglec-8, showing backbone atoms (N, C α , and C') only. (C) Cartoon representation of the lowest-energy structure of Siglec-8, with the disulfide bond depicted as sticks. (D) Superposition of the Siglec-8 solution structure with crystal structures of Siglec-1 (PDB entry: 1QFP), Siglec-5 (2ZG2), and Siglec-7 (1O7V), colored as indicated, with rms deviations ranging from 0.9 to 1.5 Å over 82–112 aligned C α positions. (E) The $\{^1\text{H}\}$ - ^{15}N heteronuclear NOE values for Siglec-8 backbone amides plotted versus the residue number, measured at 750 MHz and 293 K. (F) Ensemble of the 20 lowest-energy structures of Siglec-8 in complex with 6'S sLe^x. (G) Cartoon representation of the lowest-energy structure of the Siglec-8–6'S sLe^x complex. (H) Molecular surface of the Siglec-8 lectin domain with bound 6'S sLe^x, colored by electrostatic potential: red, negative; blue, positive. (I) Close-up view of 6'S sLe^x in the positively charged glycan-binding pocket. Selected residues at the interface are indicated. The carbohydrate is represented as sticks and colored as follows: Neu5Ac (purple), Gal6S (yellow), GlcNAc (violet), Fuc (light red), and the chemical linker (cyan).

2D ^1H , ^{15}N -HSQC titration spectra revealed that most of the perturbed cross-peaks followed similar trajectories, consistent with a similar ligand orientation. However, dependent on the presence or absence of the sulfate at the C6 position of the Gal, the cross-peaks of Tyr58 and Gln59 moved into different directions (Fig. S1B), suggesting their proximity to the sulfate recognition site.

NMR titration curves of significantly perturbed and well-resolved resonances were used to extract equilibrium dissociation constants (K_d) (Fig. 2C and Fig. S2), revealing a distinct hierarchy of binding strengths: sLe^x binds with the weakest affinity (K_d : 8.3 ± 1.9 mM), and the additional presence of a sulfate group at the C6 position of the GlcNAc moiety of 6S sLe^x resulted in only a modest gain in affinity (K_d : 2.7 ± 0.8 mM; threefold compared with sLe^x). In contrast, the presence of a single sulfate group at the C6 position of the Gal moiety of 6'S sLe^x—the presumed Siglec-8 ligand (26, 27)—produced a dramatic increase in affinity (K_d : 295 ± 26 μM ; 28-fold compared with sLe^x). A similar, however, slightly higher, affinity was determined for the disulfated 6,6'S sLe^x (K_d : 185 ± 18 μM ; 1.6-fold compared with 6'S sLe^x).

To quantify further Siglec-8 binding to 6'S sLe^x and to elucidate the thermodynamics of this interaction, we used isothermal titration calorimetry (ITC) (Fig. 2D) and a competitive binding

assay (31) (Fig. 2E). The obtained K_d of 279 μM and IC_{50} of 303 μM are both in excellent agreement with the results from NMR (K_d : 295 μM). With a ΔH of -7.8 kcal/mol the interaction is driven by favorable enthalpic forces that overcome an entropic cost of $-\text{T}\Delta S$ of ~ 3 kcal/mol.

Collectively, these data evidence a significantly stronger binding of Siglec-8 toward Gal-6-sulfated sLe^x glycan epitopes (6'S sLe^x and 6,6'S sLe^x), compared with GlcNAc-6-sulfated (6S sLe^x) and nonsulfated sLe^x. The largest affinity enhancement resulted clearly from the addition of the sulfate group at the C6 position of the Gal moiety of 6'S sLe^x (28-fold compared with sLe^x). Based on these observations, we concluded that 6'S sLe^x contains the minimum epitope specifically recognized by Siglec-8.

Structure of the Human Siglec-8 Lectin Domain. To gain insights into the structural determinants underlying this narrow target specificity, we determined the 3D structure of the human Siglec-8 lectin domain by solution NMR spectroscopy (Fig. 3A–C, Fig. S3B, and Table S1). The structure revealed a canonical Siglec lectin domain fold, which is a V-set Ig-like β -sandwich of two antiparallel β -sheets formed by β -strands ABED and C'CFG. Characteristic features of Siglec lectin domains were observed: the conserved intrasheet disulfide bond between adjacent β -strands

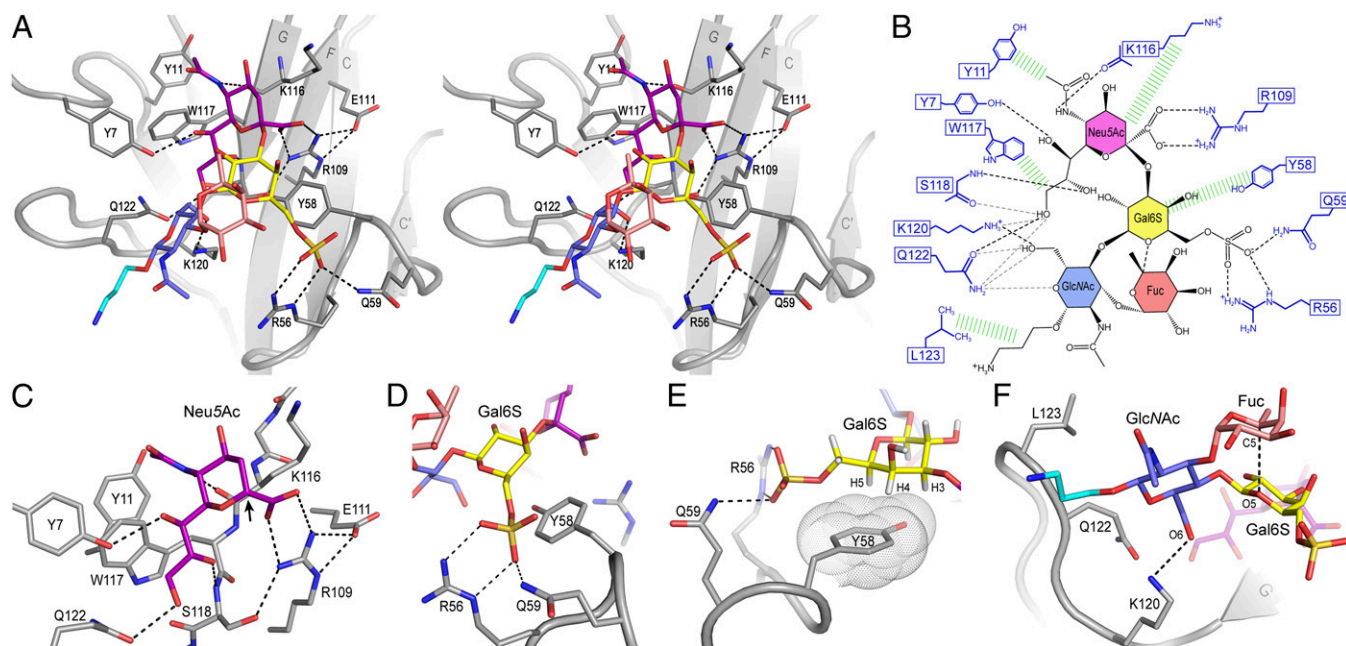


Fig. 4. Structural basis for 6'S sLe^x recognition by human Siglec-8 illustrated by a representative structure (lowest energy) of the NMR ensemble. (A) Stereoview of the Siglec-8–6'S sLe^x interface. The carbohydrate and key interacting amino acids are shown as sticks. Hydrogen bonds are represented by dashed lines. (B) Schematic illustration of the Siglec-8–6'S sLe^x interaction network. Black dashed lines indicate hydrogen bonds in the depicted structure; gray dashed lines indicate hydrogen bonds abundantly observed in other structures of the ensemble. Hydrophobic contacts are shown in green. (C) Close-up view of the recognition of the terminal Neu5Ac. The arrow points to the anomeric carbon atom (C2) of Neu5Ac to which the underlying carbohydrate moieties (omitted, for visual clarity) are linked. (D) Recognition of the Gal6S moiety involving a salt bridge and hydrogen bonds between the sulfate and the Arg56 and Gln59 side chains and (E) CH/ π -stacking interactions between the apolar B-face of Gal6S and the aromatic ring of Tyr58, for which the van der Waals surface is displayed as dotted spheres. (F) Contacts between Siglec-8 and the GlcNAc moiety, also depicted is the intracarbohydrate hydrogen bond between the Fuc C5-H5 and the O5 pyranose ring oxygen of Gal6S.

B and E; the strictly conserved essential arginine (Arg109) on β -strand F, known to provide a key salt bridge interaction for sialic acid recognition; and the splitting of the G-strand into two shorter β -strands (G and G'). Structural alignment with available crystal structures of Siglec-1 (32), Siglec-5 (33), and Siglec-7 (34) revealed, besides a nearly invariant core, large conformational differences in the N-terminal segments (before β -strand A) and the B/C intersheet loops, as well as the CC' and GG' interstrand loops (Fig. 3D). These four regions show the highest sequence diversity between the lectin domains of the different human Siglec family members, while being highly conserved among primate Siglec-8 orthologs (sequence alignment shown in Fig. S3A). Remarkably, the GG' loop of Siglec-8, consisting of 11 residues, is substantially extended compared with the typically 5 residue-spanning GG' loops of most Siglecs. Notably, in Siglec-7 and Siglec-9, the CC' and GG' loops had been found to contribute interactions to glycan moieties underlying a terminal sialic acid (35, 36) and are thus the presumed key determinants for the distinct fine carbohydrate specificities of individual Siglecs. Consistently, both loops displayed large chemical shift perturbations upon ligand binding (Fig. 2B). Backbone $\{^1\text{H}\}$ - ^{15}N -heteronuclear nuclear Overhauser effect (hetNOE) measurements (Fig. 3E) indicated that both loops are well ordered and do not exhibit conformational dynamics on the here probed picosecond-to-nanosecond time scale, which may in part originate from a stabilizing network of hydrogen bonds and salt bridges (Fig. S3 C–F).

Structure of the Human Siglec-8 Lectin Domain in Complex with 6'S sLe^x. To understand the structural basis for glycan recognition and discrimination by Siglec-8, we next determined the solution structure of the human Siglec-8 lectin domain in complex with 6'S sLe^x by NMR spectroscopy. Despite the weak binding affinity (K_d in the high micromolar range), 113 intermolecular NOEs—a large number for a tetrasaccharide ligand—could be unambiguously

identified and converted into distance restraints (Fig. S4 A and B). Based on these, plus 104 intracarbohydrate and 3,938 intra-protein NOE-derived distance restraints, we determined a precise structural ensemble of the complex (Fig. 3F and Table S1). Overall, the structure of Siglec-8 in the complex is virtually identical to that of the unliganded Siglec-8 ($C\alpha$ rms deviation of 0.8 ± 0.1 Å for residues 7–135), indicating that carbohydrate recognition is mediated by a largely preformed binding site, involving only minor side chain rearrangements. The carbohydrate lies embedded in a highly positively charged cleft formed by β -strands C, F, and G encompassed by the Siglec-8-unique CC' and GG' loops (Fig. 3 G–I). About 678 Å² or 46% of its solvent-accessible surface is buried upon binding to Siglec-8. The entire tetrasaccharide is well-ordered, showing a single major conformation throughout the ensemble with tightly clustered glycosidic torsion angles for the core Le^x trisaccharide: for the Gal β 1–4GlcNAc linkage, $47^\circ < \varphi < 54^\circ$ and $12^\circ < \psi < 19^\circ$; for Fuc α 1–3GlcNAc linkage, $47^\circ < \varphi < 51^\circ$ and $16^\circ < \psi < 20^\circ$, and there was a slightly wider spread of angles for the terminal Neu5Ac α 2–3Gal linkage, $-94^\circ < \varphi < -66^\circ$ and $18^\circ < \psi < 25^\circ$ (see *Materials and Methods* for angle definitions).

Intermolecular Contacts at the Complex Interface. Closer inspection of the binding interface (Fig. 4) revealed that specificity is achieved by two major hot spots: a primary motif for recognition of the terminal Neu5Ac, extended by a secondary motif recognizing the subterminal galactose-6-sulfate (Gal6S). A summary of observed intermolecular contacts in the Siglec-8–6'S sLe^x complex is given in Table 1. The Neu5Ac is located at the edge of the G-strand, such that its carboxyl group makes a salt bridge with the guanidinium group of the strictly conserved essential Arg109 (on the adjacent F-strand), whose orientation itself is stabilized by salt bridge/hydrogen bond interactions with the side

Table 1. Contacts between Siglec-8 and 6'S sLe^x observed in the NMR ensemble

6'S sLe ^x	Siglec-8	Type*	Occ. [†]	Supporting NMR data
Neu5Ac O1A/O1B	Arg109 N ^{η1}	SB	16/20	Large N ^ε H ^ε chemical shift changes [‡]
Neu5Ac O1A/O1B	Arg109 N ^{η2}	SB	20/20	Large N ^ε H ^ε chemical shift changes [‡]
Neu5Ac H3eq,H4	Lys116	HC	20/20	Intermolecular NOEs [§]
Neu5Ac N5	Lys116 O	HB	20/20	Large C' chemical shift change [§]
Neu5Ac CH ₃	Tyr11	HC	15/20	Intermolecular NOEs [§]
Neu5Ac H6	Trp117 H ^α	HC	20/20	Intermolecular NOEs [§]
Neu5Ac O7	Tyr7 O ^γ	HB	18/20	Intermolecular NOEs [§]
Neu5Ac O8	Ser118 N	HB	20/20	NH chemical shift disappeared upon binding [‡]
	Ser118 O ^γ	HB	5/20	
Neu5Ac H7,H91/H92	Trp117	CH/π	20/20	Intermolecular NOEs [§]
Neu5Ac O9	Ser118 O	HB	17/20	Largest C' chemical shift change [§]
	Lys120 N ^ζ	HB	9/20	
	Gln122 O ^{ε1}	HB	7/20	
	Gln122 N ^{ε2}	HB	2/20	Large N ^{ε2} H ^{ε21/22} chemical shift changes [‡]
Gal6S SO ₃ ⁻	Arg56 N ^{η1}	SB	6/20	
	Arg56 N ^{η2}	SB	6/20	
	Arg56 N ^ε	SB	6/20	N ^ε H ^ε chemical shift changes [§]
	Gln59 N ^{ε2}	HB	16/20	Large N ^{ε2} H ^{ε21/22} chemical shift changes [‡]
Gal6S H3,H4,H5	Tyr58	CH/π	20/20	Intermolecular NOEs [§]
GlcNAc O5	Gln122 N ^{ε2}	HB	4/20	Large N ^{ε2} H ^{ε21/22} chemical shift changes [‡]
GlcNAc O6	Lys120 N ^ζ	HB	13/20	Intermolecular NOEs [§]
	Gln122 N ^{ε2}	HB	5/20	Large N ^{ε2} H ^{ε21/22} chemical shift changes [‡]

*Hydrogen bond (HB) and salt bridge (SB) interactions are listed for heavy atom distances of 3.2 Å or less. HC, hydrophobic contacts; CH/π, CH/π-stacking interactions.

[†]Occ., occurrence in the NMR ensemble of the 20 lowest-energy structures of the Siglec-8–6'S sLe^x complex. Note that all intermolecular restraints used for structure calculation were derived from intermolecular NOE data, and none of the listed interactions was used as artificial input.

[‡]Shown in Fig. S1.

[§]Shown in Fig. S4.

chains of Glu111 (F-strand) and Ser118 (G-strand). The Neu5Ac pyranose ring packs tightly with its hydrophobic A-face (*Materials and Methods*) against the aliphatic portion of Lys116, which protrudes from the upper part of the G-strand, likely providing an upper boundary for the Neu5Ac recognition site. The *N*-acetyl amide and the O8 and O9 hydroxyl groups of the glycerol chain form a network of hydrogen bonds with backbone atoms of the G-strand residues Lys116 and Ser118. In addition, the *N*-acetyl methyl group, together with the H7 and the H91/H92-methylene atoms, makes intimate hydrophobic contacts with the surrounding aromatic rings of Tyr11 and Trp117 (Fig. 4 *A–C*). Notably, the interactions engaging Neu5Ac and its orientation on the Siglec-8 surface are nearly identical to those seen in structures of other Siglec family members (32, 33, 36), as was expected from sequence comparisons. However, an additional hydrogen bond is formed between the O7 hydroxyl group of Neu5Ac and the side chain hydroxyl group of Tyr7 located in the unique N-terminal extension of Siglec-8 (Fig. 4 *A–C* and Fig. S34).

The most striking and distinguishing feature of this complex is the recognition of the Gal6S, which is mediated exclusively by the side chains of three residues (Arg56, Tyr58, and Gln59) on the Siglec-8–unique CC' loop (Fig. 4 *D* and *E*). The aromatic ring of Tyr58 lies flat and exposed on top of the CC' loop, providing a platform for CH/π-stacking interactions with the apolar B-face of the galactopyranose ring, particularly with the plane formed by H3, H4, and H5. The exocyclic moiety of Gal6S approaches the edge of the CC' loop and positions the negatively charged sulfate group in the center of the flanking Arg56 and Gln59 side chains, thereby forming a salt bridge with the guanidinium group of Arg56 and/or a hydrogen bond with the NH₂ of Gln59. Remarkably, in 9 of the 20 refined structures, both side chains embrace the sulfate from opposite sides in a clamp-like configuration, thus contributing simultaneously to its recognition. These interactions are confirmed by the chemical shift

changes of the Arg56 and Gln59 side chain amides observed only upon titration with sLe^x variants containing a Gal-6-sulfate modification (Figs. S1*B* and S4*D*).

Only few intermolecular contacts were found with the GlcNAc: its apolar B-face and *N*-acetyl group are exposed to solvent, whereas its exocyclic hydroxymethylene group (directed toward the Neu5Ac glycerol chain) is interacting with the long and flexible side chains of Lys120 and Gln122 on the GG' loop. In most structures of the ensemble, the GlcNAc O6 is hydrogen bonded by the positively charged NH₃⁺ of Lys120 (Fig. 4*F*), which otherwise donates a hydrogen bond to the nearby Neu5Ac O9. In few structures, the GlcNAc O6 is hydrogen bonded by the NH₂ of Gln122, which in turn forms in some structures a hydrogen bond with the O9 of Neu5Ac (Fig. 4 *B* and *C* and Table 1).

The 3-amino-propyl linker, attached to the reducing end of the tetrasaccharide, exits the binding site toward the tip of the GG' loop (close to Gln122 and Leu123; Fig. 4*F*) and appears disordered. No intermolecular contacts were found to Fuc. Its B-face is fully solvent-exposed, whereas its A-face stacks onto the A-face of the Gal6S, stabilizing the Le^x conformation by donating an intracarbohydrate C–H...O hydrogen bond from its C5–H5 to the O5 ring oxygen of the Gal6S (Fig. 4*F*), as has been found previously in structures of Le^x (37) and sLe^x (38) oligosaccharides. The presence of this hydrogen bond is indicated by the distance between Fuc-H5 and Gal-O5 of 2.2–2.3 Å within the ensemble, which is significantly shorter than the sum of the corresponding van der Waals radii (2.6 Å), and further supported by the characteristic NMR chemical shift of Fuc-H5 (~4.8 ppm) (37, 38).

Mutations of Key Interface Residues Impair Siglec-8 Binding to 6'S sLe^x. To dissect quantitatively the contributions of individual amino acid side chains to carbohydrate recognition, we introduced several single alanine substitutions at the interface and measured

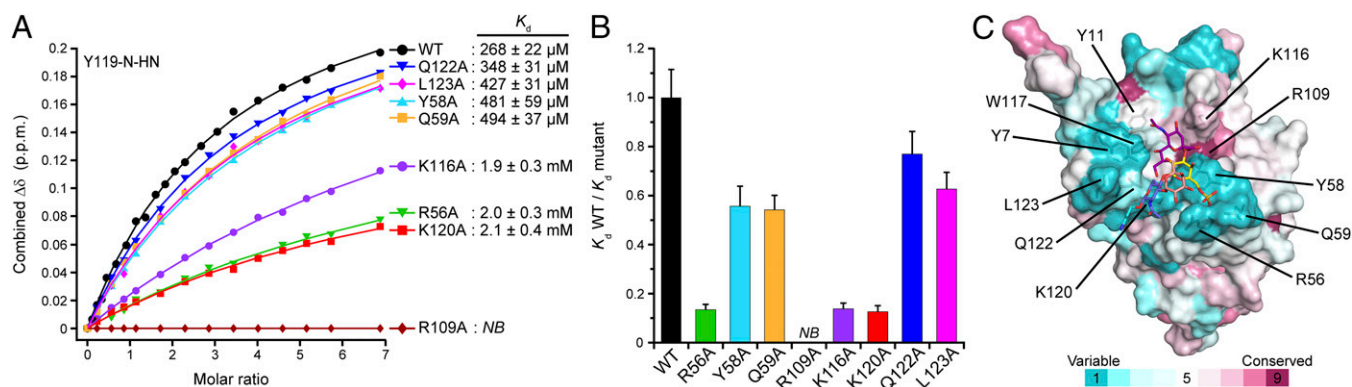


Fig. 5. Binding affinities of Siglec-8 mutants for 6'S sLe^x determined by NMR-titration, and conservation surface mapping. (A) Combined ¹H-¹⁵N chemical shift changes ($\Delta\delta$) of the representative Tyr119 amide resonance detected on ¹⁵N-labeled Siglec-8 wild-type (WT) and various alanine-substituted mutants plotted as a function of the molar ratio (carbohydrate/protein). Dissociation constants (K_d) are presented as mean \pm SD from separate fitting of the NMR-binding isotherms of 9–12 individual residues, as shown in Fig. S5. NB, no binding detected. (B) K_d values for binding of Siglec-8 wild-type and mutants to 6'S sLe^x, normalized with respect to that of Siglec-8 wild-type (set to 1.0). Error bars represent propagation of errors. (C) Conservation analysis among the human Siglec family members mapped onto the molecular surface of Siglec-8 in complex with 6'S sLe^x. Amino acid residues are color-coded according to sequence conservation scores calculated by ConSurf (39). Side chains of selected residues within the binding site are shown as sticks and labeled.

their effects on binding to 6'S sLe^x by using NMR titration experiments (Fig. 5 A and B and Fig. S5). None of these mutations affected the overall structure of Siglec-8, as indicated by very similar 2D ¹H, ¹⁵N-HSQC fingerprint spectra compared with the wild-type protein. As expected, substitution of the strictly conserved essential Arg109, which eliminated the salt bridge with the Neu5Ac carboxyl group, completely abrogated binding, confirming its indispensable role in carbohydrate recognition. Substitution of Lys116, which participates in hydrophobic contacts with the Neu5Ac pyranose ring, strongly impaired binding (sevenfold loss in affinity). A similar drastic drop in affinity (eightfold) resulted from substitution of Arg56 on the CC' loop, whereas substitutions of the adjacent Tyr58 or Gln59 only modestly affected binding (twofold affinity decrease for each mutant). Hence, a major contribution to the binding energy must originate from the salt bridge formation between the Arg56 guanidinium group and the sulfate. Substitution of the GG' loop residue Lys120, which forms hydrogen bonds with the GlcNAc O6 and/or the Neu5Ac O9, had a large effect on binding (eightfold drop in affinity). In contrast, substitutions of its neighboring GG' loop residues Gln122 or Leu123 had only minor effects on binding (less than twofold affinity decrease for each mutant).

Overall, substitutions of the positively charged residues at the binding site (Arg56, Arg109, Lys116, and Lys120) caused the largest reductions in binding affinity. Besides providing specific hydrogen-bonding contacts, these tightly clustered positive charges (Fig. 3 H and I) may contribute to a substantially enhanced association rate (k_{on}) owing to attractive Coulomb interactions (40) with the doubly negatively charged carbohydrate ligand. Electrostatic enhancement of complex formation has been described for protein–RNA interactions, where binding can be several orders of magnitude faster than in a purely diffusion-limited process (41), and is particularly conceivable for the present interaction characterized by fast-exchanging free and bound states. It is plausible that this effect is further amplified when clustered receptors on cell surfaces avidly interact with multivalently presented glycan epitopes on large polysaccharide or mucin-type glycoprotein ligands, leading to markedly higher *in vivo* binding affinities, compared with the here investigated monovalent lectin–glycan interaction.

Discussion

Here we have elucidated the molecular basis of how human Siglec-8 selectively recognizes its principal carbohydrate target

6'S sLe^x. Our comparative analysis, combined with previous functional studies (26, 27), provides clear evidence that Siglec-8 preferentially recognizes sLe^x epitopes, which display a sulfate group at the C6 of the Gal moiety, and demonstrates that both the presence and position of this modification are critical determinants of Siglec-8 specificity. Quantification revealed the sulfate group to account for a 28-fold greater affinity for binding to 6'S sLe^x ($K_d \approx 300 \mu\text{M}$), compared with nonsulfated sLe^x ($K_d \approx 8.3 \text{ mM}$). In contrast, a sulfate modification at the C6 of the GlcNAc moiety was found to make only a minor contribution to binding affinity (Fig. 2 B and C). The solution structure of the Siglec-8 lectin domain in complex with 6'S sLe^x explains these differences in affinity, revealing that tight specificity results from the unique combination of a primary recognition motif for terminal Neu5Ac, which is largely canonical among Siglecs, and a secondary motif for recognition of the underlying Gal6S, which is exclusive to Siglec-8. Conservation analysis of the Siglec-8 binding interface across the human Siglec family (Fig. 5C) shows that only residues engaging Neu5Ac have elevated levels of conservation, whereas all interactions to the underlying glycan moieties are solely mediated through nonconserved side chains. High selectivity for Gal6S is conferred by the cooperative action of three residues on the Siglec-8–unique CC' loop (Arg56, Tyr58, and Gln59), and mutational dissection revealed this interaction to be dominated by a salt bridge between the Arg56 guanidinium and the sulfate group of Gal6S (Figs. 4 D and E and 5). Although GlcNAc makes direct protein contacts that contribute to affinity, the low number and high variability of interactions observed in the NMR ensemble (Table 1) imply only a minor role for this residue in specificity. Notably, the O6 of GlcNAc is anchored in the interior of the binding groove by hydrogen bonds with the side chains of Lys120 and/or Gln122 (Fig. 4 F and Table 1), thus suggesting how the GlcNAc-6-sulfate group in 6S sLe^x and 6,6'S sLe^x ligands may be contacted by Siglec-8. However, molecular modeling of the Siglec-8–6,6'S sLe^x complex (Fig. S6) indicates that binding of a bulkier sulfate would interfere with the architecture of the binding pocket, implying that the small net gain in affinity produced by a GlcNAc-6-sulfate modification may arise from additional favorable electrostatic and hydrogen bonding interactions, largely counteracted by less favorable steric complementarity. The lack of contacts to Fuc, which projects away from the Siglec-8 surface, is consistent with a recent glycan microarray analysis (42) where a Siglec-8-IgFc fusion protein was found to bind to 6'S sLacNAc (6'-sulfo 3'-sialyl N-acetylglucosamine;

corresponding to 6'S sLe^x, minus α 1,3-linked Fuc), in addition to 6'S sLe^x, suggesting this monosaccharide to be dispensable for Siglec-8 recognition. However, Fuc stabilizes the core Le^x trisaccharide in a binding-competent conformation by forming an intracarbohydrate C–H...O hydrogen bond to Gal6S (Fig. 4F), together with hydrophobic interactions (37, 38, 43), and may thus indirectly contribute to affinity by decreasing the entropic penalty of fixing multiple conformational degrees of freedom upon binding.

Disruption of the salt bridge to the Neu5Ac carboxylate by mutation of the essential Arg109 leads to a complete loss of binding, and similarly, without Gal-6-sulfation the binding affinity is decreased to values that are likely well below the threshold needed for Siglec-8 signaling activation. Together this demonstrates that simultaneous recognition of Neu5Ac and the underlying Gal6S is absolutely required for effective Siglec-8 engagement. Considering that no other human Siglec, nor any other known endogenous receptor, shares this carbohydrate specificity, this unique binding mode may ensure a maximum control of the Siglec-8 antiinflammatory signaling pathway, while safeguarding against cross-activation by other sialylated glycans omnipresent on vertebrate cell surfaces or by hijacking pathogens decorated with sialic acids to evade host immune surveillance.

Of note, murine Siglec-F has been considered, albeit controversially, to be a functional paralog of Siglec-8, owing to a similar (although less restricted) expression pattern, glycan specificity, and ligation-induced proapoptotic effects, and was therefore proposed as an *in vivo* model for exploring Siglec-8 function in mice (3, 27). In light of our results, sequence analysis reveals that key residues of Siglec-8 responsible for the sulfate recognition are maintained only across its primate orthologs but are absent in Siglec-F (Fig. S34), suggesting either that Siglec-F recognizes 6'S sLe^x by a distinct mode or, in support of recent findings (44, 45), that Gal C6 sulfation is not required for Siglec-F ligands.

Our findings also established that a single sulfate modification at the Gal C6 renders the poor-affinity sLe^x glycan epitope into the preferred ligand for Siglec-8 and thus into a potential anti-inflammatory signal. By contrast, nonsulfated sLe^x and variants carrying a sulfate group at GlcNAc C6 are known ligands for the selectin family of cell adhesion receptors (E-, P-, and L-selectin) (28) that mediate the recruitment of leukocytes (including eosinophils and mast cells) to inflammatory sites. Given that selectins play key roles in the initiation of eosinophil inflammation, whereas Siglec-8 appears to be implicated in resolving eosinophil inflammatory responses, it is tempting to speculate that site-specific sulfation of the sLe^x epitope acts as one of the molecular switches that dictate whether eosinophil inflammation is initiated or terminated.

Our results provide unprecedented atomic-level insight on how site-specific glycan sulfation can potentially modulate glycan affinity and specificity to prime its selective recognition by its cognate receptor. This type of glycan modification adds, in addition to the nonlinearity and multitude of linkage sites per monosaccharide, yet another layer of complexity to the coding capacity of glycan structures and hence to the processes controlled by their recognition. Future studies identifying and characterizing the involved glycan-editing sulfotransferases and sulfatases will be crucial to understand further the intricate role of glycan sulfation in immune regulation and inflammation, as well as for developing new strategies for therapeutic intervention.

Collectively, we report here the complete molecular and structural description of how a Siglec receptor specifically recognizes its entire glycan ligand epitope, revealing the key determinants of glycan specificity and discrimination. Whereas the conserved sialic acid recognition motif governs general specificity, local sequence diversity in the variable loop regions defines the fine specificity and translates into unique binding modes, enabling individual Siglec family members to participate

in distinct and discrete immune-regulatory functions. This work will help elucidate the precise role and mechanism of Siglec-8 in immune cell homeostasis and inflammation resolution and provides a structural template for the rational design of Siglec-8 agonists to harness its potent signaling capacity in novel anti-inflammatory therapies for asthma.

Materials and Methods

Carbohydrates. The 6-sulfo sialyl Lewis^x, 6'-sulfo sialyl Lewis^x, and 6,6'-disulfo sialyl Lewis^x were chemically synthesized as described in *SI Appendix*. Methyl sialyl Lewis^x was purchased from Carbosynth. Identity and purity of all carbohydrates was assessed by 2D NMR spectroscopy.

Protein Expression and Purification. Proteins were expressed recombinantly and purified as described elsewhere (30). In brief, human Siglec-8 lectin domain (Met1-His139, containing a C26S point mutation to facilitate soluble expression; herein referred to as wild type), C-terminally fused to a thrombin cleavage-site and a His₆-tag, was natively expressed in the oxidative cytoplasm of *Escherichia coli* Rosetta-gami B (*trxB⁻lgor⁻*) (Novagen), using LB medium or M9 minimal medium supplemented with 1 g/L ¹⁵NH₄Cl and 4 g/L D-glucose (¹³C-labeled for ¹³C/¹⁵N-labeled proteins). Proteins were purified by Ni-NTA affinity chromatography, followed by thrombin cleavage for His₆-tag removal and size exclusion chromatography (Superdex 75; GE Healthcare). Protein purity and uniform presence of the intradomain disulfide bond were confirmed by SDS/PAGE (reducing vs. nonreducing) and ESI-TOF mass spectrometry, and the oxidized state of the disulfide bond was additionally confirmed by NMR spectroscopy (details are shown in Fig. S7). Mutants were constructed by PCR-mediated site-directed mutagenesis, confirmed by DNA sequencing and expressed and purified as the wild-type protein. The proper fold of all recombinant proteins was confirmed by the observation of well-dispersed 2D ¹H,¹⁵N-HSQC NMR spectra.

NMR Spectroscopy. NMR spectra were acquired on Bruker AVIII 500-, 600-, 700-, 750-, and 900-MHz spectrometers (all equipped with a cryogenic probe, except for AVIII 750 MHz) at 293 K, unless mentioned otherwise. Samples were measured in 20 mM potassium phosphate, 40 mM NaCl at pH 7.0 (free protein) or pH 7.4 (complex), and at protein concentrations between 0.3–1.2 mM, containing either 5% or 100% (vol/vol) D₂O. Protein-carbohydrate complexes were prepared by titrating carbohydrate solution of typically 7–17 mM into a 0.7–1.2 mM protein solution until a molar stoichiometry of 1:1, 1:1.2, or 1:2 was reached. Sequence-specific assignment of protein backbone and side-chain resonances was achieved through 2D ¹H,¹⁵N-HSQC, 2D ¹H,¹³C-HSQC, 3D HNCA, 3D HNCACB, 3D CBCA(CO)NH, 3D HNCO, 3D HN(CO)CA, 3D (H)CCH-TOCSY, 3D ¹⁵N-edited NOESY-HSQC, 2D ¹H,¹H-NOESY, and two 3D ¹³C-edited NOESY-HSQC, optimized for the observation of protons attached to aliphatic carbons and to aromatic carbons, respectively. Stereospecific assignments of Val and Leu methyl groups were obtained from 2D ¹H,¹³C-HSQC spectra by using a 10% ¹³C-labeled sample as described (46). The 2D ¹H,¹H-TOCSY and 2D ¹H,¹³C-HSQC spectra in D₂O aided the assignment of aromatic side chain resonances. For the 2D ¹H,¹H-TOCSY experiments, mixing times of 13 and 60 ms were used, and for the 3D (H)CCH-TOCSY spectrum a mixing time of 21.7 ms was used. All NOESY experiments were recorded using a mixing time of 120 ms, unless otherwise stated. Resonance assignment of the carbohydrate was achieved by using natural abundance 2D ¹H,¹³C-HSQC, 2D ¹H,¹³C-HMBC, 2D ¹H,¹³C-HMQC-COSY, and 2D ¹H,¹H-TOCSY spectra. Resonance assignments of the carbohydrate in complex with the Siglec-8 lectin domain were performed using natural abundance 2D ¹H,¹³C-HSQC, 2D ¹³C/¹⁵N F₂-filtered ¹H,¹H-TOCSY, and 2D ¹³C/¹⁵N F₁-filtered F₂-filtered NOESY spectra. Intracarbohydrate NOEs were assigned from 2D ¹³C/¹⁵N F₁-filtered F₂-filtered NOESY experiments. Protein-carbohydrate intermolecular NOEs were obtained from 2D ¹³C/¹⁵N F₂-filtered NOESY and 3D ¹³C F₁-edited F₃-filtered HSQC-NOESY (47) spectra, recorded at 293 K or 303 K, with mixing times of 60–150 ms. Spectra were processed in TopSpin 3.0 (Bruker) and analyzed in Sparky (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco). The ¹H chemical shifts are referenced to 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS). The ¹³C and ¹⁵N chemical shifts are indirectly referenced using scaling factors of 0.251449530 and 0.101329118, respectively (48).

Structure Calculation and Refinement. The software package ATNOS CANDID (49, 50) was used for picking of initial cross-peaks of NOESY spectra. The resulting peak lists were assigned using a combination of manual and automated assignments within the NOEASSIGN module of CYANA 3.0 (51). For the complex, an extended CYANA library including the carbohydrate

residues was used, for which the topology was derived from a structural model of sLe^x that was generated in SWEET2 (www.glycosciences.de) (52) and modified for addition of the 3-aminopropyl-aglycon at C1 of GlcNAc and the sulfate group at C6 of Gal within ChemBio 3D Ultra 12.0 (Cambridge Soft). Structure calculations with torsion-angle dynamics were conducted within CYANA 3.0 (51), using additional restraints for backbone torsion angles derived from chemical shifts by using the program TALOS+ (53), stereospecific assignments of Val and Leu prochiral methyl groups, restraints for the disulfide bond generated with the CYANA macro sdbond, and distance restraints for intraprotein hydrogen bonds, based on the observation of slow-exchanging amide hydrogens in D₂O. Hydrogen bond acceptors were identified from inspection of initial structures. Two hydrogen bonds involving Tyr hydroxyl groups (Tyr106 and Tyr119) were identified from detection of their H¹ resonances in an F₂-filtered NOESY measured in H₂O and also in ¹⁵N- and ¹³C-edited 3D NOESY spectra, indicating protection from proton exchange, which typically prevents the observation of such signals. Intracarbohydrate and intermolecular (protein-carbohydrate) NOEs were assigned manually and converted into distance restraints based on cross-peak signal-to-noise (S/N) ratios. From 500 structures calculated in CYANA 3.0 (51) the 50 lowest-energy conformers were refined within AMBER 12 (54), using the ff12SB force field for the protein and simultaneously the GLYCAM_06h force field (55) for the protein-carbohydrate complex, in implicit solvent. For the 3-aminopropyl-linker the topology and parameters files were computed within Antechamber (56), using the General Amber Force Field (57) and the AM1-BCC (58) model for assignment of partial atomic charges. The 20 best structures were selected and analyzed with PROCHECK-NMR (59). The Ramachandran statistics for Siglec-8 (residues 7–136) are 89.7% in most favored and 10.3% in additionally favored regions for the free protein and 90.4% in most favored and 9.6% in additionally favored regions for the protein in the complex. The NMR and refinement statistics are provided in Table S1. Carbohydrate conformations in the NMR ensemble of the complex were compared by glycosidic torsion angles (φ , ψ), defined as $\varphi = \text{H1-C1-OX-CX}$ and $\psi = \text{C1-OX-CX-HX}$, where X is the number of the carbon atom of the second monosaccharide involved in the glycosidic linkage. For the Neu5Ac(α 2-3)Gal linkage, angles are defined as $\varphi = \text{C1-C2-OX-CX}$ and $\psi = \text{C2-OX-CX-HX}$. A-face and B-face of the described monosaccharides are defined as the faces on which the numbering of the carbon atoms increases in a clockwise and anticlockwise direction, respectively. Structural figures were prepared using PyMOL (Schrödinger, LLC).

NMR Titration Experiments. NMR titrations were carried out by recording a series of 2D ¹H,¹⁵N-HSQC spectra of ¹⁵N-labeled Siglec-8 lectin domain at 200 μ M in 20 mM potassium phosphate, 40 mM NaCl, pH 7.4 (5% D₂O), with the addition of increasing amounts of carbohydrate ligands. Titrations for comparison of Siglec-8 wild type and mutants binding to 6'S sLe^x were performed by a similar procedure, except that the starting protein concentrations were 100 μ M. All NMR spectra were recorded at 20 °C and at 500 MHz. Quantifications of NMR-binding isotherms were carried out only for nonoverlapping ¹H-¹⁵N protein resonances showing significant chemical shift changes in fast chemical exchange. Equilibrium dissociation constants (K_d) were obtained from nonlinear least-squares fit in MATLAB (MathWorks) to the equation (60)

$$\Delta\delta = \Delta\delta_{\text{max}} \left\{ \frac{([P]_t + [L]_t + K_d) - \sqrt{([L]_t + [P]_t + K_d)^2 - 4[P]_t[L]_t}}{2[P]_t} \right\}$$

where $[P]_t$ and $[L]_t$ are total concentrations of protein and ligand, respectively, which were calculated for each titration point with $[P]_t = ([P]_i/V_i + V_{\text{ad}})$ and $[L]_t = ([L]_i V_{\text{ad}}/(V_i + V_{\text{ad}}))$, where $[P]_i$ and V_i are the initial concentrations and initial volume of the protein sample, V_{ad} is the total volume of added ligand, and $[L]_i$ is the concentration of the ligand stock solution. $\Delta\delta$ is the observed chemical shift change, and $\Delta\delta_{\text{max}}$ is the chemical shift change at saturation, both calculated as a combination of ¹H and ¹⁵N chemical shift changes according to

$$\Delta\delta = \sqrt{(\Delta\delta^1\text{H})^2 + (\Delta\delta^{15}\text{N}/5)^2},$$

where $\Delta\delta^1\text{H}$ and $\Delta\delta^{15}\text{N}$ denote the chemical shift differences in parts per million (ppm) of amide hydrogen and nitrogen atoms, respectively, between the free and the carbohydrate-bound states of the protein.

Isothermal Titration Calorimetry. ITC was performed using a VP-ITC microcalorimeter (MicroCal). Prior the experiment the carbohydrate was dialyzed against 20 mM potassium phosphate, 40 mM NaCl, pH 7.4 (same buffer as used for the protein), using a Micro DispoDIALYZER (100 Da MWCO; Harvard Apparatus). The microsyringe was loaded with a solution of carbohydrate (7 mM), and the sample cell was loaded with a solution of protein (200 μ M). Titration was conducted at 20 °C using one initial injection of 2 μ L with a duration of 4 s and 400 s spacing, followed by 34 identical injections of 4 μ L with a duration of 8 s per injection and 10 min spacing between injections. Data were analyzed using the MicroCal ITC module of Origin 7.0 (OriginLab), applying a single-site binding model.

Competitive Binding Assay. Polyacrylamide-based competitive binding assay (31) was used to evaluate the binding affinity of Siglec-8 for 6'S sLe^x. Microtiter plates (F96 MaxiSorp, Nunc) were coated with 100 μ L per well of a 10 μ g/mL solution of Siglec-8 lectin domain in assay buffer (20 mM HEPES, pH 7.4; 150 mM NaCl; and 1 mM CaCl₂), overnight at 4 °C. The coating solution was discarded, and the wells were blocked with 150 μ L per well of 3% BSA in assay buffer for 2 h at 4 °C. After three washing steps with assay buffer (150 μ L per well), a threefold serial dilution of the 6'S sLe^x ligand (50 μ L per well) in assay buffer and streptavidin-peroxidase coupled polyacrylamide glycopolymer [6'S sLe^x-PAA (Lectinity), 50 μ L per well of a 0.3 μ g/mL solution] were added. The plate was incubated on a thermoshaker (PHMP-4, Grant Instruments) for 3 h at 25 °C and shaking speed 350 rpm and then carefully washed four times with 150 μ L per well assay buffer. After the addition of 100 μ L per well of the horseradish peroxidase substrate 2,2'-azino-di(3-ethylbenzthiazoline-6-sulfonic acid), the colorimetric reaction was allowed to develop for 3 min, then stopped by the addition of 2% aqueous oxalic acid, before the optical density was measured at 415 nm on a microplate-reader (Spectramax 190, Molecular Devices). The IC₅₀ value of 6'S sLe^x was calculated with the Prism software (GraphPad Software, Inc.). The IC₅₀ defines the molar concentration of the test compound that reduces the maximal specific binding of 6'S sLe^x-PAA polymer to Siglec-8 by 50%.

Modeling of Human Siglec-8-6,6'S sLe^x Complex. A structural model of the Siglec-8-6,6'S sLe^x interaction was generated by using the Siglec-8-6'S sLe^x complex NMR structure as a template. The bound 6'S sLe^x was modified by addition of a sulfate group to the GlcNAc-6 position in ChemBio 3D Ultra 12.0 (Cambridge Soft). To eliminate steric clashes with the inserted sulfate moiety, Lys120 and Gln122 side chains were adjusted with Pymol (Schrödinger, LLC). For further refinement, the model was subjected to energy minimization, followed by a 100-ps simulated annealing molecular dynamics in AMBER 12 (54), with the ff12SB force field and the GLYCAM_06h parameters (55) for the carbohydrate, during which the terminal Neu5Ac (only heavy atoms) was held rigid by imposing a weak positional force constant of 2 kcal/mol.Å², without using any additional restraints.

ACKNOWLEDGMENTS. We thank M. Blatter and P. Barraud for advice and help with structure calculations; J. Boudet for support concerning the ITC instrument and helpful discussions; F. Damberger, G. Wider, T. Suter-Stahel, and C. Maris for ensuring the best performance of the NMR infrastructure; M. Aebi, S. Campagne, G. Dorn, and K. Fowler for helpful discussions and critically reading the manuscript; and R. Woods for providing GLYCAM. This work was supported by the Swiss National Science Foundation Sinergia Grant CRSII3_127333 (to M.S., F.H.T.A., B.E., and Markus Aebi).

- Varki A (2008) Sialic acids in human health and disease. *Trends Mol Med* 14(8):351–360.
- Crocker PR, Paulson JC, Varki A (2007) Siglecs and their roles in the immune system. *Nat Rev Immunol* 7(4):255–266.
- Macauley MS, Crocker PR, Paulson JC (2014) Siglec-mediated regulation of immune cell function in disease. *Nat Rev Immunol* 14(10):653–666.
- Kletter D, Singh S, Bern M, Haab BB (2013) Global comparisons of lectin-glycan interactions using a database of analyzed glycan array data. *Mol Cell Proteomics* 12(4):1026–1035.
- Varki A, Angata T (2006) Siglecs—The major subfamily of I-type lectins. *Glycobiology* 16(1):1R–27R.

- Ravetch JV, Lanier LL (2000) Immune inhibitory receptors. *Science* 290(5489):84–89.
- von Gunten S, Simon HU (2006) Sialic acid binding immunoglobulin-like lectins may regulate innate immune responses by modulating the life span of granulocytes. *FASEB J* 20(6):601–605.
- Floyd H, et al. (2000) Siglec-8. A novel eosinophil-specific member of the immunoglobulin superfamily. *J Biol Chem* 275(2):861–866.
- Kikly KK, et al. (2000) Identification of SAF-2, a novel siglec expressed on eosinophils, mast cells, and basophils. *J Allergy Clin Immunol* 105(6 Pt 1):1093–1100.
- Nutku E, Aizawa H, Hudson SA, Bochner BS (2003) Ligand of Siglec-8: A selective mechanism for induction of human eosinophil apoptosis. *Blood* 101(12):5014–5020.

11. Hudson SA, Bovin NV, Schnaar RL, Crocker PR, Bochner BS (2009) Eosinophil-selective binding and proapoptotic effect in vitro of a synthetic Siglec-8 ligand, polymeric 6'-sulfated sialyl Lewis x. *J Pharmacol Exp Ther* 330(2):608–612.
12. Yokoi H, et al. (2008) Inhibition of FcεpsilonRI-dependent mediator release and calcium flux from human mast cells by sialic acid-binding immunoglobulin-like lectin 8 engagement. *J Allergy Clin Immunol* 121(2):499–505.e1.
13. Nutku-Bilir E, Hudson SA, Bochner BS (2008) Interleukin-5 priming of human eosinophils alters siglec-8 mediated apoptosis pathways. *Am J Respir Cell Mol Biol* 38(1):121–124.
14. Na HJ, Hudson SA, Bochner BS (2012) IL-33 enhances Siglec-8 mediated apoptosis of human eosinophils. *Cytokine* 57(1):169–174.
15. Kano G, Almanan M, Bochner BS, Zimmermann N (2013) Mechanism of Siglec-8-mediated cell death in IL-5-activated eosinophils: Role for reactive oxygen species-enhanced MEK/ERK activation. *J Allergy Clin Immunol* 132(2):437–445.
16. von Gunten S, et al. (2007) Intravenous immunoglobulin preparations contain anti-Siglec-8 autoantibodies. *J Allergy Clin Immunol* 119(4):1005–1011.
17. Jia Y, et al. (2015) Expression of ligands for Siglec-8 and Siglec-9 in human airways and airway cells. *J Allergy Clin Immunol* 135(3):799–810.e7.
18. Kiwamoto T, Katoh T, Tiemeyer M, Bochner BS (2013) The role of lung epithelial ligands for Siglec-8 and Siglec-F in eosinophilic inflammation. *Curr Opin Allergy Clin Immunol* 13(1):106–111.
19. Kim HY, DeKruyff RH, Umetsu DT (2010) The many paths to asthma: Phenotype shaped by innate and adaptive immunity. *Nat Immunol* 11(7):577–584.
20. Ilmarinen P, Kankaanranta H (2014) Eosinophil apoptosis as a therapeutic target in allergic asthma. *Basic Clin Pharmacol Toxicol* 114(1):109–117.
21. Gao PS, et al. (2010) Polymorphisms in the sialic acid-binding immunoglobulin-like lectin-8 (Siglec-8) gene are associated with susceptibility to asthma. *Eur J Hum Genet* 18(6):713–719.
22. O'Reilly MK, Paulson JC (2009) Siglecs as targets for therapy in immune-cell-mediated disease. *Trends Pharmacol Sci* 30(5):240–248.
23. Kiwamoto T, Kawasaki N, Paulson JC, Bochner BS (2012) Siglec-8 as a druggable target to treat eosinophil and mast cell-associated conditions. *Pharmacol Ther* 135(3):327–336.
24. Fulkerson PC, Rothenberg ME (2013) Targeting eosinophils in allergy, inflammation and beyond. *Nat Rev Drug Discov* 12(2):117–129.
25. Travers J, Rothenberg ME (2015) Eosinophils in mucosal immune responses. *Mucosal Immunol* 8(3):464–475.
26. Bochner BS, et al. (2005) Glycan array screening reveals a candidate ligand for Siglec-8. *J Biol Chem* 280(6):4307–4312.
27. Tateno H, Crocker PR, Paulson JC (2005) Mouse Siglec-F and human Siglec-8 are functionally convergent paralogs that are selectively expressed on eosinophils and recognize 6'-sulfato-sialyl Lewis X as a preferred glycan ligand. *Glycobiology* 15(11):1125–1135.
28. Marth JD, Grewal PK (2008) Mammalian glycosylation in immunity. *Nat Rev Immunol* 8(11):874–887.
29. Campanero-Rhodes MA, et al. (2006) Carbohydrate microarrays reveal sulphation as a modulator of siglec binding. *Biochem Biophys Res Commun* 344(4):1141–1146.
30. Pröpster JM, Yang F, Ernst B, Allain FH, Schubert M (2015) Functional Siglec lectin domains from soluble expression in the cytoplasm of *Escherichia coli*. *Protein Expr Purif* 109:14–22.
31. Rabbani S, Jiang X, Schwardt O, Ernst B (2010) Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal Biochem* 407(2):188–195.
32. May AP, Robinson RC, Vinson M, Crocker PR, Jones EY (1998) Crystal structure of the N-terminal domain of sialoadhesin in complex with 3' sialyllactose at 1.85 Å resolution. *Mol Cell* 1(5):719–728.
33. Zhuravleva MA, Trandem K, Sun PD (2008) Structural implications of Siglec-5-mediated sialoglycan recognition. *J Mol Biol* 375(2):437–447.
34. Alpey MS, Attrill H, Crocker PR, van Aalten DM (2003) High resolution crystal structures of Siglec-7. Insights into ligand specificity in the Siglec family. *J Biol Chem* 278(5):3372–3377.
35. Yamaji T, Teranishi T, Alpey MS, Crocker PR, Hashimoto Y (2002) A small region of the natural killer cell receptor, Siglec-7, is responsible for its preferred binding to alpha 2,8-disialyl and branched alpha 2,6-sialyl residues. A comparison with Siglec-9. *J Biol Chem* 277(8):6324–6332.
36. Attrill H, et al. (2006) Siglec-7 undergoes a major conformational change when complexed with the alpha(2,8)-disialylganglioside GT1b. *J Biol Chem* 281(43):32774–32783.
37. Zierke M, et al. (2013) Stabilization of branched oligosaccharides: Lewis(x) benefits from a nonconventional C-H...O hydrogen bond. *J Am Chem Soc* 135(36):13464–13472.
38. Battistel MD, Azurmendi HF, Frank M, Freedberg DI (2015) Uncovering non-conventional and conventional hydrogen bonds in oligosaccharides through NMR experiments and molecular modeling: Application to sialyl Lewis-X. *J Am Chem Soc* 137(42):13444–13447.
39. Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010) ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* 38(Web Server issue):W529–W533.
40. Selzer T, Albeck S, Schreiber G (2000) Rational design of faster associating and tighter binding protein complexes. *Nat Struct Biol* 7(7):537–541.
41. Auweter SD, et al. (2006) Molecular basis of RNA recognition by the human alternative splicing factor Fox-1. *EMBO J* 25(1):163–173.
42. Kiwamoto T, et al. (2014) Mice deficient in the ST3Gal3 gene product alpha2,3 sialyltransferase (ST3Gal-III) exhibit enhanced allergic eosinophilic airway inflammation. *J Allergy Clin Immunol* 133(1):240–247.e3.
43. Taylor ME, Drickamer K (2009) Structural insights into what glycan arrays tell us about how glycan-binding proteins interact with their ligands. *Glycobiology* 19(11):1155–1162.
44. Patnode ML, et al. (2013) Galactose 6-O-sulfotransferases are not required for the generation of Siglec-F ligands in leukocytes or lung tissue. *J Biol Chem* 288(37):26533–26545.
45. Kiwamoto T, et al. (2015) Endogenous airway mucins carry glycans that bind Siglec-F and induce eosinophil apoptosis. *J Allergy Clin Immunol* 135(5):1329–1340.e9.
46. Neri D, Szyperski T, Otting G, Senn H, Wüthrich K (1989) Stereospecific nuclear magnetic resonance assignments of the methyl groups of valine and leucine in the DNA-binding domain of the 434 repressor by biosynthetically directed fractional ¹³C labeling. *Biochemistry* 28(19):7510–7516.
47. Dominguez C, Schubert M, Duss O, Ravindranathan S, Allain FH (2011) Structure determination and dynamics of protein-RNA complexes by NMR spectroscopy. *Prog Nucl Magn Reson Spectrosc* 58(1–2):1–61.
48. Markley JL, et al.; IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy (1998) Recommendations for the presentation of NMR structures of proteins and nucleic acids. *J Biomol NMR* 12(1):1–23.
49. Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J Mol Biol* 319(1):209–227.
50. Herrmann T, Güntert P, Wüthrich K (2002) Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *J Biomol NMR* 24(3):171–189.
51. Güntert P (2004) Automated NMR structure calculation with CYANA. *Methods Mol Biol* 278:353–378.
52. Bohne A, Lang E, von der Lieth CW (1999) SWEET - WWW-based rapid 3D construction of oligo- and polysaccharides. *Bioinformatics* 15(9):767–768.
53. Shen Y, Delaglio F, Cornilescu G, Bax A (2009) TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44(4):213–223.
54. Case DA, et al. (2012) AMBER 12 (University of California, San Francisco).
55. Kirschner KN, et al. (2008) GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J Comput Chem* 29(4):622–655.
56. Wang J, Wang W, Kollman PA, Case DA (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25(2):247–260.
57. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. *J Comput Chem* 25(9):1157–1174.
58. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23(16):1623–1641.
59. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8(4):477–486.
60. Williamson MP (2013) Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* 73:1–16.
61. Gargaro AR, et al. (1996) NMR detection of arginine-ligand interactions in complexes of *Lactobacillus casei* dihydrofolate reductase. *Eur J Biochem* 238(2):435–439.
62. Sharma D, Rajarathnam K (2000) ¹³C NMR chemical shifts can predict disulfide bond formation. *J Biomol NMR* 18(2):165–171.