Published in final edited form as:

Proteins. 2014 November ; 82(11): 3079–3089. doi:10.1002/prot.24668.

Terminal sialic acids on CD44 N-glycans can block hyaluronan binding by forming competing intramolecular contacts with arginine sidechains

Christina E. Faller and Olgun Guvench*

Department of Pharmaceutical Sciences, University of New England College of Pharmacy, 716 Stevens Ave., Portland, ME 04103

Abstract

Specific sugar residues and their linkages form the basis of molecular recognition for interactions of glycoproteins with other biomolecules. Seemingly small changes, like the addition of a single monosaccharide in the covalently attached glycan component of glycoproteins, can greatly affect these interactions. For instance, the sialic acid capping of glycans affects protein-ligand binding involved in cell-cell and cell-matrix interactions. CD44 is a single-pass transmembrane glycoprotein whose binding with its carbohydrate ligand hyaluronan (HA), an extracellular matrix component, mediates processes such as leukocyte homing, cell adhesion, and tumor metastasis. This binding is highly regulated by glycosylation of the N-terminal extracellular hyaluronan-binding domain (HABD); specifically, sialic acid capped N-glycans of HABD inhibit ligand binding. However, the molecular mechanism behind this sialic acid mediated regulation has remained unknown. Two of the five N-glycosylation sites of HABD have been previously identified as having the greatest inhibitory effect on HA binding, but only if the glycans contain terminal sialic acid residues. These two sites, Asn25 and Asn120, were chosen for *in silico* glycosylation in this study. Here, from extensive standard molecular dynamics simulations and biased simulations, we propose a molecular mechanism for this behavior based on spontaneously-formed charge-paired hydrogen bonding interactions between the negatively-charged sialic acid residues and positively-charged Arg sidechains known to be critically important for binding to HA, which itself is negatively charged. Such intramolecular hydrogen bonds would preclude associations critical to hyaluronan binding. This observation suggests how CD44 and related glycoprotein binding is regulated by sialylation as cellular environments fluctuate.

Keywords

molecular dynamics; carbohydrate; extracellular matrix; glycosylation; hydrogen bond; charge pairing

*oguvench@une.edu.

Introduction

N-glycosylation is a highly regulated post-translational protein modification. Changes in the cellular environment can modulate post-translational enzyme activities and glycan structures, thus influencing function and molecular recognition. Many protein-ligand interactions and cell adhesion/trafficking functions are regulated by glycosylation.¹ In particular, sialic acid capping sugars (i.e. those attached to the non-reducing end termini of N-glycans) are often implicated in enabling or inhibiting molecular recognition.^{2,3} “Sialic acid” generally refers to a family of monosaccharides, but here we use the term to refer to one member of this family: N-acetylneuraminic acid (Neu5Ac). As physiological conditions change, endogenous glycosidase and glycotransferase enzymes remodel glycans in order to meet the needs of the cell, adaptively modifying glycoproteins on the cell surface.⁴ One such example is the removal of terminal sialic acids in order to “unmask” a protein to allow it to bind its ligand. There are four known human sialidases with various cellular locations and substrate specificities;⁵ thus, activation of these specific enzymes can cause remodeling of glycoproteins in the cellular environment.⁶⁻⁸

Regulation of protein-ligand interactions by glycans, and specifically by terminal sialic acids, has been noted in a number of glycoproteins. For example, CD44 can be an E-Selectin ligand when sialylated and fucosylated to display the sialyl-Lewis^X tetrasaccharide on the non-reducing end of the glycan. This association is critical to immune functions such as leukocyte recruitment to sites of inflammation.⁹ PTX3, a pattern recognition molecule, is expressed at sites of infection and inflammation and its functions depend on the identity of a single complex type biantennary N-glycan, which is often sialylated and fucosylated.¹⁰ LYVE-1, a CD44 homolog and hyaluronan receptor is inactive when it is sialylated.¹¹ Additionally, certain integrins are regulated by sialic acid: sialylation of β_1 integrins can protect cells from apoptosis via galectin-3 binding¹² and removal of sialic acid from $\alpha_3\beta_1$ integrin improves binding of fibronectin and collagen type IV;^{13,14} conversely, in $\alpha_5\beta_1$, the α_5 subunit requires sialic acid to bind fibronectin.¹⁵ The Natural Killer cell receptor CD244 cannot bind CD48-bearing target tumor cells when sialylated.¹⁶ And, the tyrosine kinase receptor TrKA must be de-sialylated in order to be active.¹⁷

CD44 is a single-pass transmembrane glycoprotein that is implicated in cell adhesion and migration via interactions with the extracellular matrix component hyaluronan, and thus contributes to cancer invasiveness and metastasis.¹⁸⁻²⁰ CD44 has many known forms and functions, which depend on cell type and environmental factors. Alternative splicing and post-translational modifications like N- and O-glycosylation,²¹⁻²⁵ palmitoylation,^{26,27} phosphorylation,^{28,29} and glycosaminoglycan addition contribute to the structural diversity of this glycoprotein. Notably, the globular extracellular domain of CD44, which contains the binding site for its carbohydrate ligand hyaluronan, is highly conserved across the many splice variants of the protein.²⁵ Several basic residues in the CD44 hyaluronan-binding domain (HABD) are known to be critical for binding the negatively charged hyaluronan, as evidenced by single-point mutant binding assays,³⁰ crystallographic data from the mouse CD44-hyaluronan complex,³¹ and by previous computational studies.³²

N-glycans on HABD have been shown to affect the ability of CD44 to bind hyaluronan.^{21-24,33-35} N-glycosylation can occur at Asn residues in an Asn-X-Ser/Thr sequence, where X is not Pro. This sequence is considered the N-glycosylation consensus “sequon.” According to its amino acid sequence, human HABD (residues 21-181) contains five possible N-glycosylation sites: Asn residues 25, 57, 100, 110, and 120 (Fig. 1).²⁵ Sialic acid-capped glycans in particular have been identified as inhibitors of the CD44-HA interaction. Enzymatic removal of the sialic acid-capping sugar has been shown to increase HA-binding affinity in some cell lines.^{21-24,36-38} Specifically, α -2,3-linked sialic acids have been shown to negatively regulate HA binding.^{21,24} Under physiological conditions, a variety of regulators, including endogenous sialidases, are involved in remodeling glycans to meet a cell's needs as its environment changes. This may explain why CD44 has such a variable affinity for HA.

Glycans are highly flexible, and a given glycan can assume multiple conformations at physiological temperatures.^{39,40} This is in contrast to many proteins that fold to a single well-defined three dimensional (3D) conformation. Therefore, compared to the many proteins whose unique atomic coordinates have been determined by x-ray crystallography or NMR studies, relatively few 3D structures of glycan components of glycoproteins exist, especially beyond the first or second residue from the point of attachment to the protein. This applies to the case of CD44, where attempts to explain how the presence of capping sialic acid prevents CD44-HA binding have lacked the benefit of atomic-resolution 3D structural information, which is limited to the protein component of this glycoprotein.^{41,42}

Towards atomic-resolution 3D characterization of glycoproteins, tools have been developed to “build” the glycan components of glycoproteins *in silico*. These include the web-based tools from Glycosciences (<http://www.glycosciences.de>)⁴³ and CHARMM-GUI (<http://www.charmm-gui.org>).^{44,45} The Glycosciences GlyProt tool generates 3D coordinates for glycans based on one of a manifold of potential structures, and chosen using statistically common ϕ , ψ dihedral angles for glycosidic linkages from available glycoprotein structures in the PDB such that glycan and protein coordinates do not overlap.⁴³ CHARMM-GUI is capable of detecting glycans and glycosidic linkages via its Glycan Reader module.^{44,45} The glycan fragment database (GFDB), available at <http://www.glycanstructure.org>, allows users to extract 3D structural information for glycans from those present in the PDB.⁴⁶ These and similar tools⁴⁷ produce static structures, whereas the inherent flexibility of glycans suggests a dynamic viewpoint, for example by molecular dynamics (MD) simulations, is best suited to understanding their structure-function relationships.

In an effort to understand the mechanism by which sialic acid capping sugars on N-glycans can inhibit the CD44-HA interaction, we performed multiple 100-ns MD simulations of human CD44 HABD with two different N-glycans, with the covalent attachment sites at Asn25 and Asn120. These glycosylation sites were selected based on mutagenesis studies showing an Asn25Ser or an Asn120Ser point mutation enabled otherwise inducible CD44-expressing cells to constitutively bind HA by preventing N-glycan addition at either sites.²¹ The potential glycosylation sites, in red, and basic residues of HABD, in blue, are depicted in Fig. 2; arrows indicate the Asn25 and Asn120 residues used for glycosylation and Arg41, a critical HA-binding residue.^{30,48} The N-glycans considered here were complex type

biantennary structures: one with a capping α -2,3-linked sialic acid on the penultimate galactose residues (Fig. 3, top), the other without (Fig. 3, bottom). Complex-type N-glycans have been shown to inhibit HA binding by CD44, such that blocking the metabolic pathway for processing complex N-glycans restores binding.^{33,49} Capping sialic acids were included based on experimental evidence that they are inhibitors of the CD44-HA interaction.^{21-24,36-38} An α -2,3-linkage for sialic acid was used, as treatment with an α -2,3-specific sialidase has been shown to restore HA-binding in some cell lines.^{21,24}

Methods

Preparation of glycoprotein and control systems

A set of 12 glycoproteins was built: 6 glycoforms based on the human CD44 hyaluronan-binding domain PDB ID 1UUH⁴² (which has ordered C-terminal residues) and the other 6 based on the ligand-bound structure of model 18 of PDB ID 2I83 (which has a similar HA binding site geometry compared to 1UUH but disordered C-terminal residues).⁴¹ Control systems were built without glycans for both structures for a total of 14 systems. Glycoprotein structures were built using GlyProt⁴³ (available at <http://www.glycosciences.de/modeling/glyprot>): the sialic acid capped glycan, α -Neu5Ac-(2->3)- β -D-Gal-(1->4)- β -D-GlcNAc-(1->2)- α -D-Man-(1->3)-[α -Neu5Ac-(2->3)- β -D-Gal-(1->4)- β -D-GlcNAc-(1->2)- α -D-Man-(1->6)]- β -D-Man-(1->4)- β -D-GlcNAc-(1->4)-[α -L-Fuc-(1->6)]- β -D-GlcNAc-Asn, and the asialo- version, β -D-Gal-(1->4)- β -D-GlcNAc-(1->2)- α -D-Man-(1->3)-[β -D-Gal-(1->4)- β -D-GlcNAc-(1->2)- α -D-Man-(1->6)]- β -D-Man-(1->4)- β -D-GlcNAc-(1->4)-[α -L-Fuc-(1->6)]- β -D-GlcNAc-Asn (as in glycan #8388 and #8386 from the GlyProt database, respectively) (Fig. 3).

The *in silico* glycosylated proteins were then solvated in boxes of TIP3P^{50,51} water molecules and charges were balanced with sodium ions as needed using CHARMM-GUI (<http://www.charmm-gui.org/input/glycan>),⁴⁵ a user-friendly, web-based tool which implements CHARMM software^{52,53} on the backend of its web-based interface for generating input files for MD simulations using, in this case, the glycoprotein coordinate file from GlyProt. Selenomethionine residues in 1UUH were reverted to their biological identities of Met63 and Met66. His residues were protonated at the δ -position. The 12 glycoprotein structures were visually examined for proper glycosidic linkages and stereochemistry. In some cases coordinate files were manually adjusted to correct minor errors. Table I summarizes all 14 systems. Fig. 4 depicts the starting structure for the sialylated, doubly glycosylated HABD simulations.

Standard MD Simulations

Systems were represented using the CHARMM36 (C36) all-atom additive force fields for proteins⁵⁴⁻⁵⁶ and carbohydrates⁵⁷⁻⁵⁹ and the TIP3P^{50,51} three-site rigid water molecule, and incorporated counter ions as needed to balance charges for an electrically neutral system. Each of the 14 unique systems (12 glycoproteins along with 2 control proteins lacking glycans) was simulated in triplicate using different random seeds for assignment of initial velocities. MD simulations were done with the NAMD software version 2.9.⁶⁰ Periodic boundary conditions were used for energy calculations with a cubic periodic unit. Bonds

involving hydrogen were constrained to their equilibrium lengths with the SHAKE algorithm⁶¹ and the SETTLE⁶² algorithm was used to maintain a rigid water molecule equilibrium geometry. The Langevin thermostat⁶³ and Langevin barostat⁶⁴ were used with the Brunker-Brooks-Karplus integrator and a 0.002 ps time-step.⁶⁵ Lennard-Jones interactions were smoothly truncated using an 8 Å to 10 Å switching distance.⁶⁶ Electrostatic interactions beyond the 10 Å cutoff were computed via the particle-mesh Ewald method, using a ~1 Å grid-spacing.⁶⁷ Each system was energy minimized for 1000 steps. Systems were heated to and equilibrated at 310 K and 1 atm for 20,000 steps (40 ps). During heating/equilibration, velocities were reassigned every 1000 steps, and harmonic potentials were used to restrain glycan and protein non-hydrogen atom positions. Production simulations were fully unrestrained.

Standard MD Simulation Analysis

In production simulations, snapshots were saved every 5000 steps (10 ps) for analysis, each of the 42 simulation trajectories was 100 ns in length (4.2 μs total sampling), and analysis was based on the final 65 ns of each simulation. The CORREL facility in CHARMM^{52,53} was used to calculate minimum contact distances between Arg/Lys polar side chain hydrogen atoms and either the carboxylate oxygen atoms of the sialic acids or the set of terminal galactose atoms, for sialic acid terminated glycans and their asialo- counterparts, respectively (a note about semantics: in this manuscript, we refer to the interaction between Arg sidechains and sialic acid as “charge-paired hydrogen bonding” when the distance between Arg sidechain polar hydrogen atoms and sialic acid carboxylate oxygen atoms is less than 3 Å). Each Arg/Lys residue was individually paired with each terminal monosaccharide for analysis. Hence for singly glycosylated structures 24 hydrogen bond donor-acceptor pairs were analyzed ((8 Arg + 4 Lys in HABD) × 2 terminal monosaccharides per biantennary glycan). In the case of doubly glycosylated structures, 48 hydrogen bond donor-acceptor pairs were analyzed. Employing these hydrogen bond donor and acceptor definitions, each resultant time series represents the minimum distance between the nearest pair of donor/acceptor atoms for the stated amino acid/monosaccharide pairs.

The C_α RMSDs of protein and glycoprotein structures relative to the starting structures were calculated for each of the 42 trajectories using snapshots every 0.5 ns. HABD residues included in analysis were residues 21-150, which excludes the C-terminal residues that are ordered in 1UUH but disordered in 2I83.

The percent of surface area of HA-binding groove occupied by glycans was calculated using the CORREL and COOR SURF⁶⁸ facilities in CHARMM, with a probe radius of 1.6 Å. The HA-binding residues here are defined by those within 3.5 Å of HA from the final snapshot of the 200 ns ABF trajectory of HABD in complex with HA⁶⁹ and in the absence of direct Arg41-HA contact: residues 42, 77, 78, 79, 88, 94, 96, 97, and 98.

Biased MD Simulations

MD simulations implementing the Adaptive Biasing Force (ABF)^{70,71} sampling method were performed to probe the strength of the sialic acid-Arg41 charge paired hydrogen

bonding interaction as a function of distance between the two groups. Triplicate simulations were seeded with a starting conformation extracted from snapshots with sialic acid-Arg41 binding in “ordered” trajectories. The collective variable module of NAMD was used to determine the potential of mean force $G(r)$, r being the distance between the Arg guanidinium carbon and the sialic acid carboxylate carbon from 3.82 Å to 7.42 Å, with bin sizes of 0.05 Å. The computed biasing force was not fully applied until 200 force samples had been collected for that bin.

Results and Discussion

The “ordered” versus “partially-disordered” distinction for CD44 structures refers to the conformational state of the C-terminal residues of HABD. The unfolding transition of these residues that occurs in going from the “ordered” to the “partially-disordered” form is associated with an increase in HA-binding affinity.^{41,72} Additionally, truncation mutation of HABD that deletes these C-terminal residues results in reduced HA-binding affinity,³⁰ while a Tyr161Ala mutant that adopts only the partially-disordered conformation has increased HA-binding relative to wild type.⁷² NMR data clearly show an unfolding transition in residues 153-169 associated with HA binding⁴¹ (*n.b.*: HABD residues beyond 169 are always disordered,^{41,42} and therefore, due to lack of structural information, were not included in the present study). From computational studies, it was discovered that the unfolding transition of these seventeen C-terminal residues does not have any effect on the thermodynamics of Arg41-HA contact.⁶⁹ Rather, the transition allows basic residues in the C-terminal region to gain conformational freedom to form additional direct basic residue-HA contacts. Therefore, the effects of N-glycosylation on Arg41-HA association can be analyzed independently of the transition.

Arg41 of CD44 has repeatedly been demonstrated to be the central basic amino acid in HA-binding, such that a single Arg41Ala point mutation effectively renders HABD incapable of binding HA.^{30,31} Recent computational studies described a pathway for HABD loop closure that enables direct Arg41 sidechain-HA contact⁷³ and provided strong thermodynamic evidence that this direct contact stabilizes binding.³² In the present work, the negatively-charged carboxylate group of capping sialic acids on glycans attached at Asn25 spontaneously form direct charge-paired hydrogen bonds with the positively-charged Arg41 sidechain guanidinium moiety. Such sialic acid-Arg41 contact would reduce the likelihood of Arg41 forming the above critical direct contact with HA. It is important to note that arginine has five polar sidechain hydrogen atoms, so it is capable of interacting with multiple partners simultaneously.

In two independent unbiased simulations of glycosylated CD44, one in the “ordered” S25 system and one in the “partially-disordered” S25S120 system, the Arg41 sidechain guanidinium formed spontaneous charge-paired hydrogen bonds with the capping sialic acid of the glycan attached at Asn25. The plot of the distance between sialic acid carboxylate oxygen atoms and Arg41 guanidinium hydrogen atoms reveals spontaneous and robust charge-paired hydrogen bonds between these two functional groups (Fig. 5). In simulations of both ordered (Fig. 5, top panel) and partially-disordered (Fig. 5, middle panel) CD44,

these respective atoms from sialic acid carboxylate and Arg41 guanidinium remain in close contact (separated by as little as 1.6 Å) for 32 ns and 35 ns, respectively.

A snapshot from the ordered S25 system trajectory yielding Fig. 5, top panel green (x) plot gives perspective on the size and bulk of the biantennary complex glycan relative to HABD: when the sialic acid-Arg41 interaction is present (Fig. 6, panel A white box; close-up in panel B), the HA-binding face of CD44 is partially covered by glycan. However, visual inspection reveals that this interaction does not position the glycan such that it directly occludes the HA-binding groove. Surface area analysis of the ordered S25 system, as detailed in Methods, shows that anywhere between 2 and 28% of HA-binding residue surface area is occupied by glycans at any moment during Arg41-sialic acid binding, with an average value of 12.4% during binding (Supporting Information Fig. S1, top panel, green (x) plot). Meanwhile, 0 to almost 50% occlusion of HA-binding residues may occur when the binding interaction is not present (Supporting Information Fig. S1, top panel, red (+) and blue (*) plots). However, across all the “ordered” system simulations, no glycosylation state results in more than 12% average occlusion (Supporting Information Table S1). That is, in this conformational sample steric hindrance by the various glycans is both independent of Arg41-sialic acid binding and relatively limited in nature, suggesting it is not the major contributor to binding inhibition. Rather, the observation of lengthy interactions of sialic acid with the critical HA-binding residue Arg41 points to a molecular mechanism for sialylated glycan inhibition of HA binding where HA would need to compete with the capping sialic acids of covalently attached glycans for the thermodynamically favorable interaction with Arg41.

Arg41Ala mutants of CD44 lack the guanidinium sidechain that stabilizes HA binding. Experimentally, it has been determined that this mutation reduces binding affinity for HA by 2.5 kcal/mol.³¹ Recent computational data show that the Arg41Ala mutation reduces the binding affinity in the ordered and disordered protein conformations by 2.2 kcal/mol and 2.3 kcal/mol, respectively⁶⁹, in quantitative agreement with the experimental data. Here, triplicate potential of mean force data (Fig. 7), though still not converged after 100 ns of sampling, suggest the Arg41-sialic acid interaction may be similarly favorable and can therefore compete with HA for binding to the functionally-critical Arg41 sidechain.

Arg154 is the most important single residue for HA-binding in the membrane proximal basic residue cluster of HABD, as mutation to Ala greatly reduces HA binding.³⁰ This residue is within the seventeen-residue HABD C-terminal sequence (residues 153-169) that unfolds during the transition from “ordered” to “partially disordered.” From the MD trajectories here, it is seen that terminal sialic acids on glycans attached at Asn25 can directly hydrogen bond with this basic residue in partially-disordered HABD. Spontaneous direct Arg154-sialic acid interactions were observed in two trajectories of partially-disordered glycosylated CD44. These close associations are maintained for 10+ ns at a time (Fig. 5, bottom panel, red(+) and blue(*) plots). A representative snapshot demonstrates the association of the Arg154 guanidinium sidechain with a glycan terminal sialic acid (Fig. 8, panel A). As with Arg41, hydrogen bonds form between the positively-charged Arg guanidinium group and the negatively-charged sialic acid carboxylate group (Fig. 8, panel B).

The Arg154-sialic acid interaction involving one glycan arm localizes the glycan near the HA-binding face, while the other glycan arm that is not directly involved retains enough flexibility to partially occlude the HA binding groove. In one trajectory, 20 to 40% of surface area of HA binding residues, as defined in Methods, is occupied by glycans during sialic acid-Arg154 binding in the red (+) plot, while another trajectory (blue (*)) plot has essentially no loss of HA binding residue surface area during the sialic acid-Arg154 binding interaction (Supporting Information Fig. S1, bottom panel). Taking the observations for Arg41 and Arg154 together, the common theme is sequestration of HABD basic sidechains that would otherwise be free to bind HA, with binding site occlusion due to steric hindrance a possible secondary consequence of the charge-paired hydrogen bonding that is the basis for sequestration. Table II summarizes the four trajectories with lengthy spontaneous association between capping sialic acids and either Arg41 or Arg154 sidechains. We note that, in contrast to the conformational flexibility of the glycans that enables interaction with the protein component while minimally obscuring the HA-binding groove (Supporting Information Table S1 and Table S2), visual inspection of final snapshots and plots of C_{α} RMSDs (Supporting Information Fig. S2) demonstrate that the protein structure is stable and relatively static in all cases in relation to starting conformation.

From previous computational work, it is Arg154 in the partially-disordered HABD state that forms direct contact with bound oligomeric HA.⁶⁹ And, such contact with Arg154 is the most prevalent among all possible basic residue-HA contacts involving the unfolded conformation of HABD residues 153-169, which include the basic residues Arg154, Lys158, and Arg162. Sialic acid-Arg154 contacts were not observed in the ordered state trajectories here, likely due to a lack of conformational freedom for these C-terminal residues. Therefore, the observed hydrogen bonding between capping sialic acids and Arg154 implies that the sialic acid-terminating N-glycans reduce HA-binding by sequestering disordered HABD C-terminal basic residues having sufficient conformational freedom to otherwise bind HA.

The Asn25Ser or Asn120Ser mutation, which inhibits N-glycosylation at either residue, causes an otherwise “inducible” cell line to constitutively bind HA.²¹ From the above-discussed data, direct interaction of sialic acid-capped glycans at Asn25 with important HA-binding residues is a likely underlying molecular mechanism of such glycan-mediated negative regulation of binding. However, in our study glycans attached at Asn120, whether capped with sialic acid or not, did not exhibit interactions with known HA-binding residues. Electrostatic repulsion of negatively charged HA by the negatively charged terminal sialic acids, or interactions of sialic acid capping sugars with HA-binding residues on nearby CD44 molecules could explain the experimental observation of improved HA binding in the absence of an N-glycan at Asn120. Alternatively, other regulatory mechanisms may be at play to fine tune HA-binding ability of CD44 in the cellular environment. And of course there is the possibility that such interactions would in fact be observed if computational resources allowed for significantly greater sampling.

With respect to the asialo glycoproteins studied here, the penultimate sugar of sialic acid-capped complex-type N-glycans is galactose. Therefore, enzymatic removal of sialic acid renders galactose the terminal monosaccharide in the asialo glycoform. In the cellular

environment, endogenous sialidases perform this function alongside a variety of enzymes that adaptively modify glycans as physiological conditions change, as described in the Introduction. In the case of CD44, removal of capping sialic acids has been associated with improved HA-binding.^{21-24,36-38} In this study, we observed that sialic acid, but not galactose, forms robust interactions with Arg41 and Arg154 residues that otherwise are available to stabilize HA binding. Terminal galactose sugars had very little interaction with any of the basic residues of HABD. The highest level of galactose-basic residue interaction (of 3 Å or less) occurred with Arg150, which accounted for a total of 13% of the trajectory. Qualitatively, however, the galactose-Arg150 interaction is very different from the sialic acid-Arg interactions discussed above. The robust interactions of sialic acid and basic residues of HABD were characterized by sustained hydrogen-bonding interactions for 10 ns or more, whereas the galactose-Arg interaction was only sustained for a few ns at a time as the association was frequently interrupted by separations of 7 to 20 Å. This trend of brief associations with many interruptions is shared by other instances of galactose-Arg or -Lys interactions. From these data, in contrast to sialic acid-Arg interactions, it seems unlikely that galactose-Arg interactions would have any significant impact on the formation HA-binding stabilizing contacts.

Conclusion

The results presented here from extensive unbiased and ABF all-atom explicit-solvent MD simulations provide an underlying molecular mechanism for previous experimental results that point to sialic acid-capped N-glycans as negative regulators of HA binding by CD44 HABD. This mechanism entails the capping sialic acids of glycans at Asn25 forming robust hydrogen bonds with HA-binding residues Arg41 and Arg154, precluding these residues from forming stabilizing interactions with HA. This is further supported by the apparent thermodynamic preference of direct charge-paired hydrogen bonding between Arg and sialic acid. In contrast, galactose-terminating glycans on HABD do not form such robust contacts. And although this study did not reveal interactions between sialic acid-capped glycans at Asn120 with HA-binding residues, it is possible that such interactions may occur with neighboring CD44 molecules in the context of multiple membrane-bound CD44 molecules.

Because of computational expense, the CD44 HABD glycoforms studied here do not include all possible biologically relevant states of CD44, and investigated two of five potential HABD N-glycosylation sites. However, given the complexity of the glycosylation process and the differences in glycosylation patterns among cell lines, the particular structures employed here encompassed two of the five N-glycosylation sites and the glycans that appear to be most relevant to HA binding regulation based on a survey of existing literature. Naturally, as computational resources become progressively more robust, an even more thorough survey of HABD glycoforms, supported by the experimental literature, becomes possible.

While sialic acid-capped glycans are known regulators of HA binding to CD44 HABD, it is likely that there are additional factors in the cellular environment. CD44 is a membrane protein that is alternatively spliced to include “variant exons” in its stalk region,²⁵ and these additional stalk residues provide attachment points for additional post-translational

modification such as glycosaminoglycan (GAG) addition.⁷⁴ Such modification potential in the extracellular stalk of CD44 is in contrast to the HABD, which is not known to have any sites for covalent GAG attachment. As GAGs carry significant amounts of negative charge, both from the carboxylate groups of their glucuronic and iduronic acid constituents and as a result of variable sulfation,⁷⁵ an interesting possibility is a similar mechanism to that noted here. That is, negatively-charged groups on CD44 stalk GAGs with high flexibility might act as hydrogen-bond acceptors to sequester basic amino acid sidechains involved in HA binding. Another possibility is alteration of the overall electric field of CD44 by such negatively-charged GAGs to repel HA at a distance. Complicating such pictures is the effect of cations,⁷⁶⁻⁷⁸ and in particular divalent cations,^{79,80} which are known to bind to GAGs and would neutralize the relevant negative charges.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant sponsor: National Institutes of Health; Grant number: R15 GM099022; Grant sponsor: This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575; Grant number: TG-MCB120007.

References

1. Marth JD, Grewal PK. Mammalian glycosylation in immunity. *Nat Rev Immunol.* 2008; 8(11):874–87. [PubMed: 18846099]
2. Varki, A.; Cummings, R.; Esko, J., et al. *Essentials of Glycobiology*. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1999.
3. Schauer R. Sialic acids as regulators of molecular and cellular interactions. *Curr Opin Struct Biol.* 2009; 19(5):507–514. [PubMed: 19699080]
4. Dennis JW, Lau KS, Demetriou M, Nabi IR. Adaptive regulation at the cell surface by N-glycosylation. *Traffic.* 2009; 10(11):1569–78. [PubMed: 19761541]
5. Miyagi, T.; Yamaguchi, K. Sialic Acids. In: Kamerling, J.; Boona, G.; Lee, Y.; Suzuki, A.; Taniguchi, N.; Voragen, A., editors. *Biochemistry of Glycans: sialic acids in comprehensive glycoscience*. Amsterdam: Elsevier BV; 2007. p. 297-322.
6. Monti E, Preti A, Venerando B, Borsani G. Recent development in mammalian sialidase molecular biology. *Neurochem Res.* 2002; 27(7-8):649–63. [PubMed: 12374200]
7. Mozzi A, Mazzacuva P, Zampella G, Forcella ME, Fusi PA, Monti E. Molecular insight into substrate recognition by human cytosolic sialidase NEU2. *Proteins.* 2012; 80(4):1123–1132. [PubMed: 22228546]
8. Monti E, Bonten E, D'Azzo A, Bresciani R, Venerando B, Borsani G, Schauer R, Tettamanti G. Sialidases in vertebrates: a family of enzymes tailored for several cell functions. *Adv Carbohydr Chem Biochem.* 2010; 64:403–79. [PubMed: 20837202]
9. Dimitroff CJ, Lee JY, Rafii S, Fuhlbrigge RC, Sackstein R. CD44 is a major E-selectin ligand on human hematopoietic progenitor cells. *J Cell Biol.* 2001; 153(6):1277–86. [PubMed: 11402070]
10. Inforzato A, Reading PC, Barbati E, Bottazzi B, Garlanda C, Mantovani A. The “sweet” side of a long pentraxin: how glycosylation affects PTX3 functions in innate immunity and inflammation. *Front Immunol.* 2012; 3:407. [PubMed: 23316195]
11. Nightingale TD, Frayne ME, Clasper S, Banerji S, Jackson DG. A mechanism of sialylation functionally silences the hyaluronan receptor LYVE-1 in lymphatic endothelium. *J Biol Chem.* 2009; 284(6):3935–45. [PubMed: 19033446]

12. Zhuo Y, Chammas R, Bellis SL. Sialylation of beta1 integrins blocks cell adhesion to galectin-3 and protects cells against galectin-3-induced apoptosis. *J Biol Chem*. 2008; 283(32):22177–85. [PubMed: 18676377]
13. Janik ME, Litynska A, Vereecken P. Cell migration-the role of integrin glycosylation. *Biochim Biophys Acta*. 2010; 1800(6):545–55. [PubMed: 20332015]
14. Pochec E, Litynska A, Bubka M, Amoresano A, Casbarra A. Characterization of the oligosaccharide component of alpha3beta1 integrin from human bladder carcinoma cell line T24 and its role in adhesion and migration. *Eur J Cell Biol*. 2006; 85(1):47–57. [PubMed: 16373174]
15. Nadanaka S, Sato C, Kitajima K, Katagiri K, Irie S, Yamagata T. Occurrence of oligosialic acids on integrin alpha 5 subunit and their involvement in cell adhesion to fibronectin. *J Biol Chem*. 2001; 276(36):33657–64. [PubMed: 11418585]
16. Margraf-Schonfeld S, Bohm C, Watzl C. Glycosylation affects ligand binding and function of the activating natural killer cell receptor 2B4 (CD244) protein. *J Biol Chem*. 2011; 286(27):24142–9. [PubMed: 21606496]
17. Woronowicz A, Amith SR, De Vusser K, Laroy W, Contreras R, Basta S, Szewczuk MR. Dependence of neurotrophic factor activation of Trk tyrosine kinase receptors on cellular sialidase. *Glycobiology*. 2007; 17(1):10–24. [PubMed: 16971381]
18. Marhaba R, Zoller M. CD44 in cancer progression: adhesion, migration and growth regulation. *J Mol Histol*. 2004; 35(3):211–31. [PubMed: 15339042]
19. Toole BP. Hyaluronan-CD44 Interactions in Cancer: Paradoxes and Possibilities. *Clin Cancer Res*. 2009; 15(24):7462–7468. [PubMed: 20008845]
20. Ponta H, Sherman L, Herrlich PA. CD44: From adhesion molecules to signalling regulators. *Nat Rev Mol Cell Biol*. 2003; 4(1):33. [PubMed: 12511867]
21. English NM, Lesley Jayne F, Hyman Robert. Site-specific De-N-glycosylation of CD44 Can Activate Hyaluronan Binding, and CD44 Activation State Shows Distinct Threshold Densities for Hyaluronan Binding. *Cancer Res*. 1998; 58(16):3736–3742. [PubMed: 9721887]
22. Katoh S, Miyagi T, Taniguchi H, Matsubara Y, Kadota J, Tominaga A, Kincade PW, Matsukura S, Kohno S. Cutting edge: an inducible sialidase regulates the hyaluronic acid binding ability of CD44-bearing human monocytes. *J Immunol*. 1999; 162(9):5058–61. [PubMed: 10227973]
23. Katoh S, Zheng Z, Oritani K, Shimoizato T, Kincade PW. Glycosylation of CD44 negatively regulates its recognition of hyaluronan. *J Exp Med*. 1995; 182(2):419–29. [PubMed: 7543137]
24. Skelton TP, Zeng C, Nocks A, Stamenkovic I. Glycosylation Provides Both Stimulatory and Inhibitory Effects on Cell Surface and Soluble CD44 Binding to Hyaluronan. *J Cell Biol*. 1998; 140(2):431–446. [PubMed: 9442118]
25. Screaton GR, Bell MV, Jackson DG, Cornelis FB, Gerth U, Bell JI. Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. *Proc Natl Acad Sci USA*. 1992; 89(24):12160–12164. [PubMed: 1465456]
26. Thankamony SP, Knudson W. Acylation of CD44 and its association with lipid rafts are required for receptor and hyaluronan endocytosis. *J Biol Chem*. 2006; 281(45):34601–34609. [PubMed: 16945930]
27. Lee JL, Wang MJ, Sudhir PR, Chen JY. CD44 engagement promotes matrix-derived survival through the CD44-SRC-integrin axis in lipid rafts. *Mol Cell Biol*. 2008; 28(18):5710–5723. [PubMed: 18644869]
28. Uff CR, Neame SJ, Isacke CM. Hyaluronan binding by CD44 is regulated by a phosphorylation-independent mechanism. *Eur J Immunol*. 1995; 25(7):1883–7. [PubMed: 7542594]
29. Neame SJ, Isacke CM. Phosphorylation of CD44 in vivo requires both Ser323 and Ser325, but does not regulate membrane localization or cytoskeletal interaction in epithelial cells. *EMBO J*. 1992; 11(13):4733–8. [PubMed: 1281449]
30. Peach RJ, Hollenbaugh D, Stamenkovic I, Aruffo A. Identification of hyaluronic acid binding sites in the extracellular domain of CD44. *J Cell Biol*. 1993; 122(1):257–64. [PubMed: 8314845]
31. Banerji S, Wright AJ, Noble M, Mahoney DJ, Campbell ID, Day AJ, Jackson DG. Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction. *Nat Struct Mol Biol*. 2007; 14(3):234–9. [PubMed: 17293874]

32. Favreau AJ, Faller CE, Guvench O. CD44 receptor unfolding enhances binding by freeing basic amino acids to contact carbohydrate ligand. *Biophys J*. 2013; 105(5):1217–26. [PubMed: 24010665]
33. Lesley J, English N, Perschl A, Gregoroff J, Hyman R. Variant cell lines selected for alterations in the function of the hyaluronan receptor CD44 show differences in glycosylation. *J Exp Med*. 1995; 182(2):431–7. [PubMed: 7543138]
34. Lesley J, He Q, Miyake K, Hamann A, Hyman R, Kincade PW. Requirements for hyaluronic acid binding by CD44: a role for the cytoplasmic domain and activation by antibody. *J Exp Med*. 1992; 175(1):257–66. [PubMed: 1730918]
35. Lesley J, Hyman R. CD44 can be activated to function as an hyaluronic acid receptor in normal murine T cells. *Eur J Immunol*. 1992; 22(10):2719–23. [PubMed: 1382996]
36. Gee K, Kozlowski M, Kumar A. Tumor necrosis factor- α induces functionally active hyaluronan-adhesive CD44 by activating sialidase through p38 mitogen-activated protein kinase in lipopolysaccharide-stimulated human monocytic cells. *J Biol Chem*. 2003; 278(39):37275–87. [PubMed: 12867430]
37. Kincade PW, Zheng Z, Katoh S, Hanson L. The importance of cellular environment to function of the CD44 matrix receptor. *Curr Opin Cell Biol*. 1997; 9(5):635–42. [PubMed: 9330866]
38. Katoh S, Maeda S, Fukuoka H, Wada T, Moriya S, Mori A, Yamaguchi K, Senda S, Miyagi T. A crucial role of sialidase Neu1 in hyaluronan receptor function of CD44 in T helper type 2-mediated airway inflammation of murine acute asthmatic model. *Clin Exp Immunol*. 2010; 161(2):233–41. [PubMed: 20491786]
39. Woods R. Computational carbohydrate chemistry: what theoretical methods can tell us. *Glycoconj J*. 1998; 15(3):209–216. [PubMed: 9579797]
40. Jo S, Lee HS, Skolnick J, Im W. Restricted N-glycan Conformational Space in the PDB and Its Implication in Glycan Structure Modeling. *PLoS Comput Biol*. 2013; 9(3):e1002946. [PubMed: 23516343]
41. Takeda M, Ogino S, Umemoto R, Sakakura M, Kajiware M, Sugahara KN, Hayasaka H, Miyasaka M, Terasawa H, Shimada I. Ligand-induced structural changes of the CD44 hyaluronan-binding domain revealed by NMR. *J Biol Chem*. 2006; 281(52):40089–95. [PubMed: 17085435]
42. Teriete P, Banerji S, Noble M, Blundell CD, Wright AJ, Pickford AR, Lowe E, Mahoney DJ, Tammi MI, Kahmann JD, et al. Structure of the regulatory hyaluronan binding domain in the inflammatory leukocyte homing receptor CD44. *Mol Cell*. 2004; 13(4):483–96. [PubMed: 14992719]
43. Bohné-Lang A, von der Lieth CW. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res*. 2005; 33(Web Server issue):W214–9. [PubMed: 15980456]
44. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem*. 2008; 29(11):1859–1865. [PubMed: 18351591]
45. Jo S, Song KC, Desaire H, MacKerell AD Jr, Im W. Glycan reader: Automated sugar identification and simulation preparation for carbohydrates and glycoproteins. *J Comput Chem*. 2011; 32(14):3135–3141. [PubMed: 21815173]
46. Jo S, Im W. Glycan fragment database: a database of PDB-based glycan 3D structures. *Nucleic Acids Res*. 2013; 41(Database issue):D470–4. [PubMed: 23104379]
47. Group W.. Complex Carbohydrate Research Center, University of Georgia; Athens, GA: 2005-2014 GLYCAM Web. <http://www.glycam.org>
48. Bajorath J, Greenfield B, Munro SB, Day AJ, Aruffo A. Identification of CD44 residues important for hyaluronan binding and delineation of the binding site. *J Biol Chem*. 1998; 273(1):338–43. [PubMed: 9417085]
49. Bartolazzi A, Nocks A, Aruffo A, Spring F, Stamenkovic I. Glycosylation of CD44 is implicated in CD44-mediated cell adhesion to hyaluronan. *J Cell Biol*. 1996; 132(6):1199–208. [PubMed: 8601595]
50. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys*. 1983; 79(2):926–935.
51. Durell SR, Brooks BR, Ben-Naim A. Solvent-Induced Forces between Two Hydrophilic Groups. *J Phys Chem*. 1994; 98(8):2198–2202.

52. Brooks BR, Brooks CL 3rd, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, et al. CHARMM: the biomolecular simulation program. *J Comput Chem*. 2009; 30(10):1545–614. [PubMed: 19444816]
53. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem*. 1983; 4(2):187–217.
54. MacKerell AD Jr, Bashford D, Bellott, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J Phys Chem B*. 1998; 102(18):3586–3616. [PubMed: 24889800]
55. Best RB, Zhu X, Shim J, Lopes PE, Mittal J, Feig M, Mackerell AD Jr. Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *J Chem Theory Comput*. 2012; 8(9):3257–3273. [PubMed: 23341755]
56. MacKerell AD Jr, Feig M, Brooks CL 3rd. Improved treatment of the protein backbone in empirical force fields. *J Am Chem Soc*. 2004; 126(3):698–9. [PubMed: 14733527]
57. Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW, Mackerell AD Jr. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem*. 2008; 29(15):2543–64. [PubMed: 18470966]
58. Guvench O, Hatcher ER, Venable RM, Pastor RW, Mackerell AD Jr. CHARMM Additive All-Atom Force Field for Glycosidic Linkages between Hexopyranoses. *J Chem Theory Comput*. 2009; 5(9):2353–2370. [PubMed: 20161005]
59. Guvench O, Mallajosyula SS, Raman EP, Hatcher E, Vanommeslaeghe K, Foster TJ, Jamison FW Jr, Mackerell AD Jr. CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput*. 2011; 7(10):3162–3180. [PubMed: 22125473]
60. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. Scalable molecular dynamics with NAMD. *J Comput Chem*. 2005; 26(16):1781–1802. [PubMed: 16222654]
61. Ryckaert JP, Ciccotti G, Brenderson HJC. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J Comput Phys*. 1977; 23(3):327–341.
62. Miyamoto S, Kollman PA. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J Comput Chem*. 1992; 13(8):952–962.
63. Kubo, R.; Toda, M.; Hashitsume, N. *Statistical Physics II: Nonequilibrium Statistical Mechanics*. New York: Springer; 1991.
64. Feller SE, Zhang Y, Pastor RW, Brooks BR. Constant pressure molecular dynamics simulation: The Langevin piston method. *J Chem Phys*. 1995; 103(11):4613–4621.
65. Brünger A, Brooks CL 3rd, Karplus M. Stochastic boundary conditions for molecular dynamics simulations of ST2 water. *Chem Phys Lett*. 1984; 105(5):495–500.
66. Steinbach PJ, Brooks BR. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J Comput Chem*. 1994; 15(7):667–683.
67. Darden T, York D, Pedersen L. Particle mesh Ewald: an N-log(N) method for Ewald sums in large systems. *J Chem Phys*. 1993; 98:10089–10092.
68. Lee B, Richards FM. The interpretation of protein structures: estimation of static accessibility. *J Mol Biol*. 1971; 55(3):379–400. [PubMed: 5551392]
69. Favreau AJ, Faller CE, Guvench O. CD44 receptor unfolding enhances binding by freeing basic amino acids to contact carbohydrate ligand. *Biophys J*. 2013; 105(5):1217–26. [PubMed: 24010665]
70. Hénin J, Fiorin G, Chipot C, Klein ML. Exploring Multidimensional Free Energy Landscapes Using Time-Dependent Biases on Collective Variables. *J Chem Theory Comput*. 2010; 6(1):35–47.
71. Darve E, Rodriguez-Gomez D, Pohorille A. Adaptive biasing force method for scalar and vector free energy calculations. *J Chem Phys*. 2008; 128(14):144120. [PubMed: 18412436]

72. Ogino S, Nishida N, Umemoto R, Suzuki M, Takeda M, Terasawa H, Kitayama J, Matsumoto M, Hayasaka H, Miyasaka M, et al. Two-state conformations in the hyaluronan-binding domain regulate CD44 adhesiveness under flow condition. *Structure* (Cambridge, MA, US). 2010; 18(5): 649–56.
73. Jamison FW 2nd, Foster TJ, Barker JA, Hills RD Jr, Guvench O. Mechanism of binding site conformational switching in the CD44-hyaluronan protein-carbohydrate binding interaction. *J Mol Biol*. 2011; 406(4):631–47. [PubMed: 21216252]
74. Jackson DG, Bell JI, Dickinson R, Timans J, Shields J, Whittle N. Proteoglycan forms of the lymphocyte homing receptor CD44 are alternatively spliced variants containing the v3 exon. *J Cell Biol*. 1995; 128(4):673–685. [PubMed: 7532175]
75. Mikami T, Kitagawa H. Biosynthesis and function of chondroitin sulfate. *Biochim Biophys Acta*. 2013; 1830(10):4719–4733. [PubMed: 23774590]
76. Farber SJ, Schubert M, Schuster N. The binding of cations by chondroitin sulfate. *J Clin Invest*. 1957; 36(12):1715–22. [PubMed: 13491703]
77. Horkay F, Bassar PJ, Hecht AM, Geissler E. Chondroitin Sulfate in Solution: Effects of Mono- and Divalent Salts. *Macromolecules*. 2012; 45(6):2882–2890. [PubMed: 23814316]
78. Hunter GK. Chondroitin sulfate-derivatized agarose beads: a new system for studying cation binding to glycosaminoglycans. *Anal Biochem*. 1987; 165(2):435–41. [PubMed: 3122597]
79. Hunter GK, Wong KS, Kim JJ. Binding of calcium to glycosaminoglycans: An equilibrium dialysis study. *Arch Biochem Biophys*. 1988; 260(1):161–167. [PubMed: 3124744]
80. Uchisawa H, Okuzaki Bi, Ichita J, Matsue H. Binding between calcium ions and chondroitin sulfate chains of salmon nasal cartilage glycosaminoglycan. *International Congress Series*. 2001; 1223(0):205–220.

Abbreviations

HA	hyaluronan
HABD	hyaluronan-binding domain
MD	molecular dynamics
ABF	adaptive biasing force

QIDL**NIT**CRFAGVFHVEKNGRYSISRTEAADLCKAF**NST**LPTMAQMEKALSIGFETCRY
GFIEGHVVIPRIHPNSICAA**NNT**GVYILTS**NTS**QYDTYCF**NAS**APPEEDCTSVTDLPNA
FDGPITITIVNRDGTRYVQKGEYRTNPEDIYPSNPTDDDVSSG

Figure 1.

Amino-acid sequence of human CD44 HABD, using single letter amino acid abbreviations, with N-glycosylation consensus sequences in bold. N-glycosylation sites used in the MD simulations are additionally underlined.

Figure 2. Opposite faces of human CD44 HABD (PDB ID: 1UUH). All potential N-glycosylation sites are shown in red, and basic residues Arg and Lys in blue. Red arrows indicate the glycosylation sites used in this study, and the blue arrows indicate the critical HA-binding residue Arg41 and Arg154. The left panel is the HA-binding face.

Glycan type	CD44 HABD Glycosylation site
<p>Terminal sialic acid</p>	<p>1. Asn25</p> <p>2. Asn120</p> <p>3. Asn25 and Asn120</p>
<p>Terminal galactose</p>	<p>1. Asn25</p> <p>2. Asn120</p> <p>3. Asn25 and Asn120</p>

Figure 3.

Cartoon representations (using CFG symbol nomenclature) for sialic acid-terminal and asialo-glycans and the attachment sites used for system construction. Key: purple diamond for sialic acid, yellow circle for galactose, blue square for N-acetylglucosamine, green circle for mannose, red triangle for fucose, “N” on right indicates the asparagine residue glycan attachment site.

Figure 4. HABD with sialic acid capped glycans at Asn25 and Asn120. HABD is shown as a molecular surface, with Asn 25 in ball and stick representation. Covalently attached glycans are represented as van der Waals spheres and are colored according to the scheme in Fig. 3.

Figure 5. Distances between sialic acid carboxylate oxygens and Arg/Lys polar side chain hydrogens: time series plots.

Figure 6. Snapshot from the Fig. 5, top panel (green (+) plot) trajectory illustrating two hydrogen bonds between sialic acid carboxyl oxygens (red) and two guanidinium hydrogens (white) from the Arg41 sidechain. Panel A shows the CD44 backbone in cartoon representation, with Arg41 and Asn25 as van der Waals spheres, and glycan (covalently linked to Asn25) in ball and stick representation. Panel B is a close-up view of the Arg41 sidechain and sialic acid interaction.

Figure 7. Potential of mean force for Arg41 association with sialic acid. Triplicate data are from ABF trajectories after 10 ns (top) and 100 ns (bottom) of sampling; r is the distance between the sialic acid carboxylate carbon and the Arg41 guanidinium carbon.

Figure 8. Snapshot from the Fig. 5, bottom panel (red (+) plot) trajectory illustrating two hydrogen bonds between sialic acid carboxyl oxygens (red) and two guanidinium hydrogens (white) of the Arg154 sidechain. Panel A shows the CD44 backbone in cartoon representation, Arg41 and Arg154 sidechains in line representation, and glycan (covalently linked to Asn25 (ribbon in red)) in ball and stick representation. Panel B is a close-up view of the Arg154 sidechain and sialic acid interaction.

Table I

Constructed systems for molecular dynamics simulations.

	System	Glycan position	Terminal monosaccharide
1uuh “ordered”	Control	No glycan	-
	S25	Asn25	Sialic acid
	A25		β -galactose
	S120	Asn120	Sialic Acid
	A120		β -galactose
	S25S120	Asn25 & Asn120	Sialic acid
	A25A120		β -galactose
2i83 “disordered”	Control	No glycan	-
	S25	Asn25	Sialic acid
	A25		β -galactose
	S120	Asn120	Sialic Acid
	A120		β -galactose
	S25S120	Asn25 & Asn120	Sialic acid
	A25A120		β -galactose

Table II

Simulations with long-term associations of less than 3Å between sialic acid on glycan on Asn25 and Arg residues.

Glycoprotein Structure	Residue	% of simulation
Ordered S25	Arg41	49
Disordered S25S120	Arg154	38
Disordered S25S120	Arg41	52
Disordered S25S120	Arg154	19