

Revealing the mechanisms of protein disorder and N-glycosylation in CD44-hyaluronan binding using molecular simulation

*Olgun Guvench**

Department of Pharmaceutical Sciences, University of New England College of Pharmacy, Portland, ME, USA

OPEN ACCESS

Edited by:

David Naor,
Hebrew University of Jerusalem,
Israel

Reviewed by:

Barry C. Finzel,
University of Minnesota, USA
Ichio Shimada,
The University of Tokyo, Japan

*Correspondence:

Olgun Guvench,
Department of Pharmaceutical
Sciences, University of New England
College of Pharmacy, 716 Stevens
Avenue, Portland, ME 04103, USA
oguvench@une.edu

Specialty section:

This article was submitted to
Inflammation, a section of the journal
Frontiers in Immunology

Received: 15 February 2015

Accepted: 28 May 2015

Published: 16 June 2015

Citation:

Guvench O (2015) Revealing the mechanisms of protein disorder and N-glycosylation in CD44-hyaluronan binding using molecular simulation. *Front. Immunol.* 6:305. doi: 10.3389/fimmu.2015.00305

The extracellular N-terminal hyaluronan binding domain (HABD) of CD44 is a small globular domain that confers hyaluronan (HA) binding functionality to this large transmembrane glycoprotein. When recombinantly expressed by itself, HABD exists as a globular water-soluble protein that retains the capacity to bind HA. This has enabled atomic-resolution structural biology experiments that have revealed the structure of HABD and its binding mode with oligomeric HA. Such experiments have also pointed to an order-to-disorder transition in HABD that is associated with HA binding. However, it had remained unclear how this structural transition was involved in binding since it occurs in a region of HABD distant from the HA-binding site. Furthermore, HABD is known to be N-glycosylated, and such glycosylation can diminish HA binding when the associated N-glycans are capped with sialic acid residues. The intrinsic flexibility of disordered proteins and of N-glycans makes it difficult to apply experimental structural biology approaches to probe the molecular mechanisms of how the order-to-disorder transition and N-glycosylation can modulate HA binding by HABD. We review recent results from molecular dynamics simulations that provide atomic-resolution mechanistic understanding of such modulation to help bridge gaps between existing experimental binding and structural biology data. Findings from these simulations include: Tyr42 may function as a molecular switch that converts the HA-binding site from a low affinity to a high affinity state; in the partially disordered form of HABD, basic amino acids in the C-terminal region can gain sufficient mobility to form direct contacts with bound HA to further stabilize binding; and terminal sialic acids on covalently attached N-glycans can form charge-paired hydrogen bonding interactions with basic amino acids that could otherwise bind to HA, thereby blocking HA binding to glycosylated CD44 HABD.

Keywords: CD44, hyaluronan, binding, free energy, molecular dynamics, glycosylation, inhibition, unfolding

Introduction

The structure of the cell surface protein CD44, from its N-terminus to its C-terminus, consists of a globular hyaluronan binding domain (HABD), a stalk domain, a single-pass transmembrane domain, and a cytoplasmic domain (1, 2). Amino acids located N-terminal to the transmembrane domain are on the extracellular side of the cell membrane, and amino acids located C-terminal to the

transmembrane domain are on the intracellular side (**Figure 1**). Post-translational modifications to CD44 include glycosylation of the extracellular portion (3–5), palmitoylation of amino acids immediately C-terminal to the transmembrane domain (6–9), and phosphorylation of the cytoplasmic domain (10–12). The already-complex structural biology of CD44 is further complicated by variable splicing of the RNA transcript of the *CD44* gene, which yields a variety of different patterns of amino acid insertion into the stalk domain and which modulates CD44 function (1, 13, 14), and by shedding that produces soluble CD44 (15).

Atomic-resolution structures can lead to substantial insight into the function of a biomolecule. Such high-resolution structural information is typically obtained from X-ray crystallography or NMR spectroscopy experiments, with examples being as small as a single zinc finger domain (16) and as large as ribosomes (17) and virus capsids (18). However, not all biomolecules are equally amenable to having their structures solved by these methods. Glycoproteins and proteoglycans are particularly challenging because of the difficulty in obtaining pure samples and the inherent flexibility of these two types of biomolecules. Sample homogeneity is a challenge for glycoproteins and proteoglycans because the carbohydrate component results from non-template-based enzymatic synthesis, leading to carbohydrate microheterogeneity at a given glycosylation site even though the protein component for a given

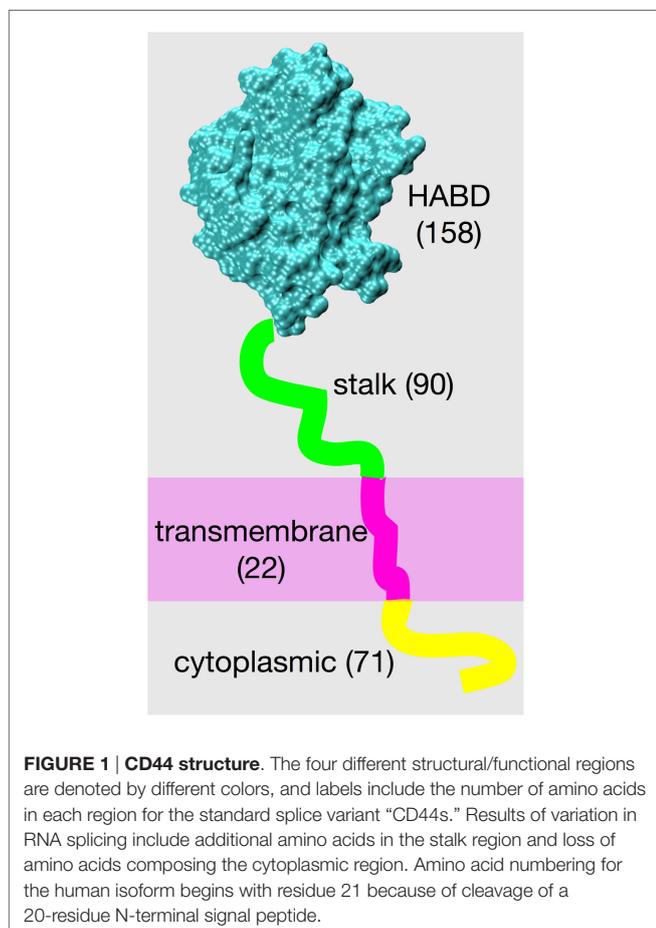
sample is identical throughout (19). The carbohydrate component is also flexible (20, 21), especially in comparison to globular proteins with their well-defined tertiary (three-dimensional) structures, and which therefore were crystallized early in the development of the field of structural biology (22) and still compose the majority of publicly available experimental atomic-resolution structures.

In contrast to globular proteins, which exist in an aqueous environment, transmembrane proteins are located in biological lipid bilayers. Therefore, this environment must be suitably reproduced in samples in order to do experimental structural biology, which can be very challenging (23, 24). Additionally, heterologous expression of transmembrane proteins and subsequent purification can be more difficult than for water-soluble globular proteins because of the limited surface area of the cell membrane for expression, the resulting toxicity to the organism being used for expression, and the subsequent need to reconstitute the protein in a lipid environment after extraction and purification (25, 26).

Intrinsically disordered proteins provide another contrast to globular proteins in that the former lack well-defined unique stable three-dimensional structures (27–29). In X-ray experiments, this results in crystallographic disorder, diffuse scattering, and therefore undefined atomic coordinates (30). While NMR has been used extensively to study intrinsically disordered proteins, solution NMR experiments yield data that represent ensemble averages, which can limit understanding of the various discrete conformations that such proteins may assume (31).

CD44, with its multiple domains, poses a number of challenges for characterization by X-ray crystallography or NMR spectroscopy. The N-terminal HABD, which is similar to globular proteins, is in its biologically relevant form a glycoprotein that has numerous glycosylation sites (1). Furthermore, HABD in the presence of hyaluronan (HA) has characteristics of an intrinsically disordered protein (32, 33). The stalk domain that connects the extracellular HABD to the cell membrane has both N- and O-glycosylation sites (34). And, with alternative splicing, the stalk domain can have proteoglycan characteristics, namely a protein core with glycosaminoglycan (GAG) attachments in the form of chondroitin sulfate and heparan sulfate (35). Recalling that CD44 is a transmembrane protein, in addition to the challenge associated with being located in the cell membrane, the transmembrane domain can be post-translationally modified by the attachment of lipids, which further complicates its structural biology since this modification can alter its interactions with the membrane bilayer (9, 36, 37). Finally, the cytoplasmic domain is likely disordered when not non-covalently bound to intracellular adapter proteins (12).

Atomic-resolution experimental structural biology on CD44 has been largely limited to the ~150 amino acid HABD in its non-glycosylated form because of the many specific challenges above. X-ray structures for HABD exist for both human and mouse isoforms. Human HABD has been reproducibly crystallized in its apo form (i.e., not bound to HA, inhibitors, or peptides) (38, 39), as well as with unidentified peptide (39) found to be non-covalently bound to the face of HABD opposite that of the location of the HA-binding site. However, there are no publicly available X-ray structures of human HABD complexed with either HA or with inhibitors of HA binding. In contrast, mouse HABD has been co-crystallized with both oligomeric HA (oHA) (40, 41) and



small-molecule inhibitors (40), as well as in its apo form (41). These co-crystals reveal the binding mode of HA with HABD, which is presumably the same for human HABD given the ~85% sequence identity between the two forms, and the 100% sequence identity of the HA-binding site (42). NMR structures exist for human HABD, both in its apo form (38) and bound to oHA (32). However, unlike X-ray structures with HA, NMR structural characterization of the bound form lacks atomic coordinates for HA and therefore does not provide comprehensive information into the non-covalent atomic contacts between HABD and bound HA. All of these previous structural biology examples are of CD44 HABD; the only non-HABD example of experimental CD44 structure is a complex consisting of a nine amino acid long portion of the 72 amino acid cytoplasmic domain complexed with the radixin FERM domain (12).

Molecular dynamics (MD) is a physics-based approach to the modeling and simulation of biomolecules (43). In all-atom explicit-solvent MD simulations, all the atoms of the system, including those for the solvent, are included as interaction sites for computing the forces in the system. The values of the forces as a function of atomic positions are determined by a combination of a mathematical expression and parameters, commonly called a “force field,” that encodes properties such as the energetic cost to stretch a bond or the energetic benefit of a van der Waals or charge-pairing interaction (44). These forces are numerically integrated to propagate the system, and this is done in an iterative manner to generate a trajectory, analogous to a movie, that shows how the positions of the atoms in the system evolve with the passing of time (45). Typical present-day simulations involve tens of thousands to hundreds of thousands of atoms with trajectory lengths of tens to hundreds of nanoseconds, which require tens to hundreds of millions of consecutive integration steps. MD simulations can be used to determine not only the time-evolution (dynamics) of the system but also the relative probabilities, and therefore free energies, associated with different states (thermodynamics) (46). As such, MD is an especially useful tool for studying flexible biomolecules at an atomic level of resolution, which makes it an ideal complement to experimental structural biology techniques like X-ray crystallography and NMR spectroscopy (47).

Over the past several years, our research group has applied all-atom explicit-solvent MD simulations to extend the understanding of the function of the CD44 HABD. These efforts have aimed to address the following scientific questions: (1) what is the mechanism and associated thermodynamics of a conformational change in an arginine-containing loop at the HABD binding site that is associated with HA binding? (41); (2) why does HABD transition from a well-ordered (folded) three-dimensional structure to one that is partially disordered when it binds to HA? (32, 33, 48); and (3) why do covalently attached sialylated N-glycans inhibit HA binding while unsialylated ones do not? (49–51) This article reviews the contribution MD simulations have made toward developing answers to these questions. We note that others have also recently applied MD simulation to the study of CD44 HABD, with topics including conformational flexibility and the microscopic structure and dynamics of water surrounding HABD (52, 53), and the mechanism and thermodynamics of the ordered to partially disordered transition (54).

Conformational Switching at the CD44 HABD Binding Site

Two lines of experimental evidence lead to the hypothesis that direct contact between CD44 Arg 41 and HA is a major source of binding affinity (“Arg41” reflects amino acid numbering in the human form of CD44; the equivalent amino acid is Arg45 in the mouse form. For simplicity, the human numbering will be used throughout this text.). From X-ray crystallography, Arg41 and the loop that contains it change conformation depending upon binding of oHA. In the apo form, the loop is in an open conformation that locates this sidechain too far away to contact oHA if it were present, and HABD is said to be in the “A” state (38, 41). In the complexed form, the loop can be either in the open conformation, or in a closed conformation that facilitates direct contact between the positively charged guanidinium group of the Arg41 sidechain and bound oHA, in which case HABD is in the “B” state (40, 41) (**Figure 2**, bottom row). From mutation data, the Arg41Ala single point mutation essentially destroys the ability of HABD to associate with oHA (41, 55, 56). This thermodynamic information demonstrates the critical nature of Arg41 in binding. However, the X-ray crystallographic data do not say anything about the thermodynamics of the A and B states, other than they are both sufficiently stable to be trapped as crystals when oHA is bound. And the mutation data do not provide information about the conformation of Arg41 when oHA is bound. We therefore applied MD simulations in an attempt to tie these two previous findings together.

All-atom explicit-solvent MD simulations can be used to determine true thermodynamic free energies for a variety of biomolecular processes (46, 57–60). This is because of two reasons: (1) a model system under study includes water molecules and allows for full conformational flexibility of the included biomolecules, which means both solvent effects and entropic effects are explicitly included (that is, not as an approximation, but as part of the system under investigation); and (2) there exist exact mathematical expressions to determine thermodynamic quantities directly from simulation data. One approach to determining the free-energy difference between two conformational states x and y of a biomolecule from MD simulation is to compute the reversible work required to transition the system from state x to state y by integrating the measured average force along the transition path (61). Since free energy is a thermodynamic state function, it does not depend on the actual path used to convert the system from state x to state y . However, in practice obtaining good numerical convergence as well as plausible biological insight both depend upon determining a physically reasonable transition path.

For CD44 HABD, the two states are A, having an open loop and the Arg41 sidechain separated from bound oHA, and B, with a closed loop and the Arg41 sidechain in direct contact with oHA. While a simple distance between Arg41 and oHA can be used to discriminate between different conformations of the sidechain, it is not immediately obvious that a similarly simple metric can be used to discriminate between the open vs. closed loop conformations. In an effort to identify such a metric, we first compared the conformations of the backbone dihedral angles ϕ , ψ for Arg41 and three amino acids on either side: -Lys-Asn-Gly-Arg41-Tyr-Ser-Ile-. This

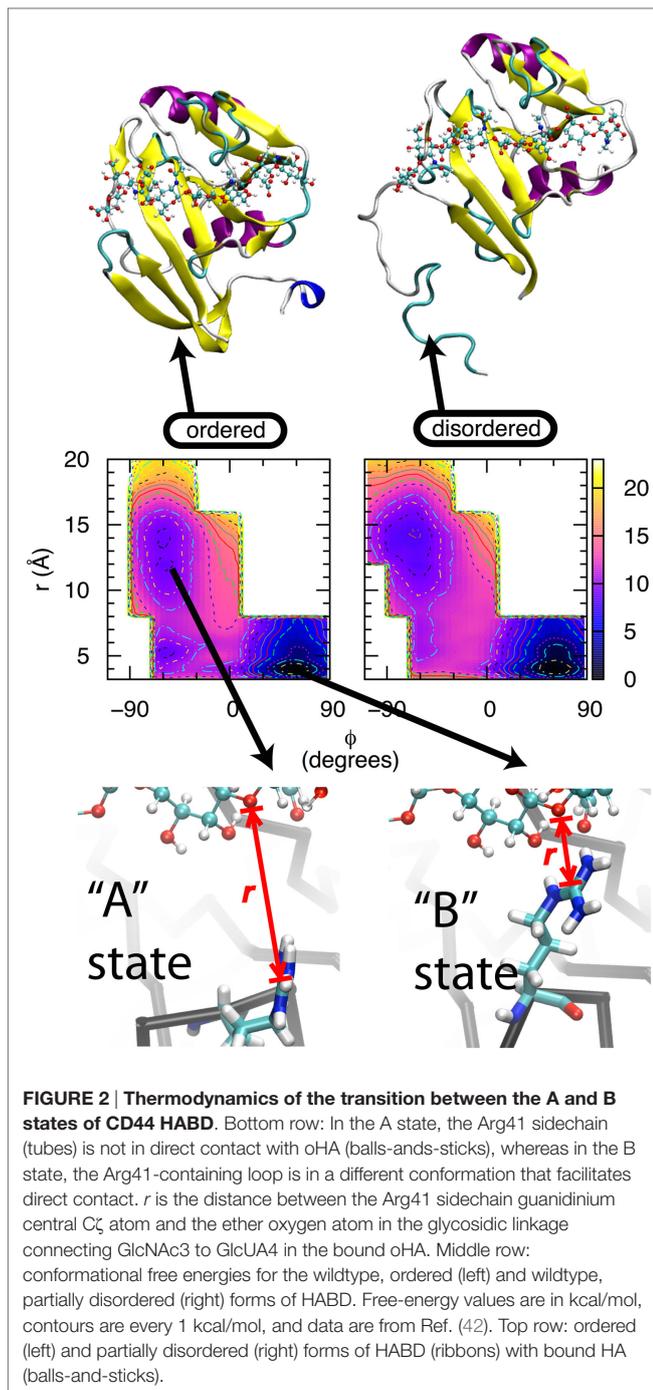


FIGURE 2 | Thermodynamics of the transition between the A and B states of CD44 HABD. Bottom row: In the A state, the Arg41 sidechain (tubes) is not in direct contact with oHA (balls-and-sticks), whereas in the B state, the Arg41-containing loop is in a different conformation that facilitates direct contact. r is the distance between the Arg41 sidechain guanidinium central C ζ atom and the ether oxygen atom in the glycosidic linkage connecting GlcNAc3 to GlcUA4 in the bound oHA. Middle row: conformational free energies for the wildtype, ordered (left) and wildtype, partially disordered (right) forms of HABD. Free-energy values are in kcal/mol, contours are every 1 kcal/mol, and data are from Ref. (42). Top row: ordered (left) and partially disordered (right) forms of HABD (ribbons) with bound HA (balls-and-sticks).

revealed a difference in the Tyr42 ϕ dihedral where in the A state $\phi = -60^\circ$ and in the B state $\phi = +60^\circ$ (62). This was followed by MD simulation where force was applied to this dihedral angle to gradually convert it from one value of ϕ to the other. In one case, starting from the A state and increasing ϕ converted the system to the B state, not just with respect to the Tyr42 backbone but also with respect to the separation distance between Arg41 and oHA. In the other case, starting from the B state and decreasing ϕ converted the system to the A state, including breaking of the

contact between Arg41 and oHA. In neither case was any force applied to directly affect the Arg41 to oHA distance; rather, changes in this distance spontaneously resulted from changes in the loop backbone conformation. Furthermore, a similar effect could not be achieved using the backbones of other amino acids in the loop (62). Thus, the reaction path for interconversion between the A and B states was defined in terms of two progress variables: the value of the Tyr42 backbone dihedral angle ϕ , and the distance r between the Arg41 sidechain guanidinium central C ζ atom and the ether oxygen atom in the glycosidic linkage connecting GlcNAc3 to GlcUA4 in the bound oHA.

Extensive simulations were subsequently done to compute the free energy of the system as a function of the progress variables ϕ and r (42). All simulations were of the human HABD (hHABD) complexed with [-4GlcUA β 1-3GlcNAc β 1-] $_4$ ("HA8"), which was computationally constructed by combining information from mouse and human structures. The four hHABD-HA8 systems considered were: wildtype, ordered; wildtype, partially disordered; Arg41Ala, ordered; and Arg41Ala, partially disordered. **Figure 2**, middle row shows the data for both wildtype systems, and the major free-energy minima, which are the most stable states for each system, correspond to A and B.

Wildtype simulation data demonstrate that the B state of the hHABD-HA8 complex is more stable than the A state by ~ 8 – 9 kcal/mol. This is true for both the ordered and partially disordered forms of hHABD (**Figure 2**, top row). Additional simulation data demonstrate that the analogous transition for the Arg41Ala mutant is substantially less favorable at ~ 6 kcal/mol. From these data, it is possible to calculate the loss in binding affinity associated with the point mutation, with values of 2.2 kcal/mol for the ordered form and 2.3 kcal/mol for the partially disordered form (42). These simulation data are in close agreement with existing experimental data measuring the loss in binding affinity to be 2.5 kcal/mol (41), which helps validate both the force field and the convergence of the simulations. Taken together, these results support the idea that formation of direct contact between HA and the Arg41 sidechain is a substantial source of favorable binding free energy (41).

In contrast to some of these findings, Plazinski and Knys-Dziciuch, in their simulation studies, found that the ϕ -related free-energy barrier was not correlated with the Arg41-HA distance (54, 63). The authors also observed a low free-energy barrier associated with separation of Arg41 from HA, with the A and B states reducing to an average dynamic structure (54, 63). They suggested that a possible explanation might be differences in the force fields used in their studies (54), since the previous analogous work found the A and B states to be distinct (62). Differences in force fields, which are the underlying physical models used to represent the bonded and non-bonded interactions in such simulations, can indeed cause such differing results. When such differences are inferred based on differing simulation results and conclusions, one possibility is to review the methodology involved in the force field development (44, 64). Another is to compare the outcomes of the particular simulations with the existing experimental data for inconsistencies. In this particular case, relevant experimental data include the crystal structures of the A and B states of HABD complexed with oHA, and the mutation data showing loss of binding affinity

in the Arg41Ala mutant. Dynamic averaging would manifest as crystallographic disorder with poorly resolved electron density for Arg41, which is in contrast to the existing crystallographic data. Additionally, rapid equilibrium between short and long separation distances between Arg41 and oHA suggests a weak interaction between them, which is in contrast to the mutation data.

Binding and Unfolding in CD44 HABD

Early experimental studies of the Arg41Ala mutation that predated the CD44 structural studies also probed the contribution of other basic amino acids by both point and truncation mutations (55). This was a logical course of action in the absence of HABD three-dimensional structure since HA contains a negatively charged carboxylate group every other monosaccharide and Arg and Lys sidechains are positively charged, suggesting the possibility of charge-charge interaction as a mechanism of binding. From that perspective, it is not surprising that these mutations all decrease the strength of binding of HA oligomers. But, taken in the context of the subsequent structural information, the explanation is less obvious, since these additional amino acids, in contrast to Arg41, are located spatially far from the now-known binding site (41). Further complicating the situation is the observation that the portion of HABD that contains these amino acids 153-169 goes from having well-defined three-dimensional structure to becoming unfolded in conjunction with HA binding (32, 33, 48). That is, the change in the conformation, which correlates with the binding of HA, in this span of amino acids located in the C-terminal most region of HABD is what defines the ordered-to-partially disordered HABD transition.

The above suggests two questions: why do basic amino acids distant from the binding site affect affinity? And why is the affinity greater when the sequence containing these amino acids unfolds and becomes disordered? To answer these two questions, we return to the previous set of free-energy data from MD simulations. For wildtype hHABD, the free-energy data associated with the interconversion between the A and B states are qualitatively the same regardless of whether hHABD is in the ordered or partially disordered form (Figure 2, middle row). The same is also the case for the Arg41Ala mutant, where the free-energy change in going from A to B is independent of whether those distant amino acids are folded or not (42) (data not shown here). Not only are there strong qualitative similarities in free-energy data between the ordered and partially disordered forms, but the quantitative values are also very similar. For the A→B transition of the binding site in the wildtype, ordered form, the associated free-energy change is -8.7 kcal/mol, and for this transition in the wildtype, partially disordered form the value is a very similar -7.8 kcal/mol (42). Likewise for the Arg41Ala mutant, in the ordered form the value is -6.5 kcal/mol and in the partially disordered form it is -5.5 kcal/mol (42). The small difference of ~ 1 kcal/mol is within the precision that can be expected from these particular computational experiments.

In an allosteric mechanism, a conformational change distant from the site affects the energetics at the binding site. In the case of HABD, the independence of the energetics of the binding site A→B transition from the ordered vs. partially disordered form of HABD contradicts the hypothesis that allostery is at work. That is,

conformational switching at the binding site appears no more or less favorable if the HABD C-terminal region is folded or unfolded.

Analysis of the MD trajectories that were generated as part of the free-energy experiments yielded a result that, in hindsight, is obvious: flexibility from partial unfolding permits favorable electrostatic interactions between HA and the C-terminal HABD amino acids that cannot occur when the domain is fully ordered. In the ordered form of HABD, the amino acids in question are locked into a folded conformation that keeps them far from bound oHA, while in the partially disordered form, this is no longer the case (Figure 2, top row). Because this span of amino acids is no longer in consistent contact with the rest of the HABD domain, it assumes the properties of a random coil peptide, which through random fluctuations can collide with bound oHA. If this collision happens in a way that brings one of the basic sidechains into close proximity with oHA, a favorable contact can be formed. Unlike the Arg41 interaction, which has a well-defined mechanism based on specific interactions of the sidechain with a particular limited section of bound oHA, the flexibility of the disordered amino acids and the repeating nature of the HA polymer permit the possibility of a wide range of basic sidechain interactions with bound HA (42), consistent with the long-standing finding that a 13-amino acid CD44 peptide spanning Arg150–Arg162 itself will bind HA (65). This wide range may conceivably include interaction with HA bound to an adjacent molecule of CD44, though the present simulations had a single copy of the complex and therefore could not directly address the possibility of such *trans* association.

The above computational experiments do suggest a plausible molecular mechanism by which the ordered to partially disordered transition confers increased binding affinity to oHA. However, the simulations involved either the ordered or the partially disordered form of CD44 hHABD, and therefore do not provide any insight into the mechanism of the ordered to partially disordered transition itself. Independent work has been done toward this end, and with the additional aims of estimating the free-energy profile of the transition and clarifying the role of select amino acids in the transition (54). Connected with the proposed transition mechanism was a free-energy change of $+25$ kcal/mol, implying the partially disordered form in the absence of HA is very unstable relative to the ordered form. While the sign of the free-energy change, associated with loss of a single beta strand at the edge of a beta sheet, agrees with experiment (32), the magnitude is substantially larger than the folding free-energy values for entire small single-domain proteins, which are typically <10 kcal/mol, including for proteins consisting exclusively of beta strands (66). A further difficulty is that the simulation free-energy data for the analogous transition in the Tyr161Ala mutant are identical to those for the wild type. This is in contrast to experimental data, where the Tyr161Ala mutant constitutively exhibits the partially disordered conformation (33, 48). One possible explanation for these apparent inconsistencies is that the study represents the initial steps of the transition (54), such that extending simulations further along the selected reaction coordinate may result in subsequent decreases in free energy. Another possibility is lack of convergence of simulations given the large scale of the transition (54). This was the case for the free energy for the A→B transition, which involves a much smaller conformational change than the ordered to partially disordered

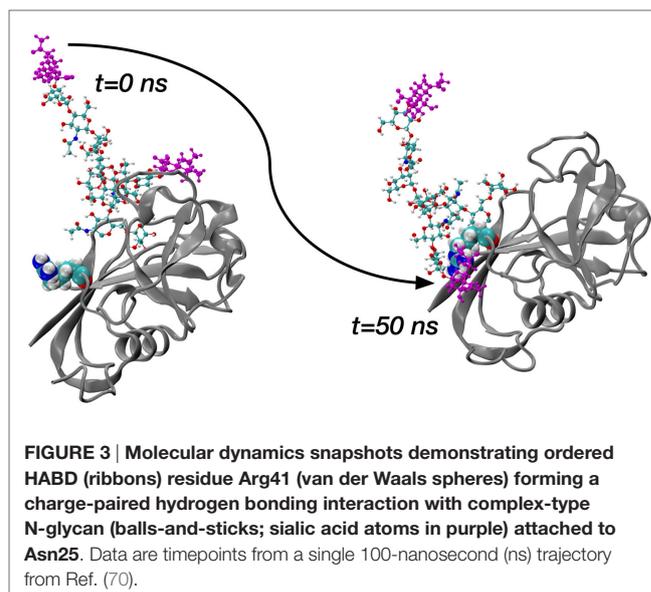
transition. The first study suggested that the A and the B states were essentially equally stable (62), while subsequent work that extended the timescale of the simulations by 40-fold showed the B state to be substantially more stable than the A (42). Importantly, in the case of the ordered to partially disordered transition, the partially disordered form is not a single, well-defined conformation. Rather, the disordered C-terminal HABD amino acids are free to take on a multitude of conformations. Therefore the partially disordered form is actually an ensemble of diverse conformations, and this further complicates computational experiments toward understanding the transition mechanism.

Inhibition by Glycosylation

N-glycosylation of CD44 HABD is known to have variable effects on CD44 function depending on the nature of the N-glycans. One effect is to block HABD binding to HA (49–51). Another is to make CD44 itself a ligand that binds to lectins (67, 68). The biochemistry behind both of these contrasting functions is related and is modulated by *N*-acetylneuraminic acid (Neu5Ac) monosaccharide, which is commonly called “sialic acid.” In one case, HABD N-glycans are capped with sialic acids, and this both blocks HA binding and makes CD44 a selectin ligand. In the other case, sialidase activity removes these terminal sialic acids, leaving the bulk of the N-glycan structures intact, and this change restores both HA-binding and removes selectin ligand functionality. While sialidase treatment removes only the terminal monosaccharide from the attached N-glycan, the functional result is the same as removal of the entire N-glycan. For example, Asn point mutation precludes N-glycan attachment and heterologous expression yields non-glycosylated protein. In both of these cases, the functional outcome is the same as sialidase treatment, which implies that inhibition of HA-binding cannot be explained as a consequence of steric blockage of the HABD binding site, since the de-sialylated N-glycan has nearly the same bulk as the sialo-glycan. This mechanism of regulation is not unique to CD44, as the related hyaladherin LYVE-1 demonstrates similar behavior (69).

The sialidase data immediately suggest two sets of simulations of glycosylated HABD: one set with sialylated N-glycan and a second set with asialo N-glycan. Asn25 and Asn120 were selected for computational N-glycosylation based on previous mutagenesis studies showing that cells expressing Asn25Ser and Asn120Ser mutants constitutively bind HA (50). Complex-type N-glycans that were selected (70) based on the existing finding that blocking the metabolic pathway for processing complex N-glycans restores HA binding (71, 72). In conjunction with the two different glycosylation sites and the sialo and asialo forms of the N-glycan, both the ordered and partially disordered forms of CD44 HABD were studied in the simulations. A representative starting conformation of ordered HABD with a sialo glycan attached to Asn25 is shown in the left frame of **Figure 3**; the analogous asialo form would be missing only the atoms colored purple.

The key finding from the comprehensive set of simulations covering the ordered and disordered HABD paired with sialo and asialo N-glycans is that only in the sialo form do stable, long-lasting non-covalent contacts form between the protein and glycan components. Furthermore, these contacts involve



the positively charged sidechains of HABD amino acids and the negatively charged carboxylate groups of the terminal sialic acids (**Figure 3**). In contrast, asialo glycans form only brief contacts, which is understandable since they lack the negative charge of the sialo form. Long-lasting contacts in the sialoglycan simulations involve Arg41 and Arg154, and these contacts form spontaneously during the simulation and can last for 40–50% of the simulation length (70). Both of these amino acids can directly associate with HA8 when it is bound, based on findings from the computational experiments on partially disordered HABD, as summarized in the previous section. However, their binding with sialic acid is an interaction that would directly compete with their binding with HA8. Therefore, the view that emerges is that, in the sialo form, the covalently attached N-glycans will form charge-paired hydrogen bonding interactions with Arg sidechains known to be important for HA binding. As further evidence, free-energy simulations, similar to the ones for Arg41–HA8 association described in the previous section, demonstrate that the Arg–sialic acid association is indeed thermodynamically favorable by ~1 kcal/mol (70). We do note that these simulations were limited to only the CD44 HABD, which is present in all splice variants of CD44, and that the simulations did not include HA. Clearly, additional work needs to be done to understand the molecular mechanisms by which glycosylation alters binding, since it has been shown that N-glycosylation of CD44 can also facilitate HA binding (73).

Conclusion

A subset of the computational experiments above suggests the following four conclusions. The first is that the Tyr42 backbone dihedral angle φ can act as a molecular switch to convert the HABD HA-binding site from the open A state to the closed B state, which includes the formation of direct contact between HA and the Arg41 sidechain (62). The second is that the B state is more thermodynamically stable, and this stability is due to direct

Arg41-HA contact (42). The third is that basic amino acids located distant from the HA-binding site in the ordered form of HABD gain sufficient mobility in the partially disordered form to be able to form direct contacts with oHA and further stabilize binding (42). And the fourth is that terminal sialic acids on covalently attached N-glycans can form charge-paired hydrogen bonding interactions with basic amino acids that could otherwise bind to HA, thereby blocking HA binding to glycosylated CD44 HABD (70). In addition to contributing to the mechanistic understanding of CD44-HA binding, these conclusions may be of utility in the future development of small-molecule modulators of CD44 function (40), especially given the potential for CD44 as a therapeutic target (74–76).

However, it should be kept in mind that the role of Tyr42 as a molecular switch, and the discrete nature of the A and B states of the Arg41-containing loop is contradicted by other computational work (54, 63). Furthermore, there do exist experimental data that are in apparent conflict with the above conclusions. As mentioned previously, N-glycosylation of CD44 can facilitate HA binding (73). And, mutation of positively charged amino acids in the disordered region has been found to enhance HA-binding affinity in the context of both purified HABD and cell surface CD44 (41). Given these differences, further investigation is warranted to achieve a comprehensive consistent view. Finally, while outside the scope of this review, there have been substantial efforts using MD simulations to understand the importance of water molecules and of biomolecular conformational entropy changes in HABD binding with HA (52, 53). Findings from these simulations that can inform development of small-molecule modulators of CD44 function include reduced translational and rotational freedom of water molecules in contact with HABD and HA, and loss of HA flexibility associated with binding to HABD.

References

- Ponta H, Sherman L, Herrlich PA. CD44: from adhesion molecules to signalling regulators. *Nat Rev Mol Cell Biol* (2003) 4(1):33–45. doi:10.1038/nrm1004
- Naor D, Wallach-Dayana SB, Zahalka MA, Sionov RV. Involvement of CD44, a molecule with a thousand faces, in cancer dissemination. *Semin Cancer Biol* (2008) 18(4):260–7. doi:10.1016/j.semcancer.2008.03.015
- Han H, Stapels M, Ying W, Yu Y, Tang L, Jia W, et al. Comprehensive characterization of the N-glycosylation status of CD44s by use of multiple mass spectrometry-based techniques. *Anal Bioanal Chem* (2012) 404(2):373–88. doi:10.1007/s00216-012-6167-4
- 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. doi:10.1016/S0955-0674(97)80116-0
- Naor D, Sionov RV, Ish-Shalom D. CD44: structure, function, and association with the malignant process. *Adv Cancer Res* (1997) 71:241–319. doi:10.1016/S0065-230X(08)60101-3
- Thorne RF, Legg JW, Isacke CM. The role of the CD44 transmembrane and cytoplasmic domains in co-ordinating adhesive and signalling events. *J Cell Sci* (2004) 117(Pt 3):373–80. doi:10.1242/jcs.00954
- Bourguignon LY, Kalamiris EL, Lokeshwar VB. Acylation of the lymphoma transmembrane glycoprotein, GP85, may be required for GP85-ankyrin interaction. *J Biol Chem* (1991) 266(18):11761–5.
- Guo YJ, Lin SC, Wang JH, Bigby M, Sy MS. Palmitoylation of CD44 interferes with CD3-mediated signaling in human T lymphocytes. *Int Immunol* (1994) 6(2):213–21. doi:10.1093/intimm/6.2.213
- Babina IS, McSherry EA, Donatello S, Hill AD, Hopkins AM. A novel mechanism of regulating breast cancer cell migration via palmitoylation-dependent alterations in the lipid raft affiliation of CD44. *Breast Cancer Res* (2014) 16(1):R19. doi:10.1186/bcr3614
- 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. doi:10.1002/eji.1830250714
- Legg JW, Lewis CA, Parsons M, Ng T, Isacke CM. A novel PKC-regulated mechanism controls CD44 ezrin association and directional cell motility. *Nat Cell Biol* (2002) 4(6):399–407. doi:10.1038/ncb797
- Mori T, Kitano K, Terawaki S, Maesaki R, Fukami Y, Hakoshima T. Structural basis for CD44 recognition by ERM proteins. *J Biol Chem* (2008) 283(43):29602–12. doi:10.1074/jbc.M803606200
- Brown RL, Reinke LM, Damerow MS, Perez D, Chodosh LA, Yang J, et al. CD44 splice isoform switching in human and mouse epithelium is essential for epithelial-mesenchymal transition and breast cancer progression. *J Clin Invest* (2011) 121(3):1064–74. doi:10.1172/JCI44540
- 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):12160–4. doi:10.1073/pnas.89.24.12160
- Cichy J, Kulig P, Pure E. Regulation of the release and function of tumor cell-derived soluble CD44. *Biochim Biophys Acta* (2005) 1745(1):59–64. doi:10.1016/j.bbamer.2005.02.006
- Lee MS, Gippert GP, Soman KV, Case DA, Wright PE. Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* (1989) 245(4918):635–7. doi:10.1126/science.2503871
- Simonovic M, Steitz TA. A structural view on the mechanism of the ribosome-catalyzed peptide bond formation. *Biochim Biophys Acta* (2009) 1789(9–10):612–23. doi:10.1016/j.bbagr.2009.06.006

18. Rossmann MG. Structure of viruses: a short history. *Q Rev Biophys* (2013) **46**(2):133–80. doi:10.1017/S0033583513000012
19. An HJ, Froehlich JW, Lebrilla CB. Determination of glycosylation sites and site-specific heterogeneity in glycoproteins. *Curr Opin Chem Biol* (2009) **13**(4):421–6. doi:10.1016/j.cbpa.2009.07.022
20. Woods RJ, Tessier MB. Computational glycoscience: characterizing the spatial and temporal properties of glycans and glycan-protein complexes. *Curr Opin Struct Biol* (2010) **20**(5):575–83. doi:10.1016/j.sbi.2010.07.005
21. 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. doi:10.1371/journal.pcbi.1002946
22. Jaskolski M, Dauter Z, Wlodawer A. A brief history of macromolecular crystallography, illustrated by a family tree and its Nobel fruits. *FEBS J* (2014) **281**(18):3985–4009. doi:10.1111/febs.12796
23. Loll PJ. Membrane proteins, detergents and crystals: what is the state of the art? *Acta Crystallogr F Struct Biol Commun* (2014) **70**(Pt 12):1576–83. doi:10.1107/S2053230X14025035
24. Judge PJ, Taylor GF, Dannatt HR, Watts A. Solid-state nuclear magnetic resonance spectroscopy for membrane protein structure determination. *Methods Mol Biol* (2015) **1261**:331–47. doi:10.1007/978-1-4939-2230-7_17
25. Mus-Veteau I. Heterologous expression of membrane proteins for structural analysis. *Methods Mol Biol* (2010) **601**:1–16. doi:10.1007/978-1-60761-344-2_1
26. Lee JK, Stroud RM. Unlocking the eukaryotic membrane protein structural proteome. *Curr Opin Struct Biol* (2010) **20**(4):464–70. doi:10.1016/j.sbi.2010.05.004
27. Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys* (2008) **37**:215–46. doi:10.1146/annurev.biophys.37.032807.125924
28. Dyson HJ. Expanding the proteome: disordered and alternatively folded proteins. *Q Rev Biophys* (2011) **44**(4):467–518. doi:10.1017/S0033583511000060
29. Chen J. Towards the physical basis of how intrinsic disorder mediates protein function. *Arch Biochem Biophys* (2012) **524**(2):123–31. doi:10.1016/j.abb.2012.04.024
30. Welberry TR, Goossens DJ. Diffuse scattering and partial disorder in complex structures. *IUCr* (2014) **1**(Pt 6):550–62. doi:10.1107/S205225251402065X
31. Konrat R. NMR contributions to structural dynamics studies of intrinsically disordered proteins. *J Magn Reson* (2014) **241**:74–85. doi:10.1016/j.jmr.2013.11.011
32. Takeda M, Ogino S, Umemoto R, Sakakura M, Kajiwara M, Sugahara KN, et al. Ligand-induced structural changes of the CD44 hyaluronan-binding domain revealed by NMR. *J Biol Chem* (2006) **281**(52):40089–95. doi:10.1074/jbc.M608425200
33. Ogino S, Nishida N, Umemoto R, Suzuki M, Takeda M, Terasawa H, et al. Two-state conformations in the hyaluronan-binding domain regulate CD44 adhesiveness under flow condition. *Structure* (2010) **18**(5):649–56. doi:10.1016/j.str.2010.02.010
34. Zoller M. CD44: can a cancer-initiating cell profit from an abundantly expressed molecule? *Nat Rev Cancer* (2011) **11**(4):254–67. doi:10.1038/nrc3023
35. Jackson DG, Bell JJ, 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):673–85. doi:10.1083/jcb.128.4.673
36. Olausson BE, Grossfield A, Pitman MC, Brown MF, Feller SE, Vogel A. Molecular dynamics simulations reveal specific interactions of post-translational palmitoyl modifications with rhodopsin in membranes. *J Am Chem Soc* (2012) **134**(9):4324–31. doi:10.1021/ja2108382
37. 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):34601–9. doi:10.1074/jbc.M601530200
38. Teriete P, Banerji S, Noble M, Blundell CD, Wright AJ, Pickford AR, et al. Structure of the regulatory hyaluronan binding domain in the inflammatory leukocyte homing receptor CD44. *Mol Cell* (2004) **13**(4):483–96. doi:10.1016/S1097-2765(04)00080-2
39. Liu LK, Finzel B. High-resolution crystal structures of alternate forms of the human CD44 hyaluronan-binding domain reveal a site for protein interaction. *Acta Crystallogr F Struct Biol Commun* (2014) **70**(Pt 9):1155–61. doi:10.1107/S2053230X14015532
40. Liu LK, Finzel BC. Fragment-based identification of an inducible binding site on cell surface receptor CD44 for the design of protein-carbohydrate interaction inhibitors. *J Med Chem* (2014) **57**(6):2714–25. doi:10.1021/jm5000276
41. Banerji S, Wright AJ, Noble M, Mahoney DJ, Campbell ID, Day AJ, et al. Structures of the Cd44-hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction. *Nat Struct Mol Biol* (2007) **14**(3):234–9. doi:10.1038/nsmb1201
42. 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. doi:10.1016/j.bpj.2013.07.041
43. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol* (2002) **9**(9):646–52. doi:10.1038/nsb0902-646
44. Monticelli L, Tieleman DP. Force fields for classical molecular dynamics. *Methods Mol Biol* (2013) **924**:197–213. doi:10.1007/978-1-62703-017-5_8
45. Hug S. Classical molecular dynamics in a nutshell. *Methods Mol Biol* (2013) **924**:127–52. doi:10.1007/978-1-62703-017-5_6
46. Wereszczynski J, McCammon JA. Statistical mechanics and molecular dynamics in evaluating thermodynamic properties of biomolecular recognition. *Q Rev Biophys* (2012) **45**(1):1–25. doi:10.1017/S0033583511000096
47. Dror RO, Dirks RM, Grossman JP, Xu H, Shaw DE. Biomolecular simulation: a computational microscope for molecular biology. *Annu Rev Biophys* (2012) **41**:429–52. doi:10.1146/annurev-biophys-042910-155245
48. Osawa M, Takeuchi K, Ueda T, Nishida N, Shimada I. Functional dynamics of proteins revealed by solution NMR. *Curr Opin Struct Biol* (2012) **22**(5):660–9. doi:10.1016/j.sbi.2012.08.007
49. 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):431–46. doi:10.1083/jcb.140.2.431
50. English NM, Lesley JF, Hyman R. Site-specific de-N-glycosylation of CD44 can activate hyaluronan binding, and CD44 activation states show distinct threshold densities for hyaluronan binding. *Cancer Res* (1998) **58**(16):3736–42.
51. Katoh S, Zheng Z, Oritani K, Shimozato T, Kincaid PW. Glycosylation of CD44 negatively regulates its recognition of hyaluronan. *J Exp Med* (1995) **182**(2):419–29. doi:10.1084/jem.182.2.419
52. Jana M, Bandyopadhyay S. Restricted dynamics of water around a protein-carbohydrate complex: computer simulation studies. *J Chem Phys* (2012) **137**(5):055102. doi:10.1063/1.4739421
53. Jana M, Bandyopadhyay S. Conformational flexibility of a protein-carbohydrate complex and the structure and ordering of surrounding water. *Phys Chem Chem Phys* (2012) **14**(18):6628–38. doi:10.1039/c2cp24104h
54. Plazinski W, Knys-Dzieciuch A. The 'order-to-disorder' conformational transition in CD44 protein: an umbrella sampling analysis. *J Mol Graph Model* (2013) **45**:122–7. doi:10.1016/j.jmgm.2013.08.002
55. 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. doi:10.1083/jcb.122.1.257
56. Wallach-Dayana SB, Grabovsky V, Moll J, Sleeman J, Herrlich P, Alon R, et al. CD44-dependent lymphoma cell dissemination: a cell surface CD44 variant, rather than standard CD44, supports in vitro lymphoma cell rolling on hyaluronic acid substrate and its in vivo accumulation in the peripheral lymph nodes. *J Cell Sci* (2001) **114**(19):3463–77.
57. Guvench O, MacKerell AD Jr. Computational evaluation of protein-small molecule binding. *Curr Opin Struct Biol* (2009) **19**(1):56–61. doi:10.1016/j.sbi.2008.11.009
58. Knight JL, Brooks CL III. Lambda-dynamics free energy simulation methods. *J Comput Chem* (2009) **30**(11):1692–700. doi:10.1002/jcc.21295
59. Deng Y, Roux B. Computations of standard binding free energies with molecular dynamics simulations. *J Phys Chem B* (2009) **113**(8):2234–46. doi:10.1021/jp807701h
60. Bernardi RC, Melo MC, Schulten K. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim Biophys Acta* (2015) **1850**(5):872–77. doi:10.1016/j.bbagen.2014.10.019
61. Comer J, Gumbart JC, Henin J, Lelievre T, Pohorille A, Chipot C. The adaptive biasing force method: everything you always wanted to know but were afraid to ask. *J Phys Chem B* (2015) **119**(3):1129–51. doi:10.1021/jp506633n
62. Jamison FW II, 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. doi:10.1016/j.jmb.2010.12.040
63. Plazinski W, Knys-Dzieciuch A. Interactions between CD44 protein and hyaluronan: insights from the computational study. *Mol Biosyst* (2012) **8**(2):543–7. doi:10.1039/c2mb05399c

64. Guvench O, MacKerell AD Jr. Comparison of protein force fields for molecular dynamics simulations. *Methods Mol Biol* (2008) **443**:63–88. doi:10.1007/978-1-59745-177-2_4
65. Yang B, Yang BL, Savani RC, Turley EA. Identification of a common hyaluronan binding motif in the hyaluronan binding proteins RHAMM, CD44 and link protein. *EMBO J* (1994) **13**(2):286–96.
66. Jackson SE. How do small single-domain proteins fold? *Fold Des* (1998) **3**(4):R81–91. doi:10.1016/S1359-0278(98)00033-9
67. Dimitroff CJ, Lee JY, Fuhlbrigge RC, Sackstein R. A distinct glycoform of CD44 is an L-selectin ligand on human hematopoietic cells. *Proc Natl Acad Sci U S A* (2000) **97**(25):13841–6. doi:10.1073/pnas.250484797
68. 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. doi:10.1083/jcb.153.6.1277
69. 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. doi:10.1074/jbc.M805105200
70. Fallner CE, Guvench O. Terminal sialic acids on CD44 N-glycans can block hyaluronan binding by forming competing intramolecular contacts with arginine sidechains. *Proteins* (2014) **82**(11):3079–89. doi:10.1002/prot.24668
71. 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. doi:10.1084/jem.182.2.431
72. 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. doi:10.1083/jcb.132.6.1199
73. Girbl T, Hinterseer E, Grossinger EM, Asslaber D, Oberascher K, Weiss L, et al. CD40-mediated activation of chronic lymphocytic leukemia cells promotes their CD44-dependent adhesion to hyaluronan and restricts CCL21-induced motility. *Cancer Res* (2013) **73**(2):561–70. doi:10.1158/0008-5472.CAN-12-2749
74. Orian-Rousseau V. CD44, a therapeutic target for metastasising tumours. *Eur J Cancer* (2010) **46**(7):1271–7. doi:10.1016/j.ejca.2010.02.024
75. Orian-Rousseau V, Ponta H. Perspectives of CD44 targeting therapies. *Arch Toxicol* (2015) **89**(1):3–14. doi:10.1007/s00204-014-1424-2
76. Johnson P, Ruffell B. CD44 and its role in inflammation and inflammatory diseases. *Inflamm Allergy Drug Targets* (2009) **8**(3):208–20. doi:10.2174/187152809788680994
77. Towns J, Cockerill T, Dahan M, Foster I, Gaither K, Grimshaw A, et al. XSEDE: accelerating scientific discovery. *Comput Sci Eng* (2014) **16**(5):62–74. doi:10.1109/MCSE.2014.80
78. Brooks BR, Brooks CL III, MacKerell AD Jr, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. *J Comput Chem* (2009) **30**(10):1545–614. doi:10.1002/jcc.21287
79. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. *J Comput Chem* (2005) **26**(16):1781–802. doi:10.1002/jcc.20289
80. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, et al. The Amber biomolecular simulation programs. *J Comput Chem* (2005) **26**(16):1668–88. doi:10.1002/jcc.20290
81. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem* (2005) **26**(16):1701–18. doi:10.1002/jcc.20291
82. Lopes PE, Guvench O, MacKerell AD Jr. Current status of protein force fields for molecular dynamics simulations. *Methods Mol Biol* (2015) **1215**:47–71. doi:10.1007/978-1-4939-1465-4_3
83. Guvench O, MacKerell AD Jr. Quantum mechanical analysis of 1,2-ethanediol conformational energetics and hydrogen bonding. *J Phys Chem A* (2006) **110**(32):9934–9. doi:10.1021/jp0623241
84. Guvench O, Greene SN, Kamath G, Brady JW, Venable RM, Pastor RW, et al. Additive empirical force field for hexopyranose monosaccharides. *J Comput Chem* (2008) **29**(15):2543–64. doi:10.1002/jcc.21004
85. Guvench O, MacKerell AD Jr. Automated conformational energy fitting for force-field development. *J Mol Model* (2008) **14**(8):667–79. doi:10.1007/s00894-008-0305-0
86. 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):2353–70. doi:10.1021/ct900242e
87. Hatcher E, Guvench O, MacKerell AD Jr. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *J Phys Chem B* (2009) **113**(37):12466–76. doi:10.1021/jp905496e
88. Hatcher E, Guvench O, Mackerell AD Jr. CHARMM additive all-atom force field for acyclic polyalcohols, acyclic carbohydrates and inositol. *J Chem Theory Comput* (2009) **5**(5):1315–27. doi:10.1021/ct9000608
89. Raman EP, Guvench O, MacKerell AD Jr. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *J Phys Chem B* (2010) **114**(40):12981–94. doi:10.1021/jp105758h
90. Guvench O, Mallajosyula SS, Raman EP, Hatcher E, Vanommeslaeghe K, Foster TJ, et al. 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):3162–80. doi:10.1021/ct200328p
91. Mallajosyula SS, Guvench O, Hatcher E, MacKerell AD Jr. CHARMM additive all-atom force field for phosphate and sulfate linked to carbohydrates. *J Chem Theory Comput* (2012) **8**(2):759–76. doi:10.1021/ct200792v

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Guvench. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.