



Article

Constructing 3-Dimensional Atomic-Resolution Models of Nonsulfated Glycosaminoglycans with Arbitrary Lengths Using Conformations from Molecular Dynamics

Elizabeth K. Whitmore ^{1,2} , Devon Martin ^{1,2} and Olgun Guvench ^{1,2,*}

¹ Department of Pharmaceutical Sciences and Administration, University of New England School of Pharmacy, 716 Stevens Avenue, Portland, ME 04103, USA; ewhitmore@une.edu (E.K.W.); dmartin11@une.edu (D.M.)

² Graduate School of Biomedical Science and Engineering, University of Maine, 5775 Stodder Hall, Orono, ME 04469, USA

* Correspondence: oguvench@une.edu; Tel.: +1-207-221-4171

Received: 10 September 2020; Accepted: 15 October 2020; Published: 18 October 2020



Abstract: Glycosaminoglycans (GAGs) are the linear carbohydrate components of proteoglycans (PGs) and are key mediators in the bioactivity of PGs in animal tissue. GAGs are heterogeneous, conformationally complex, and polydisperse, containing up to 200 monosaccharide units. These complexities make studying GAG conformation a challenge for existing experimental and computational methods. We previously described an algorithm we developed that applies conformational parameters (i.e., all bond lengths, bond angles, and dihedral angles) from molecular dynamics (MD) simulations of nonsulfated chondroitin GAG 20-mers to construct 3-D atomic-resolution models of nonsulfated chondroitin GAGs of arbitrary length. In the current study, we applied our algorithm to other GAGs, including hyaluronan and nonsulfated forms of dermatan, keratan, and heparan and expanded our database of MD-generated GAG conformations. Here, we show that individual glycosidic linkages and monosaccharide rings in 10- and 20-mers of hyaluronan and nonsulfated dermatan, keratan, and heparan behave randomly and independently in MD simulation and, therefore, using a database of MD-generated 20-mer conformations, that our algorithm can construct conformational ensembles of 10- and 20-mers of various GAG types that accurately represent the backbone flexibility seen in MD simulations. Furthermore, our algorithm efficiently constructs conformational ensembles of GAG 200-mers that we would reasonably expect from MD simulations.

Keywords: molecular dynamics; glycosaminoglycan; proteoglycan; hyaluronan; dermatan; keratan; heparan; carbohydrate conformation; carbohydrate flexibility; glycosidic linkage; ring pucker; force field; explicit solvent

1. Introduction

Proteoglycans (PGs) are a diverse group of proteins in the extracellular matrix (ECM), as well as on and within cells in animal tissue. PGs play key roles in signal transduction [1,2], tissue morphogenesis [3–7], and matrix assembly [7–10] by binding growth factors [3–7,11–17], enzymes [7,17], membrane receptors [17], and ECM molecules [2,7,17]. Many of these functions are mediated by glycosaminoglycans (GAGs), which are linear, highly negatively-charged, and structurally diverse carbohydrate chains covalently-linked to PGs. Specifically, GAGs can form covalent and noncovalent complexes with proteins or inhibit complex formation with other biomolecules. All of these functions allow GAGs to modulate disease. For example, hyaluronan can predict disease outcome and is used as

a treatment for osteoarthritis [18,19]; deficiency of dermatan sulfate has been linked to Ehlers-Danlos syndrome, so screening of dermatan sulfate in urine may present a method of diagnosis [20,21]; reduced cerebral cell keratan sulfate levels have been implicated in Alzheimer's disease [22]; and heparan sulfate has been shown to induce septic shock [23,24].

GAG-protein binding is determined by the sequence, structure, shape, and charge of the GAG and GAG-binding region of the protein [17,25]. Therefore, the three-dimensional structure and conformation of GAGs determine their bioactivity, and even subtle structural differences can change their function. For example, while hyaluronan and chondroitin have many functional differences, the only structural difference is in the chirality of carbon four in the amide monosaccharide derivatives (i.e., N-acetylglucosamine and N-acetylgalactosamine, respectively). The structural and conformational complexities of GAGs make studying GAG conformational thermodynamics at atomic resolution a challenge using existing experimental methods. Structural complexities arise from nontemplate-based synthesis [26] and variable sulfation [27], which means a single GAG type can have variable lengths and sulfation patterns, making it polydisperse and heterogeneous. Conformational complexities result from glycosidic linkage flexibility [28–33] and ring puckering [33]. There are several experimental techniques for studying GAGs, including liquid chromatography-mass spectrometry (LC-MS) [34–36], X-ray crystallography [37–42], and nuclear magnetic resonance (NMR) [43–46], but no single technique is able to account for all of these complexities. While each has its limitations, these experimental techniques provide a critical means to validate molecular dynamics (MD) simulations [46–52]. MD can be used to generate atomic-resolution conformational ensembles of GAGs under physiological conditions, and agreement between MD-generated results and available experimental data builds confidence that MD does indeed generate realistic three-dimensional atomic-resolution GAG models [53–57]. However, all-atom explicit-solvent MD is not a feasible tool for routine simulation of large systems, and GAG biopolymers may consist of up to 200 monosaccharide units [4], which, when fully solvated, would result in a system with over 10^6 atoms. Coarse-grained (CG) MD is a feasible alternative for simulations of long GAG polymers, but may be limited because it lacks the atomic resolution and the accurate treatment of solvation provided by all-atom explicit-solvent MD [48,58–60].

As an alternative approach to obtaining atomic-resolution, three-dimensional, conformational ensembles of long GAG biopolymers, we developed an algorithm that uses glycosidic linkage and monosaccharide ring conformations from unbiased all-atom explicit-solvent MD simulations [57,61–63] of GAG 20-mers (i.e., 20 monosaccharide units) to rapidly generate GAG polymer conformational ensembles of user-specified length and with a user-specified number of conformations [33]. In our previous work, we examined nonsulfated chondroitin GAGs. Here, we aim to expand the application of our algorithm to hyaluronan and to nonsulfated forms of dermatan, keratan, and heparan. Nonsulfated forms were chosen for their simplicity and homogeneity. The algorithm is expected to produce conformational ensembles that mimic the backbone flexibility observed in simulations only if individual ring and linkage conformations of GAG 20-mers behave randomly and independently in simulation. This means that the algorithm will not produce polymer conformations with higher-order structures.

It has been suggested that hyaluronan may take on higher order structures [64], and while it has been shown that helical structures occur in hyaluronan oligosaccharides [65,66], this behavior may not necessarily be extrapolated to long polysaccharides under physiological conditions. Hyaluronan oligosaccharides have been shown to exhibit rigid structures in aqueous solution, possibly resulting from electrostatic effects and hydrogen-bonding between neighboring monosaccharides [65,67–73]. MD studies of hyaluronan di- and tri- saccharides in aqueous solution (without ions) revealed that hydrogen-bonds form but do not remain for long periods of time, and do not always occur at the same time in different regions of the oligosaccharide [67,68]. It has been shown that under physiological conditions, the extended conformations and large hydrodynamic radius of hyaluronan can be attributed primarily to hydrogen-bonding between adjacent monosaccharides (and less so to electrostatic effects) [69]. It is important to note that each of these studies examined short hyaluronan

oligosaccharides; thus, atomic-resolution conformational properties of long polysaccharides under physiological conditions have not been reported.

We sought to determine (1) if longer polysaccharides of hyaluronan and nonsulfated dermatan, keratan, and heparan took on a higher-order structure during MD simulations in an aqueous solution with physiological ion strength and, consequently, (2) if our algorithm could construct conformational ensembles of GAGs that mimic the backbone flexibility observed in MD simulations. To this end, we ran a microsecond-scale, all-atom, explicit-solvent MD on 20-mers of each GAG type with sequences [-4 glucuronate β 1-3 N-acetylglucosamine β 1-]₁₀ (hyaluronan), [-4 iduronate α 1-3 N-acetylgalactosamine β 1-]₁₀ (dermatan), [-3 galactose β 1-4 N-acetylglucosamine β 1-]₁₀ (keratan), and [-4 iduronate α 1-4 N-acetylglucosamine α 1-]₁₀ (heparan). Analysis of glycosidic linkage and monosaccharide ring conformations revealed no apparent correlation between individual rings and flanking linkages in any of the 20-mers. Glycosidic linkages took on experimentally-observed helical conformations, and some linkages sampled secondary conformations that caused a kink in the polymer. No patterns of interdependence between adjacent linkage conformations were seen. While ring and linkage conformations did not reach equilibrium in hyaluronan simulations on the microsecond timescale, overall hyaluronan 20-mer appeared to behave randomly. End-to-end distance distributions revealed mostly extended conformations (n.b. we use the term “extended” throughout to describe conformations with high end-to-end distances; this does not necessarily contradict helical structure) with some compact conformations resulting from non-⁴C₁ ring puckering, which brought the oxygen atoms in sequential glycosidic linkages closer together, and certain glycosidic linkage conformations that brought neighboring monosaccharides closer together. Thus, these ring and linkage conformations both introduce a kink in the polymer. These conformations are independent and occur rarely on the microsecond timescale. The variability in individual ring and linkage conformations, as well as end-to-end distances and the radii of gyration, demonstrate the flexibility and lack of stable higher-order structures of each of these nonsulfated GAG polymers during simulation in aqueous solution.

For each nonsulfated GAG type, monosaccharide ring and glycosidic linkage conformations from MD-generated 20-mers were used as inputs in our algorithm to generate conformational ensembles of 10-, 20-, and 200-mers. For each GAG type, 10- and 20-mer constructed ensembles were compared to 10- and 20-mer MD-generated ensembles. In each case, end-to-end distances in constructed and MD-generated ensembles match, with minimal differences in most probable end-to-end distances. The similarities in backbone conformation of constructed and MD-generated 10- and 20-mer ensembles provide further evidence that there is little correlation between individual glycosidic linkage and ring conformations, and thus, that nonsulfated GAG 10- and 20-mers do not exhibit a higher-order structure. This suggests that our program can be used to construct realistic conformations of long GAG biopolymers (e.g., 200 monosaccharides), and the resulting atomic-resolution models may be used to develop further insights into GAG binding properties, and thus, GAG-protein complex formation. These will be especially useful for studying complexes in which there are multiple biomolecules bound to the same GAG.

Other programs exist which are capable of constructing three-dimensional, atomic-resolution GAG polymer models including Glycam GAG Builder [74], POLYS Glycan Builder [75], CarbBuilder [76], MatrixDB GAG Builder [77,78], and CHARMM-GUI [79–82]. These are useful tools for constructing the initial structures of GAGs having user-specified sequences and lengths for MD simulations. Additional features of Glycam and POLYS Glycan Builder include user specification of glycosidic linkage dihedral angle values or construction using conformations from their databases. Glycam, POLYS Glycan Builder, and Carb Builder databases consist of GAG mono- and di- saccharide structures sampled in MD and/or determined by molecular mechanics calculations. The MatrixDB databases contain GAG disaccharide conformations from GAG-protein complex crystal structures. CHARMM-GUI uses the CHARMM software [83–85] to construct GAGs with user-specified sequences, using default force field parameters for carbohydrates [57,61–63], adding explicit solvents with user-specified water box size and ion concentrations, and generating inputs for MD equilibration and production simulations.

Although the user has the option to choose the GAG length, these tools are best used for short GAG oligosaccharides. Our algorithm differs from these tools in that it draws upon our databases containing full MD-generated unbound GAG 20-mer atomic-resolution conformational ensembles. Furthermore, our algorithm is developed for modeling long GAG biopolymers (e.g., 200-mers), and can efficiently produce large conformational ensembles (e.g., 10,000 3-D models) of long GAG biopolymers. Therefore, our algorithm provides an alternative to MD simulation that takes much less time and computational resources.

2. Materials and Methods

The following protocol was conducted for each of the hyaluronan and nonsulfated dermatan, keratan, and heparan GAGs.

2.1. Molecular Dynamics: System Construction, Energy Minimization, Heating, and Production Simulations

System construction, energy minimization, and heating followed the same protocol we described previously [33]. Briefly, all polymers were constructed with an explicit solvent using the CHARMM software [83–85] v. c41b2 with the CHARMM36 (C36) biomolecular force field for carbohydrates [57,61–63]. The initial conformation of each GAG 10- and 20-mer had glycosidic linkage dihedrals (with IUPAC definitions: $\phi = \text{O}_5\text{-C}_1\text{-O-C}_n$ and $\psi = \text{C}_1\text{-O-C}_n\text{-C}_{(n-1)}$) set to $\phi = -83.75^\circ$ and $\psi = -156.25^\circ$ in all GlcA β 1-3GlcNAc (hyaluronan), IdoA α 1-3GalNAc (dermatan), and Gal β 1-4GlcNAc (keratan) linkages and $\phi = -63.75^\circ$ and $\psi = 118.75^\circ$ in all GlcNAc β 1-4GlcA (hyaluronan), GalNAc β 1-4IdoA (dermatan), and GlcNAc β 1-3Gal (keratan) linkages. These dihedral values were the most energetically stable in nonsulfated chondroitin disaccharide glycosidic linkages during MD simulation [55]. In heparan 10- and 20-mers, all IdoA α 1-4GlcNAc glycosidic linkage dihedrals were set to $\phi = -69^\circ$ and $\psi = 147^\circ$ and all GlcNAc α 1-4IdoA linkage dihedrals were set to $\phi = 77^\circ$ and $\psi = 30^\circ$. These dihedral values were found in an NMR and force field study of heparan disaccharides [30], and gave extended heparan polymer conformations. With these glycosidic linkage conformations, 20-mer end-to-end distances were 99.0 Å, 89.7 Å, 85.8 Å, and 87.5 Å for hyaluronan, dermatan, keratan, and heparan, respectively. The GAG 20-mers were solvated in cubic periodic unit cells with edge lengths of 124.3 Å (63,347 water molecules), 115.0 Å (50,066 water molecules), 111.9 Å (46,103 water molecules), and 111.9 Å (46,122 water molecules) for hyaluronan, dermatan, keratan, and heparan, respectively. Each of these edge lengths is at least 10 Å longer than the maximum dimension of the corresponding 20-mer in a fully-extended (i.e., maximum end-to-end distance) conformation. The TIP3 water model [86,87], as implemented in the CHARMM force field, was used for all simulations, and neutralizing Na⁺ counterions and 140 mM NaCl [88] were added.

The NAMD program [89] v. 2.12 (<http://www.ks.uiuc.edu/Research/namd/>) was used to minimize the potential energy and heat the system to the target temperature of 310 K under constant pressure at 1 atm. An unbiased constant particle number/constant pressure/constant temperature (NPT) MD run followed, and average periodic cell parameters from the last half of this trajectory were used as cell basis vectors for the quadruplicate canonical (NVT) ensemble MD simulations. NVT MD simulations were preceded by minimization and heating at constant volume with box edge lengths of 123.7 Å, 114.4 Å, 111.3 Å, and 111.3 Å for hyaluronan, dermatan, keratan, and heparan, respectively.

Production simulations were performed using a similar protocol as previously described [33] but with a longer timescale and lower frequency for saving snapshots. Briefly, using the CHARMM software, each of four replicates (per GAG system) was equilibrated. Then, unbiased canonical (NVT) ensemble MD simulations were run for 500,000,000 steps (1 μ s) and atomic coordinates were saved at 50,000-step (100-ps) intervals for analyses (10,000 snapshots per simulation, 40,000 snapshots per GAG).

2.2. Conformational Analysis

GAG conformational parameters include bond length, bond angle, and dihedral angle values. The glycosidic linkage conformational parameters considered were bond lengths for C₁-O and O-C_n,

bond angles for O_5-C_1-O and C_1-O-C_n , and dihedral angles ϕ and ψ (Figure 1). To characterize potential patterns in glycosidic linkage conformations, dihedral angle free energies $\Delta G(\phi, \psi)$ were analyzed. As ϕ, ψ dihedral values from MD-generated ensembles were taken to have uniform probabilities, $\Delta G(\phi, \psi)$ was calculated by sorting dihedral angle values into $2.5^\circ \times 2.5^\circ$ bins and then using the formula $\Delta G(\phi_i, \psi_j) = -RT \ln(n_{ij}) - k$, where n_{ij} is the bin count for the bin corresponding to ϕ_i, ψ_j , R is the universal gas constant, T is the temperature of the MD simulations, and k was chosen so that the global minimum was located at $\Delta G = 0$ kcal/mol.

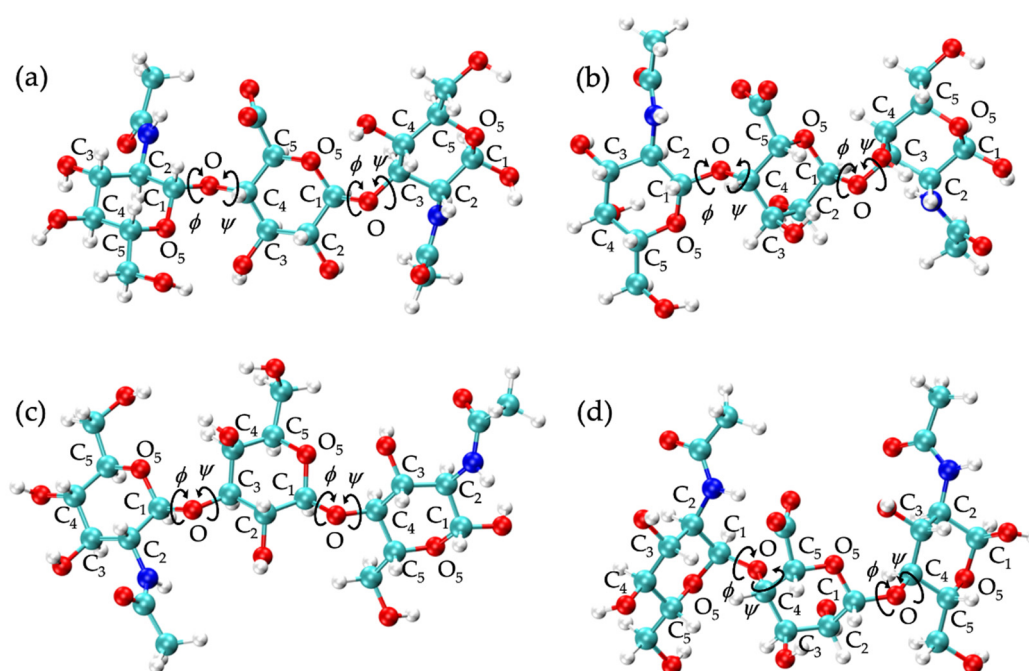


Figure 1. GAG trisaccharides: (a) hyaluronan [GlcNAc β 1-4GlcA β 1-3GlcNAc], (b) nonsulfated dermatan [GalNAc β 1-4IdoA α 1-3GalNAc], (c) nonsulfated keratan [GlcNAc β 1-3Gal β 1-4GlcNAc], (d) nonsulfated heparan [GlcNAc α 1-4IdoA α 1-4GlcNAc]; endocyclic ring atoms and oxygen atoms linking adjacent rings are labeled; arrows indicate rotations about glycosidic linkage bonds C_1-O and $O-C_n$ (quantified by ϕ and ψ dihedral angles, respectively); glycosidic linkage parameters used in the construction algorithm include C_1-O and $O-C_n$ bond lengths, O_5-C_1-O and C_1-O-C_n bond angles, and $\phi = O_5-C_1-O-C_n$ and $\psi = C_1-O-C_n-C_{(n-1)}$ dihedral angles. Molecular graphics throughout are produced with the VMD program [90].

Monosaccharide ring geometries are quantified by all bond lengths, bond angles, and dihedral angles within the ring and in exocyclic functional groups that are not part of a glycosidic linkage. Cremer-Pople (C-P) ring puckering parameters (ϕ, θ, Q) [91] of each monosaccharide ring in the MD-generated 20-mer ensembles were computed to characterize potential ring puckering patterns. Conformations of each monosaccharide and each linkage of the 20-mers were extracted separately from each snapshot of the MD trajectories. First, to determine if conformational data in different runs matched, and if there was interdependency in the individual linkage and ring conformations, data were separated by run and monosaccharide/linkage number and aggregated across all snapshots in each MD simulation run. Next, all individual conformations were aggregated across all snapshots in all runs (e.g., 10,000 snapshots * 4 runs * 10 GlcA monosaccharides = 400,000 samples of GlcA monosaccharide conformations in hyaluronan 20-mer) to generate a single dataset for each of the two monosaccharide types and two linkage types in each GAG (e.g., GlcA monosaccharide, GlcNAc monosaccharide, β 1-3 glycosidic linkage, and β 1-4 glycosidic linkage conformations in hyaluronan 20-mer).

2.3. Construction Algorithm to Generate GAG Conformational Ensembles

The MD-generated conformational datasets described above made up the database used for the algorithm we developed to generate GAG polymer conformational ensembles of user-specified length and with a user-specified number of conformations (previously described) [33]. Essentially, our algorithm (1) incorporates all bond, bond angle, and dihedral angle conformational parameters from MD simulation of GAG 20-mers, (2) treats monosaccharide rings and glycosidic linkages independently, (3) performs a restrained energy minimization on each constructed conformation to relieve steric overlap while maintaining polymer conformation, and (4) applies a bond potential energy cutoff to exclude conformations that remain nonphysical after minimization. A nonphysical conformation results from either overlapping bonds or a bond that pierces the center of another monosaccharide ring in the initial constructed conformation (Figure S1). As dihedral angles are restrained, minimization addresses these issues by stretching the overlapping or ring-piercing bonds to nonphysical lengths, which increases the post-minimization bond potential energy by more than 132 kcal/mol [33]. The energy cutoff is the sum of a 100 kcal/mol buffer and a polymer-length specific cutoff, which is equal to the post-minimization bond potential energy of the fully-extended conformation. This conformation was constructed by assigning energetically-favorable glycosidic linkage dihedrals from the corresponding MD simulations of each GAG type that give a fully-extended conformation (i.e., with the maximum end-to-end distance observed in simulation). We used the algorithm to construct 20-mer conformational ensembles of each nonsulfated GAG.

For internal validation of the algorithm, glycosidic linkage dihedral free energies $\Delta G(\phi, \psi)$, monosaccharide ring C-P parameters, end-to-end distance distributions, and radii of gyration from MD-generated ensembles and constructed ensembles were compared. Additionally, post-minimization bond potential energy distributions from constructed ensembles were analyzed to validate the calculated bond potential energy cutoff and verify that the ensemble had expected energy distributions for the given polymer length.

To evaluate application of MD-generated 20-mer conformations to construct GAG polymers of variable length, we constructed 10-mer ensembles using the algorithm and compared them to 10-mer ensembles generated using the same protocol as 20-mer MD simulations. To assess the efficacy and efficiency of our algorithm to construct conformational ensembles of GAG polymers with biologically-relevant chain lengths, we also constructed 200-mer ensembles.

3. Results and Discussion

3.1. Hyaluronan

3.1.1. Molecular Dynamics Simulations: Glycosidic Linkage and Monosaccharide Ring Geometry Effects on Polymer Backbone Flexibility

To examine the backbone flexibility of hyaluronan 20-mers, end-to-end distances and radii of gyration were analyzed in each of the four MD simulation runs. One of the four runs produced more compact conformations, i.e., with lower end-to-end distances and radii of gyration than the other three (Figures 2 and S2a and Table 1). The system potential energy distributions were identical for all four runs, suggesting that this outlying run was energetically stable (Figure S3). To uncover the conformational factors that contributed to end-to-end distance and explain the differences in backbone flexibility in the different MD runs, the monosaccharide ring and glycosidic linkage conformations were examined.

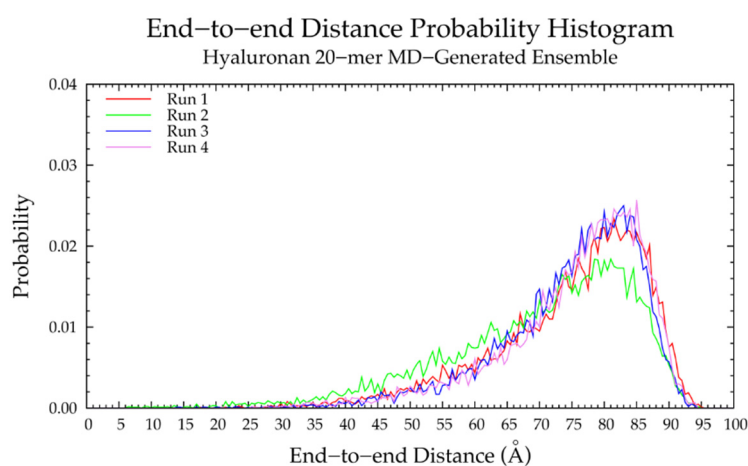


Figure 2. End-to-end distance probability distribution of MD-generated hyaluronan 20-mer ensemble; each of the four runs includes 10,000 conformations; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins.

Table 1. Most Probable End-to-End Distances (d) in MD-Generated and Constructed Hyaluronan Ensembles ¹.

	20-mer Ensembles			10-mer Ensembles		
	MD-Generated d (Å)	Constructed d (Å)	% Difference	MD-Generated d (Å)	Constructed d (Å)	% Difference
Run 1	81.5	80.5		43.00	42.25	
Run 2	81.0	78.0		42.50	41.75	
Run 3	83.0	80.5		44.00	42.75	
Run 4	85.0	78.0		43.50	42.75	
All ²	80.0	78.0	2.53%	43.50	41.75	4.11%

¹ Probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer ensembles and 0.25 Å bins for the 10-mer ensembles. ² All = end-to-end distance distribution aggregated across all four runs.

Analysis of C-P parameters of GlcNAc monosaccharide rings revealed mostly ⁴C₁ chair conformations, as found in crystal structures [92–99] and NMR and force field studies [67,100–102], with a minority of samples in boat and skew-boat conformations, namely B_{3,O}, ¹S₃, ^{1,4}B, and ¹S₅ (Figures 3a and 4a). While non-⁴C₁ conformations of GlcNAc are rare, they have been observed in crystal structures of protein co-complexes [95,103–106]. For example, one GlcNAc ring in the crystal structure of a hyaluronan tetramer in complex with hyaluronidase was found in the ^{1,4}B boat conformation [95]. Additionally, GlcNAc boat/skew-boat conformations have been sampled in biased MD [107], 1-μs unbiased MD [108], and 10-μs unbiased MD simulations of GlcNAc monosaccharides [100]. One NMR and force field study suggested that these conformations were important for GAG-protein interactions [100]. Analysis of individual GlcNAc rings in each run revealed that these non-⁴C₁ conformations were not specific to any particular region of the 20-mer and did not occur simultaneously in multiple rings in the same snapshot (Figure S4). Furthermore, GlcNAc rings that sampled non-⁴C₁ puckers returned to ⁴C₁ chair conformations relatively quickly, i.e., in under 25 ns (Figure S4).

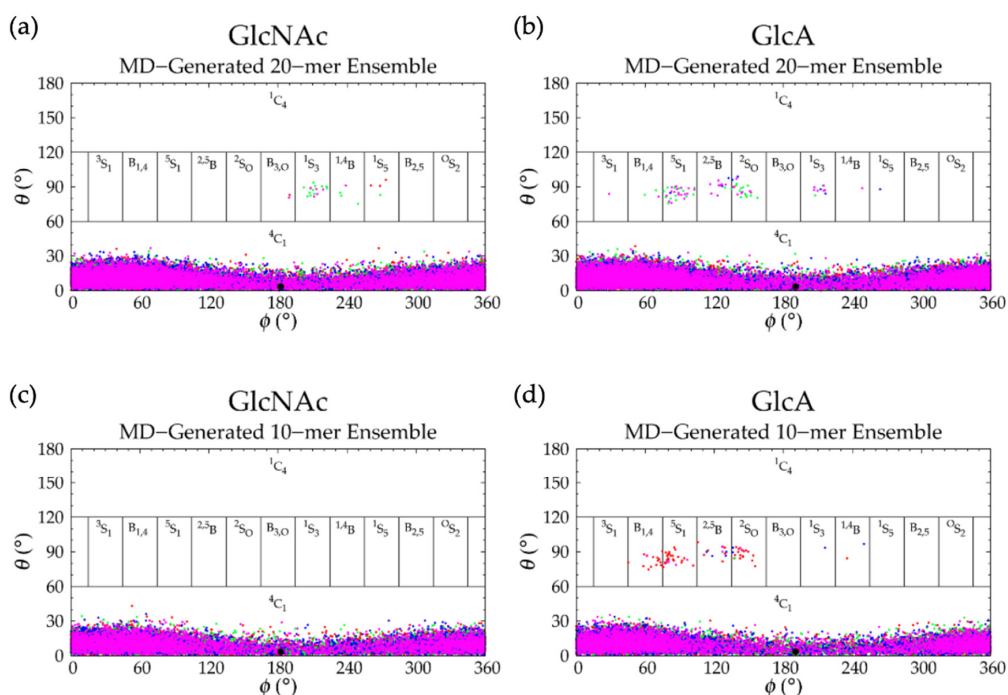


Figure 3. Cremer–Pople data for MD-generated hyaluronan ensembles: (a) GlcNAc and (b) GlcA in the 20-mer and (c) GlcNAc and (d) GlcA in the 10-mer; geometries from the four sets of each ensemble are represented by red, green, blue, and magenta dots, respectively and the force-field geometry is represented by a single large black dot; Cremer–Pople parameters (ϕ , θ) for all rings in every tenth snapshot from each ensemble were plotted (i.e., 10 rings \times 1000 snapshots per run \times 4 runs = 40,000 parameter sets for the 20-mer and 20,000 parameter sets for the 10-mer).

C-P analysis of GlcA monosaccharide rings in the hyaluronan 20-mer showed that GlcA is found mostly in 4C_1 chair conformations (Figure 3b), as seen in crystal structures [92–95] and NMR and force field studies [51,67,101,102,107,109,110], with some boat and skew-boat conformations, including 3S_1 , $B_{1,4}$, 5S_1 , ${}^{2,5}B$, 2S_O , 1S_3 , ${}^{1,4}B$, and 1S_5 (Figures 3b and 4b,c), which were also observed in GlcA rings in unbiased MD simulations of nonsulfated chondroitin 20-mers [33]. Some of these conformations were also sampled rarely in unbiased MD simulations of GlcA monosaccharides [51,107]. The slight differences between boat/skew-boat conformations in our simulations and monosaccharide simulations may have arisen from intramolecular interactions in the hyaluronan 20-mer. As with GlcNAc rings, these conformations did not occur consistently in any one region of the 20-mer, did not occur simultaneously in multiple rings in the same snapshot, and returned to 4C_1 chair conformations quickly, i.e., within 40 ns (Figure S5). Except for 2S_O , each of these GlcA and GlcNAc boat and skew-boat conformations brought the linker oxygen atoms closer together, causing a kink in the polymer chain (Figure 4), which may have contributed to more compact polymer conformations. To investigate this, end-to-end distances and radii of gyration of 20-mer conformations with non- 4C_1 ring puckers were analyzed (Figure S6) and both compact and extended conformations were observed. This can be explained by the flexibility in the glycosidic linkages. Of note, MD-generated atomic-resolution snapshots before and after a monosaccharide ring underwent a conformational change between 4C_1 and non- 4C_1 were visualized in VMD to check for ionic interactions (e.g., between Na^+ and carboxylate O^- atoms and/or partially-charged hydroxyl O atoms in GlcA, between Cl^- and hydroxyl H atoms in GlcA, between Na^+ and partially-charged N-acetyl N or O atoms and/or hydroxyl O atoms in GlcNAc, or between Cl^- and H atoms of methyl and/or hydroxyl groups in GlcNAc). Specifically, we looked for ions within 5 Å of the exocyclic atoms of the monosaccharide of interest and found that ionic interactions were most often absent during, and thus, did not significantly contribute to, ring puckering in the MD simulation of hyaluronan 20-mer.

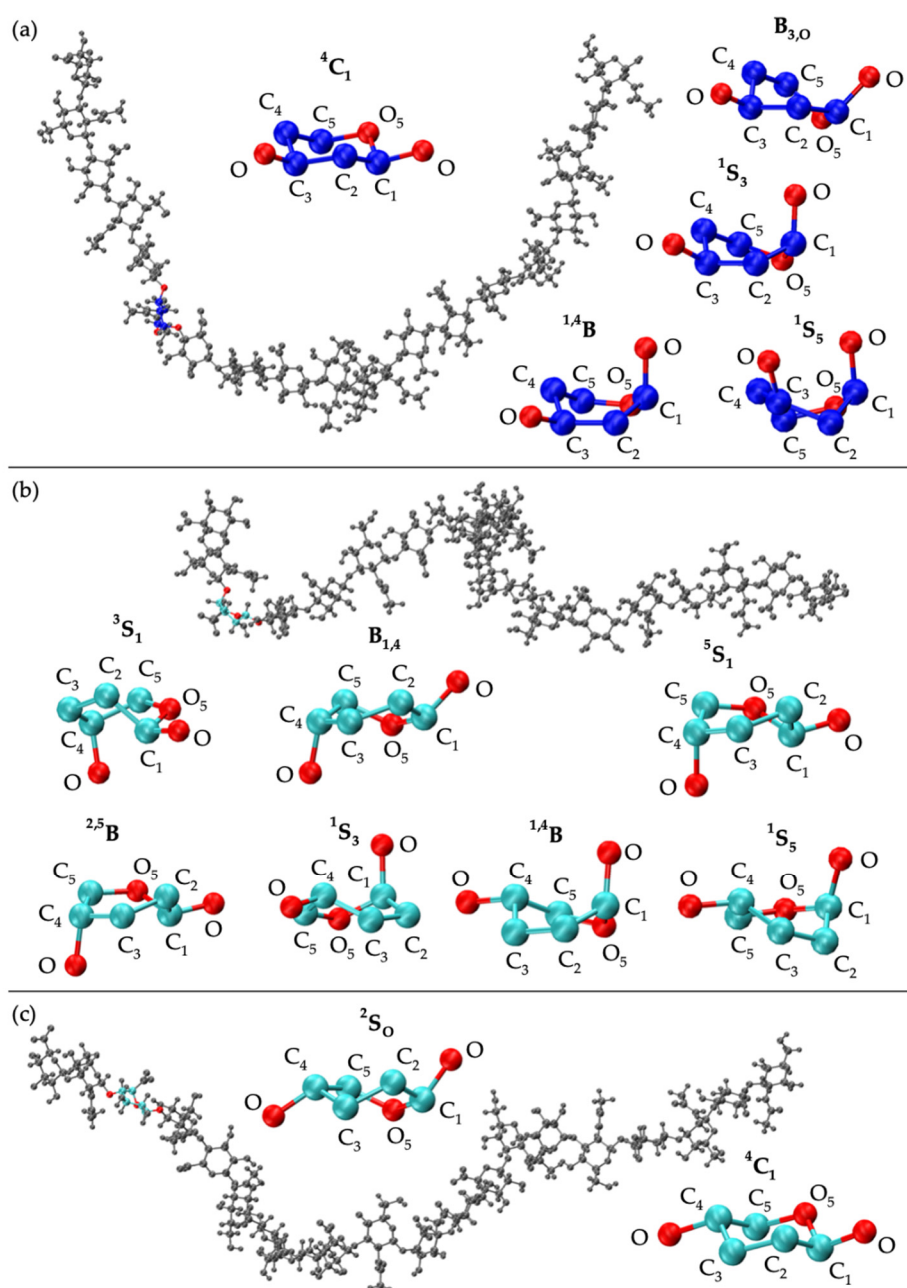


Figure 4. MD-generated hyaluronan conformations: (a) 20-mer with a 1^4B GlcNAc conformer (ring atoms in blue; linkage atoms in red) and 4^1C_1 , $B_{3,0}$, 1^1S_3 , 1^4B , and 1^1S_5 GlcNAc monosaccharide conformations (endocyclic ring atoms and linker oxygen atoms only); (b) 20-mer with a 5^1S_1 GlcA conformer (ring atoms in cyan; linkage atoms in red) and 3^1S_1 , $B_{1,4}$, 5^1S_1 , 2^5B , 1^1S_3 , 1^4B , and 1^1S_5 GlcA monosaccharide conformations (endocyclic ring atoms and linker oxygen atoms only); (c) 20-mer with a 2^2S_0 GlcA conformer (ring atoms in cyan; linkage atoms in red) and 4^1C_1 and 2^2S_0 GlcA monosaccharide conformations (endocyclic ring atoms and linker oxygen atoms only); in each 20-mer snapshot, all gray rings are in 4^1C_1 conformations.

To analyze glycosidic linkage conformations, free energies of glycosidic linkage dihedrals $\Delta G(\phi, \psi)$ were calculated (Figure 5a,b and Table 2). If we assume $O_5-C_1-O-C_n = H_1-C_1-O-C_n - 120^\circ$ for ϕ values and $C_1-O-C_3-C_2 = C_1-O-C_3-H_3 - 120^\circ$ and $C_1-O-C_4-C_3 = C_1-O-C_4-H_4 + 120^\circ$ for ψ values, our data match well with X-ray diffraction and force field data from hyaluronan tetrasaccharides [49], as well as NMR and molecular mechanics data from hyaluronan tetrasaccharides [111], hexasaccharides [112],

and octasaccharides [113]. If we assume $C_1-O-C_3-C_2 = C_1-O-C_3-C_4 + 120^\circ$ and $C_1-O-C_4-C_3 = C_1-O-C_4-C_5 - 120^\circ$, our data are also in close agreement with X-ray diffraction and quantum molecular modeling data [65,114–116], as well as NMR and MD data [102]. This agreement is with conformations sampled about the global minimum in each linkage in our MD simulations. All of these conformations were helical, so it stands to reason that 2-, 3-, and 8-fold helices were observed in hyaluronan oligosaccharides. This does not necessarily mean that adjacent glycosidic linkage conformations were interdependent. In fact, variation in the types of helices observed in the force field and experimental studies in only short oligosaccharide units suggests a lack of consistent and uniform helical structures in long hyaluronan polymers.

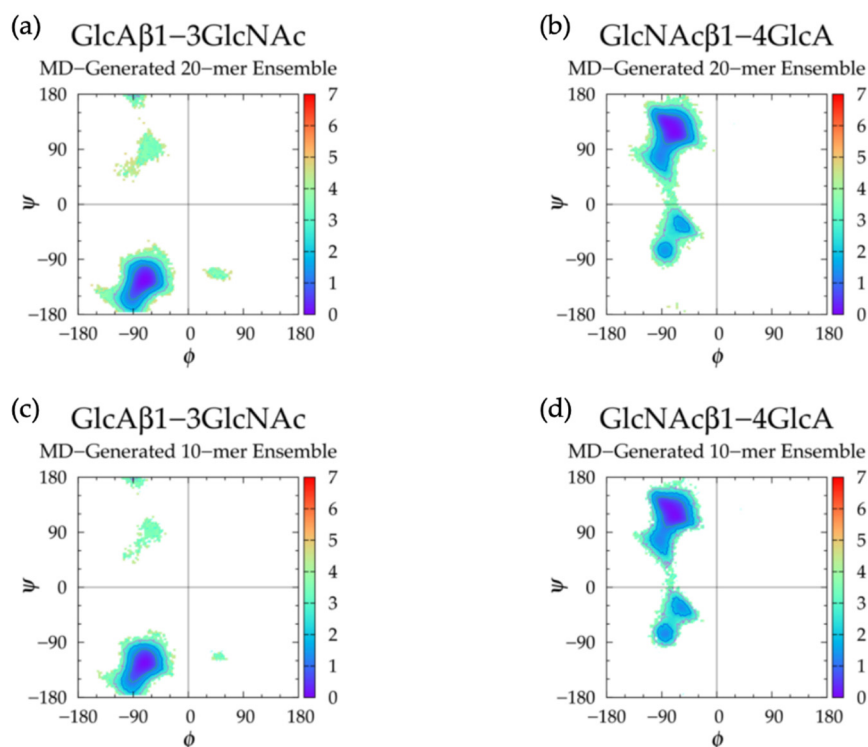


Figure 5. $\Delta G(\phi, \psi)$ in the MD-generated hyaluronan ensembles for aggregated GlcA β 1-3GlcNAc and GlcNAc β 1-4GlcA glycosidic linkage data in the (a,b) 20-mer and (c,d) 10-mer, respectively; contour lines every 1 kcal/mol.

Table 2. Hyaluronan Glycosidic Linkage Dihedrals $(\phi, \psi)^1$ and Free Energy ($\Delta G(\phi, \psi)$ kcal/mol) Minima.

	MD-Generated 20-mer Ensemble				MD-Generated 10-mer Ensemble				Constructed 20-mer Ensemble			
	GlcA β 1-3 GlcNAc		GlcNAc β 1-4 GlcA		GlcA β 1-3 GlcNAc		GlcNAc β 1-4 GlcA		GlcA β 1-3 GlcNAc		GlcNAc β 1-4 GlcA	
Min	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG
I	$-71.25^\circ,$ -123.75°	0.00	$-68.75^\circ,$ 116.25°	0.00	$-71.25^\circ,$ -123.75°	0.00	$-68.75^\circ,$ 118.75°	0.00	$-71.25^\circ,$ -123.75°	0.00	$-71.25^\circ,$ 116.25°	0.00
II	$-53.75^\circ,$ 91.25°	3.06	$-83.75^\circ,$ -73.75°	1.31	$-58.75^\circ,$ 86.25°	3.08	$-81.25^\circ,$ -76.25°	1.16	$-56.25^\circ,$ 91.25°	3.01	$-83.75^\circ,$ -73.75°	1.29
II'			$-58.75^\circ,$ -33.75°	1.42			$-53.75^\circ,$ -36.25°	1.27			$-56.25^\circ,$ -36.25°	1.40

¹ ϕ, ψ dihedral angles were sorted into 2.5° bins.

Our data also show that a secondary basin was sampled in each linkage type in our MD simulations. In GlcA β 1-3GlcNAc linkages, conformations in the upper left quadrant ($-\phi, +\psi$) were sampled. A molecular modeling study that reportedly sampled the full allowable range of

glycosidic linkage conformations in agreement with experimental data for several disaccharide units also revealed hyaluronan GlcA β 1-3GlcNAc linkage conformations in a secondary basin similar to ours [117]. Furthermore, crystal structures of hyaluronan oligosaccharides in complexes with proteins [92–95] were examined and we found that most glycosidic linkage conformations were similar to the most energetically-favorable conformations observed in MD simulation, with the exception of one GlcA β 1-3GlcNAc linkage conformation with $-\phi, +\psi$ dihedrals, which was seen in a hyaluronan hexasaccharide in complex with hyaluronidase [94]. This suggests that, although rare, this is a physical conformation. Analysis of individual glycosidic linkages in individual runs (Figure S7) showed that this basin was sampled mostly in two particular GlcA β 1-3GlcNAc linkages in the MD run with the lowest end-to-end distance distribution. Therefore, 20-mer MD snapshots with these glycosidic linkage conformations were analyzed; we found that these linkage conformations were associated with lower end-to-end distances (Figure S8 and Table S1). Visual analysis of these snapshots showed that GlcA β 1-3GlcNAc glycosidic linkages with $-\phi, +\psi$ dihedrals caused a kink in the polymer, which explains the more compact polymer conformations in comparison to those containing only linkage conformations at the global free energy minimum (Figure S8). To find out if there was a connection between these glycosidic linkage conformations and non- 4C_1 puckers of adjacent monosaccharide rings, glycosidic linkages flanking non- 4C_1 ring puckers were plotted, and they were all centered about the global minimum, i.e., $-\phi, -\psi$ (Figure S9). In fact, there did not appear to be any correlation between GlcA β 1-3GlcNAc linkage $-\phi, +\psi$ conformations and a non- 4C_1 ring pucker in any region of the polymer.

Analysis of GlcNAc β 1-4GlcA glycosidic linkage conformations (Figure 5b and Table 2) revealed secondary minima in the lower left quadrant ($-\phi, -\psi$), which was also seen in a molecular modeling study of GAG disaccharides with results validated by experimental data [117]. In hyaluronan 20-mer MD simulations, these conformations were sampled more in the GlcNAc β 1-4GlcA linkages neighboring the GlcA β 1-3GlcNAc linkages with $-\phi, +\psi$ dihedrals in the MD run with the lowest end-to-end distances. The same analyses to those performed on 20-mer conformations with GlcA β 1-3GlcNAc linkages with $-\phi, +\psi$ dihedrals were performed on MD-generated 20-mer snapshots with GlcNAc β 1-4GlcA linkages with $-\phi, -\psi$ dihedrals. Similarly, these glycosidic linkage conformations caused a kink in the polymer chain and were associated with lower end-to-end distances (Figure S10 and Table S1). To determine if there was a correlation between GlcA β 1-3GlcNAc $-\phi, +\psi$ linkage dihedrals and GlcNAc β 1-4GlcA $-\phi, -\psi$ linkage dihedrals, snapshots with both of these linkage conformations flanking the same monosaccharide ring were examined. It was found that there were very few snapshots in which this was the case. GlcNAc β 1-4GlcA linkage conformations flanking non- 4C_1 monosaccharide rings were also analyzed, and these conformations were centered about the global minimum, i.e., $-\phi, +\psi$ (Figure S9). Additionally, ionic interactions (e.g., between Na $^+$, O $^-$ in GlcA carboxylate group, and O in C=O of adjacent GlcNAc N-acetyl group, between Cl $^-$, H in GlcNAc methyl group, and H in GlcA hydroxyl groups, or between Na $^+$ and O atoms or Cl $^-$ and H atoms of hydroxyls in adjacent monosaccharides) did not appear to play a role in the transition of glycosidic linkage conformations between their primary and secondary ϕ, ψ basins.

Hyaluronan 10-mers were also simulated and showed the same conformations as the 20-mer in MD simulation except for GlcNAc monosaccharide rings, which sampled only 4C_1 chair conformations in the 10-mer (Figures 3c,d and 5c,d and Table 2). This was expected, as these conformations occurred very few times in 20-mer MD simulations, and there were half as many samples (i.e., monosaccharide rings) in the 10-mer.

Together, these results suggest that (1) realistic helical glycosidic linkage conformations in hyaluronan 10- and 20-mers are seen in MD simulation, (2) non- 4C_1 ring puckers, $-\phi, +\psi$ dihedrals in GlcA β 1-3GlcNAc linkages, and $-\phi, -\psi$ dihedrals in GlcNAc β 1-4GlcA linkages all contribute to more compact conformations, and (3) monosaccharide rings and glycosidic linkages in hyaluronan 10- and 20-mers behave randomly and independently in MD simulation.

3.1.2. Construction Algorithm

Our construction algorithm was used to construct four sets of hyaluronan 20-mer ensembles with 10,000 conformations each (40,000 conformations total). A comparison of end-to-end distance distributions (Figure 6a and Table 1) and radii of gyration (Figure S2a,b) showed that our construction algorithm produces ensembles that mimic backbone flexibility observed in MD simulations. Furthermore, the C-P parameters (Figure S11) and glycosidic linkage dihedral free energies $\Delta G(\phi, \psi)$ (Figure S12) of constructed conformations post-minimization matched those of MD-generated ensembles (i.e., construction algorithm input data; Figures 3a,b and 5a,b and Table 2), indicating that dihedral angles were properly restrained and that minimization did not change the overall ring and linkage conformation.

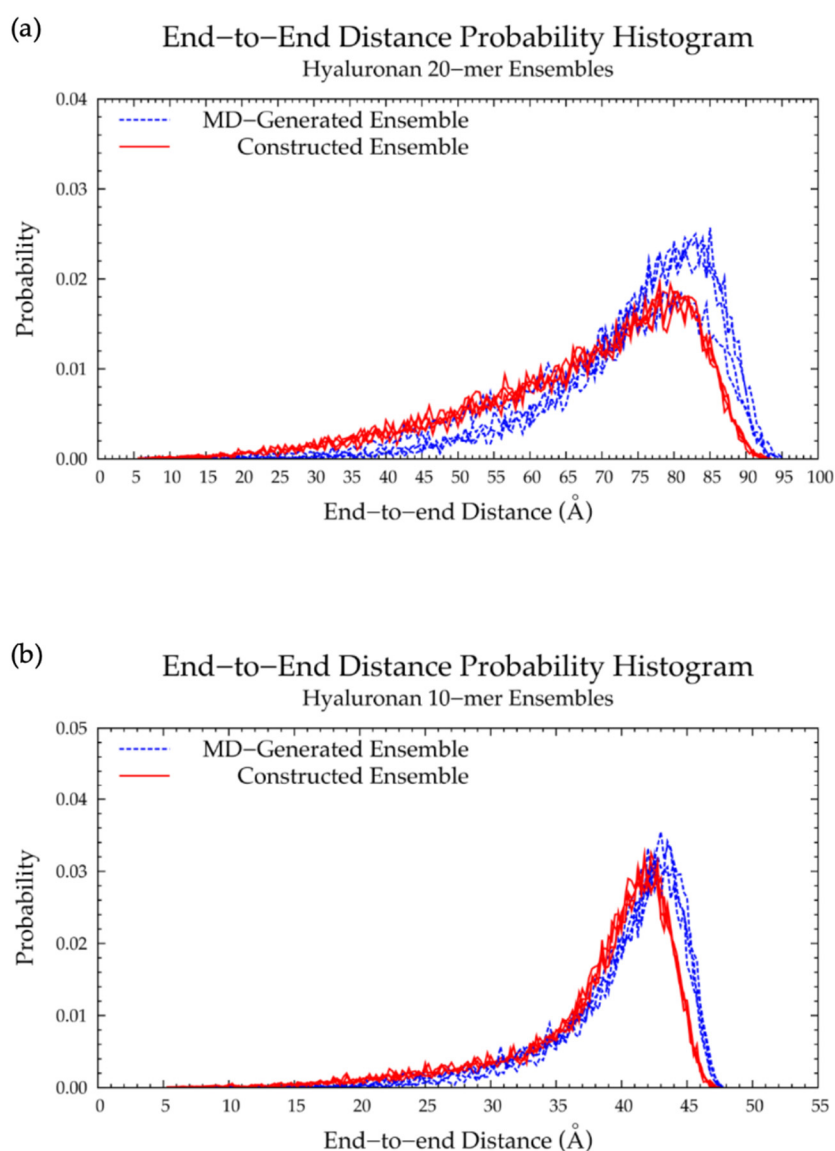


Figure 6. End-to-end distance probability distributions of MD-generated (blue dashed lines) and constructed (red solid lines) hyaluronan ensembles: (a) 20-mer and (b) 10-mer; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer and 0.25 Å bins for the 10-mer; each ensemble includes four sets of 10,000 conformations.

A hyaluronan 10-mer ensemble with 40,000 conformations was also constructed using the algorithm; similarly, the end-to-end distances and radii of gyration of the constructed 10-mer

ensemble matched those of the MD-generated 10-mer ensemble (Figures 6b and S2c,d and Table 1). Both ensembles contained mostly extended conformations, as expected for short oligosaccharides with fewer opportunities for kinks or curves.

Finally, a hyaluronan 200-mer ensemble containing four sets of 10,000 conformations was constructed using the algorithm. The construction procedure produced end-to-end distances, radii of gyration, and bond potential energies of all conformations, as well as PDB files of conformations with most probable end-to-end distances, bond potential energies above that of the fully-extended conformation, and all excluded conformations (for a total of about 1000 PDBs for each of the four sets) for visualization and validation of the bond potential energy cutoff. As seen in MD-generated hyaluronan 10- and 20-mer ensembles, we expected the end-to-end distance distribution curve skewness to shift toward the right (i.e., lower end-to-end distances indicating more compact conformations) with increasing polymer length, as there were more opportunities for kinks and curves resulting from ring puckering and the flexibility of the glycosidic linkages. This pattern was seen in the end-to-end distance distributions of the hyaluronan 200-mer constructed ensembles (Figure 7) and in those of nonsulfated chondroitin 100- and 200-mer constructed ensembles [33]. The relationship between end-to-end distance and radius of gyration was decreasingly linear with increasing polymer length. In other words, conformations with the same end-to-end distance have a wider range of radii of gyration in longer polymers. This was evidenced by comparison of R^2 values of the end-to-end distance vs. radius of gyration regression lines of hyaluronan 20- and 10-mer ensembles, both MD-generated and constructed (Figure S2a–d). As expected, the end-to-end distance and radius of gyration relationship was decreasingly linear with increasing polymer length, as shown by comparison of hyaluronan 10-, 20-, and 200-mer constructed ensembles (Figure S2b,d,e). This suggests that our algorithm constructs hyaluronan 200-mer ensembles with the backbone conformations that we would expect to see in MD simulations.

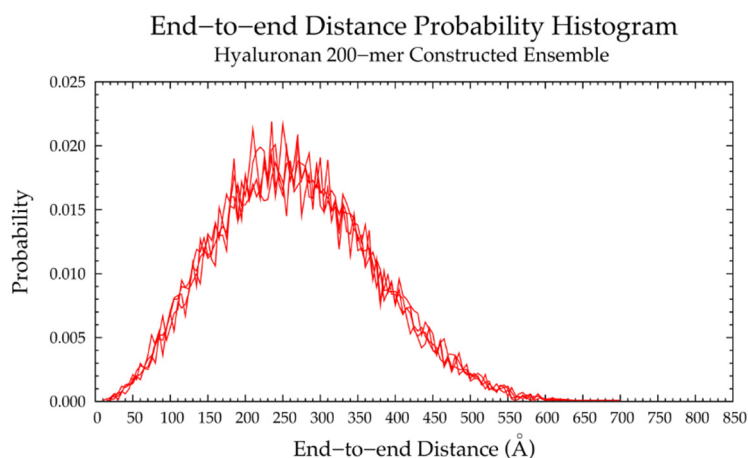


Figure 7. End-to-end distance probability distribution of constructed ensemble of hyaluronan 200-mer; most probable end-to-end distance across all four sets is 235 Å; probabilities were calculated for end-to-end distances sorted into 5 Å bins; the ensemble contains four sets of 10,000 conformations.

3.2. Nonsulfated Dermatan

3.2.1. Molecular Dynamics Simulations: Glycosidic Linkage and Monosaccharide Ring Geometry Effects on Polymer Backbone Flexibility

MD simulations of nonsulfated dermatan 20-mer revealed relatively rigid and linear backbone conformations, as evidenced by the narrow and highly left-skewed end-to-end distance distributions, which match in all four MD runs (Figure 8 and Table 3), and the linear relationship between end-to-end distance and radius of gyration (Figure S13a). To understand the factors contributing to this rigid

behavior of the dermatan 20-mer in simulation, conformations of monosaccharide rings and glycosidic linkages were analyzed.

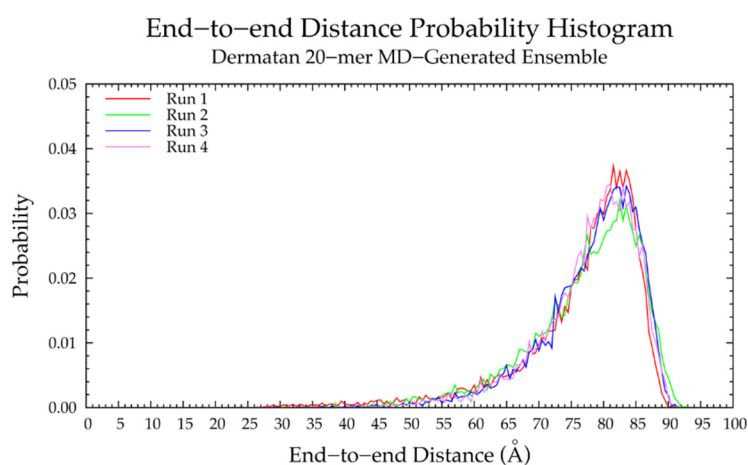


Figure 8. End-to-end distance probability distribution of MD-generated nonsulfated dermatan 20-mer ensemble; each of the four runs includes 10,000 conformations; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins.

Table 3. Most Probable End-to-End Distances (d) in MD-Generated and Constructed Nonsulfated Dermatan Ensembles ¹.

	20-mer Ensembles			10-mer Ensembles		
	MD-Generated d (Å)	Constructed d (Å)	% Difference	MD-Generated d (Å)	Constructed d (Å)	% Difference
Run 1	81.5	81.5		42.25	42.75	
Run 2	82.5	83.5		42.50	42.00	
Run 3	83.5	84.0		42.00	42.75	
Run 4	81.0	81.5		42.25	42.75	
All ²	83.5	81.5	2.42%	42.25	42.75	1.18%

¹ Probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer ensembles and 0.25 Å bins for the 10-mer ensembles. ² All = end-to-end distance distribution aggregated across all four runs.

GalNAc monosaccharide ring C-P parameters (Figure 9a) show mostly ⁴C₁ conformations, as well-established by X-ray diffraction [49,106,118], NMR [47,119], and force field [33,47,49,107] data. Biased MD simulations of β-GalNAc monosaccharides showed that nonsulfated β-GalNAc sampled boat and skew-boat conformations, namely B_{3,O}, ¹S₃, and ^{1,4}B, with relatively high free energies [107]. In line with this study, our simulations showed very few boat and skew-boat puckers, namely ¹S₃, ^{1,4}B, ¹S₅ (Figure 4a; -3GalNAcβ1- endocyclic ring and linker oxygen atoms were identical to those of -3GlcNAcβ1-), and B_{2,5}, on the microsecond timescale. These puckers were sampled by different rings of the 20-mer in different MD runs and returned to ⁴C₁ within 10 ns (Figure S14), suggesting that this behavior was random, and confirming that non-⁴C₁ GalNAc puckers are not energetically favorable. While these boat and skew-boat conformations caused a kink in the polymer chain, there were so few occurrences of this that they did not impact the overall end-to-end distance distribution.

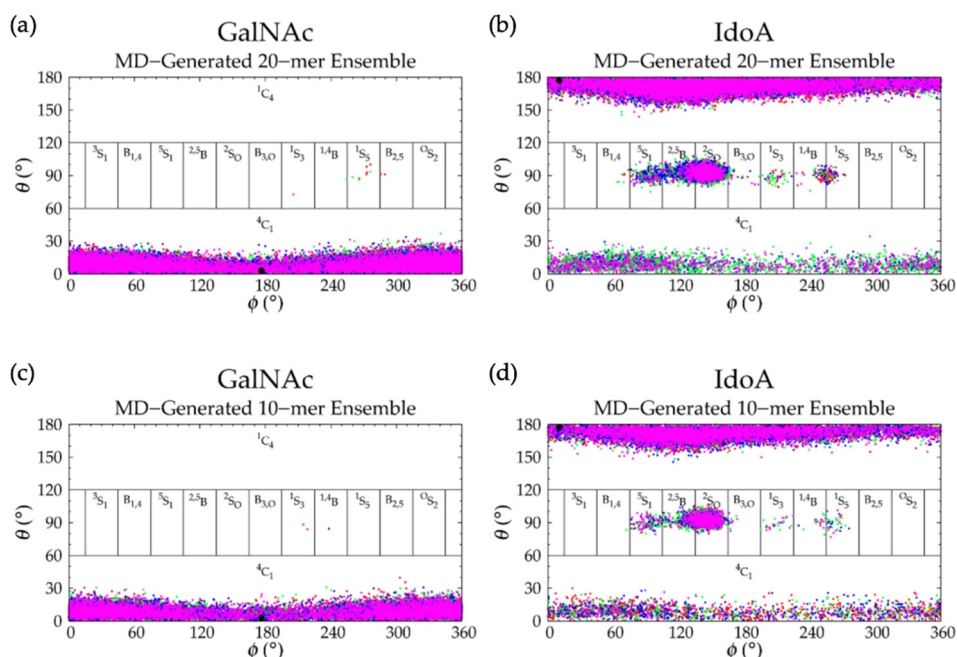


Figure 9. Cremer–Pople data for MD-generated nonsulfated dermatan ensembles: (a) GalNAc and (b) IdoA in the 20-mer and (c) GalNAc and (d) IdoA in the 10-mer; geometries from the four sets of each ensemble are represented by red, green, blue, and magenta dots, respectively and the force-field geometry is represented by a single large black dot; Cremer–Pople parameters (ϕ , θ) for all rings in every tenth snapshot from each ensemble were plotted (i.e., 10 rings \times 1000 snapshots per run \times 4 runs = 40,000 parameter sets for the 20-mer and 20,000 parameter sets for the 10-mer).

IdoA monosaccharide rings in 20-mer MD sampled a majority of 1C_4 and, to a lesser degree, 2S_0 and 4C_1 puckers (Figure 9b and Table S2), as observed in NMR and force field studies [29,45,47,51,107,120–125]. Literature reports of the relative proportions of these three ring puckers in nonsulfated IdoA vary, which may be explained by differences in the structure of neighboring residues (or lack thereof) [45,120,123,124] and differences in ion concentrations [123]. Furthermore, existing studies were conducted on IdoA monosaccharides and short IdoA-containing GAG oligosaccharides (i.e., <20 monosaccharides). NMR data have shown that, when at the nonreducing terminal in GAG oligosaccharides, nonsulfated IdoA is in equilibrium with 1C_4 , 2S_0 , and 4C_1 puckers, and when flanked by nonsulfated N-acetylated sugar derivatives (i.e., GalNAc or GlcNAc), nonsulfated IdoA is primarily in 1C_4 and 2S_0 conformations [123], with fewer than 10% of samples in 4C_1 [45,124]. Indeed, we found this to be the case in our dermatan 20-mer MD simulations (Table S2). Importantly, none of the 1C_4 , 2S_0 , or 4C_1 puckers introduced a kink in the polymer chain, so we would not expect variations in the relative proportions of these ring puckers to alter overall polymer backbone conformations. IdoA also sampled other boat and skew-boat conformations in 20-mer MD simulations, specifically $B_{1,4}$, 5S_1 , ${}^{2,5}B$, $B_{3,O}$, 1S_3 , ${}^{1,4}B$, and 1S_5 (Figure 9b and Table S2), which is in agreement with results from MD simulations of IdoA monosaccharides [107,126]. Another force field study showed that ${}^{2,5}B$ and $B_{3,O}$ IdoA conformers are also feasible interpretations of existing NMR data [127]. Furthermore, a force field study that mapped X-ray diffraction and NMR data for dermatan sulfate [122] and a molecular modeling study of IdoA monosaccharides [126] showed that interconversion between different boat/skew-boat forms of IdoA is much more common than that between boat/skew-boat and chair forms, as it is less energetically costly, which helps explain the sampling of multiple different boat/skew-boat IdoA conformations. As each of these boat/skew-boat puckers (except for 2S_0) introduces a kink in the polymer chain (Figures 10 and 4b,c; -4IdoA α 1-endocyclic ring and linker oxygen atoms are identical to those of -4GlcAc β 1-), end-to-end distances of 20-mer conformations with these puckers were analyzed to determine if they were associated

with more compact conformations. The end-to-end distance distribution of conformations with boat/skew-boat ring puckers (other than 2S_O) matched that of the average of the four runs from the full MD-generated 20-mer ensemble (Figure S15), indicating that these ring puckers do not necessarily give compact polymer conformations. Additionally, the end-to-end distance distribution of 20-mer conformations with 2S_O IdoA conformations was compared to that of the full MD ensemble (Figure S15). These distributions were similar, further suggesting that 2S_O conformations are not associated with more or less compact 20-mer conformations. To determine if IdoA ring puckering in nonsulfated dermatan 20-mer MD was random, an analysis was performed of C-P parameters of individual IdoA monosaccharides in each of the four MD runs (Figure S16). This revealed that (1) not all IdoA rings overcame the energy barrier to sample a 4C_1 pucker, (2) no single IdoA ring sampled 4C_1 in all four MD runs, (3) each IdoA ring overcame the energy barrier to sample boat/skew-boat puckers (i.e., C-P $\theta \sim 90^\circ$) in at least one MD run, and (4) after sampling boat/skew-boat puckers, some IdoA rings returned to 1C_4 in as little as 165 ns, while others remained in boat/skew-boat conformations for the remainder of the 1- μ s trajectory. These observations support the idea that monosaccharide rings in a nonsulfated dermatan 20-mer behave randomly and independently in unbiased MD simulation.

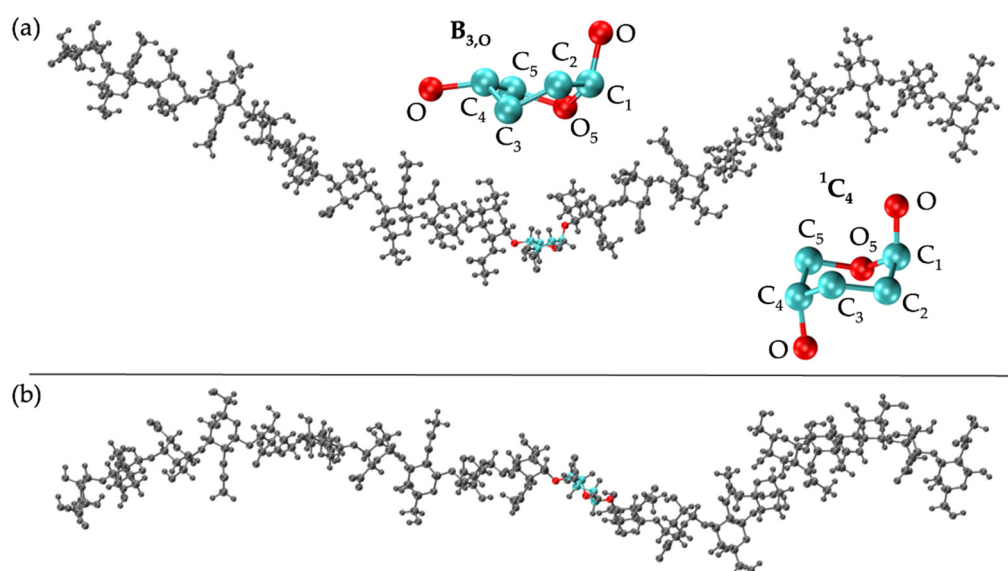


Figure 10. MD-generated nonsulfated dermatan conformations: (a) 20-mer with a $B_{3,O}$ IdoA conformer (ring atoms in cyan; linkage atoms in red) and $B_{3,O}$ IdoA and 1C_4 IdoA monosaccharide conformations (endocyclic ring atoms and linker oxygen atoms only); (b) 20-mer with a 2S_O IdoA conformer highlighted (ring atoms in cyan; linkage atoms in red; 2S_O α -IdoA is identical to 2S_O β -GlcA in Figure 4c); in both 20-mer conformations, all IdoA monosaccharides in gray are in 1C_4 and all GalNAc monosaccharides are in 4C_1 .

Analysis of dihedral free energies revealed a single $\Delta G(\phi, \psi)$ minimum for each of IdoA α 1-3GalNAc and GalNAc β 1-4IdoA glycosidic linkages (Figure 11a,b and Table 4). This is consistent across all linkages in all runs for each linkage type. The lack of secondary basins in ϕ, ψ dihedral samples may explain the higher degree of rigidity and tendency toward extended backbone conformations of the nonsulfated dermatan 20-mer in comparison to hyaluronan and nonsulfated chondroitin [33] 20-mers in MD simulation. An X-ray diffraction study of sodium dermatan sulfate observed three different helical forms (i.e., right-handed, left-handed, and achiral) [118] and quantified their dihedral angles, which were defined by: $\phi = O_5-C_1-O-C_n$ and $\psi = C_1-O-C_n-C_{(n+1)}$. Assuming $C_1-O-C_3-C_2 = C_1-O-C_3-C_4 + 120^\circ$ and $C_1-O-C_4-C_3 = C_1-O-C_4-C_5 - 120^\circ$, we see that the right-handed and achiral helical forms corresponded to high $\Delta G(\phi, \psi)$ values in our nonsulfated dermatan 20-mer MD simulations, and the left-handed helical form corresponded to glycosidic linkage conformations very close to the $\Delta G(\phi, \psi)$ global minimum. In line with our results, force field studies that compared MD results to existing

NMR and X-ray diffraction data of dermatan sulfate dismissed right-handed and achiral helical forms and confirmed the presence of left-handed helical structure [49,122]. A quantitative comparison of our MD-generated glycosidic linkage data to their findings and other NMR data for dermatan tetrasaccharides [47] showed close agreement if we assume $O_5-C_1-O-C_n = H_1-C_1-O-C_n - 120^\circ$ for all ϕ values, $C_1-O-C_3-C_2 = C_1-O-C_3-H_3 - 120^\circ$ for α 1-3 linkage ψ values, and $C1-O-C_4-C_3 = C_1-O-C_4-H_4 + 120^\circ$ for β 1-4 linkage ψ values.

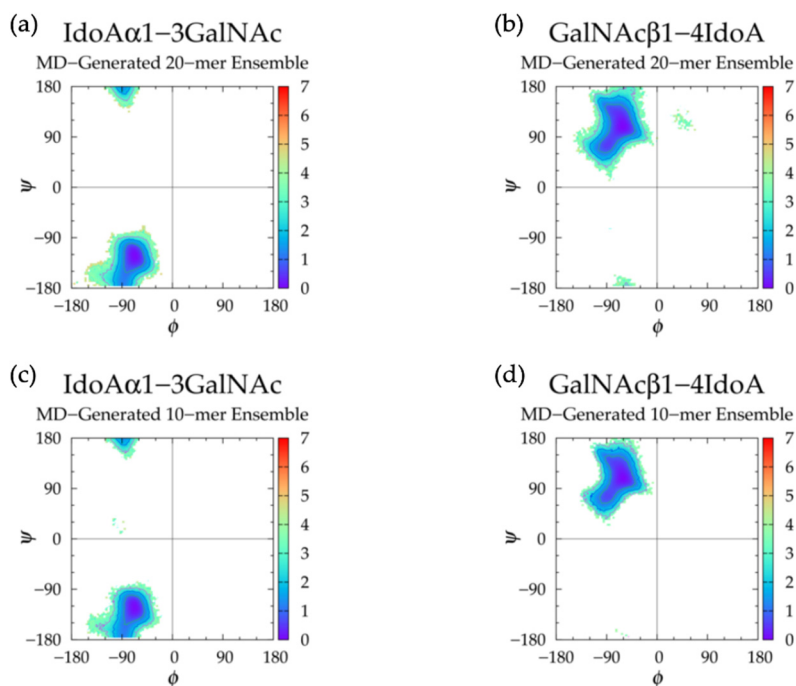


Figure 11. $\Delta G(\phi, \psi)$ in the MD-generated nonsulfated dermatan ensembles for aggregated IdoA α 1-3GalNAc and GalNAc β 1-4IdoA glycosidic linkage data in the (a,b) 20-mer and (c,d) 10-mer, respectively; contour lines every 1 kcal/mol.

Table 4. Nonsulfated Dermatan Glycosidic Linkage Dihedrals (ϕ, ψ)¹ and Free Energy ($\Delta G(\phi, \psi)$ kcal/mol) Minima.

	MD-Generated 20-mer Ensemble				MD-Generated 10-mer Ensemble				Constructed 20-mer Ensemble			
	IdoA β 1-3 GalNAc		GalNAc β 1-4 IdoA		IdoA β 1-3 GalNAc		GalNAc β 1-4 IdoA		IdoA β 1-3 GalNAc		GalNAc β 1-4 IdoA	
Min	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG
I	-66.25°, -123.75°	0.00	-61.25°, 106.25°	0.00	-68.75°, -121.25°	0.00	-58.75°, 108.75°	0.00	-68.75°, -121.25°	0.00	-61.25°, 108.75°	0.00

¹ ϕ, ψ dihedral angles were sorted into 2.5° bins.

MD simulations of a nonsulfated dermatan 10-mer produced mostly linear backbone conformations (Figure S13c) and similar monosaccharide ring (Figure 9) and glycosidic linkage (Figure 11 and Table 4) geometries to the nonsulfated dermatan 20-mer. The major difference was that in 10-mer MD simulations, 9% of all IdoA rings in all runs were found in 4C_1 form; this was almost exclusively in the nonreducing terminal IdoA (i.e., ring #10; Figure S17 and Table S2), which is mostly in line with the aforementioned NMR study of IdoA ring conformations [45]. However, based on the findings from the NMR study and from our dermatan 20-mer MD simulations, we would still expect the nonterminal IdoA rings in a GAG polymer to sample some 4C_1 conformations. This supports our position that GAG 20-mers are better-suited for predictions of long-chain backbone conformations than short GAG oligosaccharides.

These findings suggest that (1) monosaccharide rings and glycosidic linkages in nonsulfated dermatan GAG 20-mers behave randomly and independently in MD simulation, (2) nonsulfated dermatan polymers take on rigid left-handed helical structure with a tendency toward linear backbone conformations in unbiased MD simulations, and (3) nonsulfated dermatan 20-mer conformations in MD simulation provide a realistic representation of longer nonsulfated dermatan polymer conformations.

3.2.2. Construction Algorithm

Four sets of 10,000 conformations of a nonsulfated dermatan 20-mer were constructed using our algorithm and compared to conformations seen in nonsulfated dermatan 20-mer MD. As expected, the C-P parameter plots (Figure S18) and glycosidic linkage free energies (Figure S19) in the constructed 20-mer ensemble matched those in the MD-generated 20-mer ensemble (Figures 9a,b and 11a,b and Table 4). Furthermore, the end-to-end distances (Figure 12a) and radii of gyration (Figure S13a,b) in the constructed ensemble matched those in the MD-generated ensemble, with only a 2.42% difference in most probable end-to-end distance (Table 3). This demonstrates that our algorithm produces nonsulfated dermatan 20-mer ensembles that mimic 20-mer backbone flexibility seen in MD simulations.

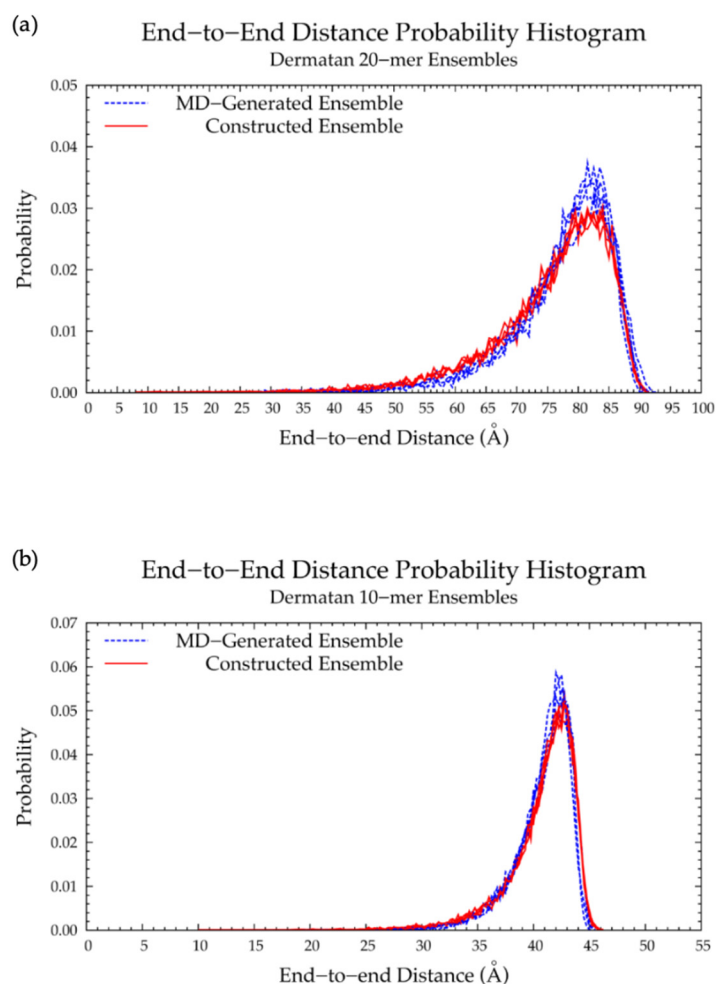


Figure 12. End-to-end distance probability distributions of MD-generated (blue dashed lines) and constructed (red solid lines) nonsulfated dermatan ensembles: (a) 20-mer and (b) 10-mer; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer and 0.25 Å bins for the 10-mer; each ensemble includes four sets of 10,000 conformations.

The nonsulfated dermatan 10-mer ensemble with 40,000 conformations constructed by the algorithm had very similar end-to-end distance and radius of gyration data compared to the MD-generated ensemble (Figures 12b and S13c,d and Table 3). The algorithm was also used to create a nonsulfated dermatan 200-mer ensemble. As expected, the 200-mer end-to-end distance distribution (Figure 13) and radii of gyration (Figure S13e) showed that conformations tended to be more compact as polymer length increased. The shift in skewness of the end-to-end distance distribution with increasing polymer length was more subtle in nonsulfated dermatan than in hyaluronan, as there were fewer linkage and ring conformations causing kinks in the polymer.

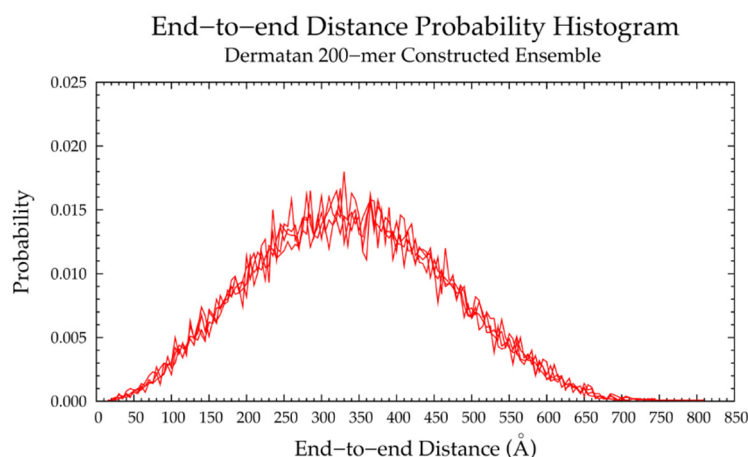


Figure 13. End-to-end distance probability distribution of constructed ensemble of nonsulfated dermatan200-mer; most probable end-to-end distance across all four sets is 365 Å; probabilities were calculated for end-to-end distances sorted into 5 Å bins; the ensemble contains four sets of 10,000 conformations.

3.3. Nonsulfated Keratan

3.3.1. Molecular Dynamics Simulations: Glycosidic Linkage and Monosaccharide Ring Geometry Effects on Polymer Backbone Flexibility

In MD simulations, nonsulfated keratan 20-mers showed rigid backbone behavior and favored extended backbone conformations, as evidenced by end-to-end distance distributions (Figure 14 and Table 5) which were identical in all four MD runs, and the high correlation of end-to-end distance to radius of gyration (Figure S20a). To understand the causes of this rigid behavior, monosaccharide ring and glycosidic linkage conformations were analyzed.

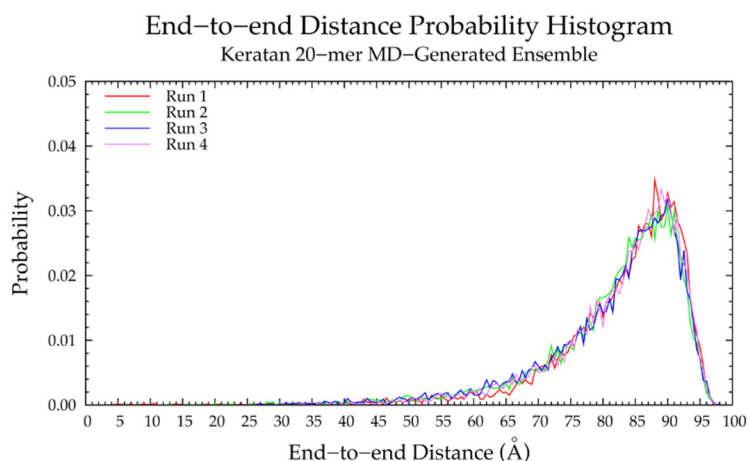


Figure 14. End-to-end distance probability distribution of MD-generated nonsulfated keratan 20-mer ensemble; each of the four runs includes 10,000 conformations; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins.

Table 5. Most Probable End-to-End Distances (d) in MD-Generated and Constructed Nonsulfated Keratan Ensembles ¹.

	20-mer Ensembles			10-mer Ensembles		
	MD-Generated d (Å)	Constructed d (Å)	% Difference	MD-Generated d (Å)	Constructed d (Å)	% Difference
Run 1	88.0	86.0		45.50	45.00	
Run 2	90.0	87.5		46.00	45.75	
Run 3	90.0	87.0		45.75	45.00	
Run 4	89.0	86.0		45.50	45.00	
All ²	90.0	87.0	3.39%	45.50	45.00	1.10%

¹ Probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer ensembles and 0.25 Å bins for the 10-mer ensembles. ² All = end-to-end distance distribution aggregated across all four runs.

C-P parameters of GlcNAc monosaccharides in keratan 20-mer MD (Figure 15a) showed similar conformations to GlcNAc in hyaluronan 20-mer. There were predominantly ⁴C₁ chair conformations, which are also found in a keratan sulfate tetrasaccharide crystal structure [128], with very few transitions (i.e., no more than one per ring or two per run) to boat/skew-boat conformations (i.e., ²S_O, ¹S₃, ^{1,4}B, and ¹S₅), which were sampled for no more than 20 ns before returning to ⁴C₁ chair (Figure S21). Analysis of end-to-end distance and radius of gyration of 20-mer conformations containing non-⁴C₁ ring puckers showed that these ring puckers were not associated with compact 20-mer conformations (Figure S22). C-P analysis of Gal monosaccharides showed that for the duration of the keratan 20-mer MD simulations, Gal remained in ⁴C₁ chair (Figure 15b), which was the predominant conformation found in NMR [119,129–131] and force field data [107] for Gal in general, and in a keratan sulfate tetrasaccharide crystal structure [128].

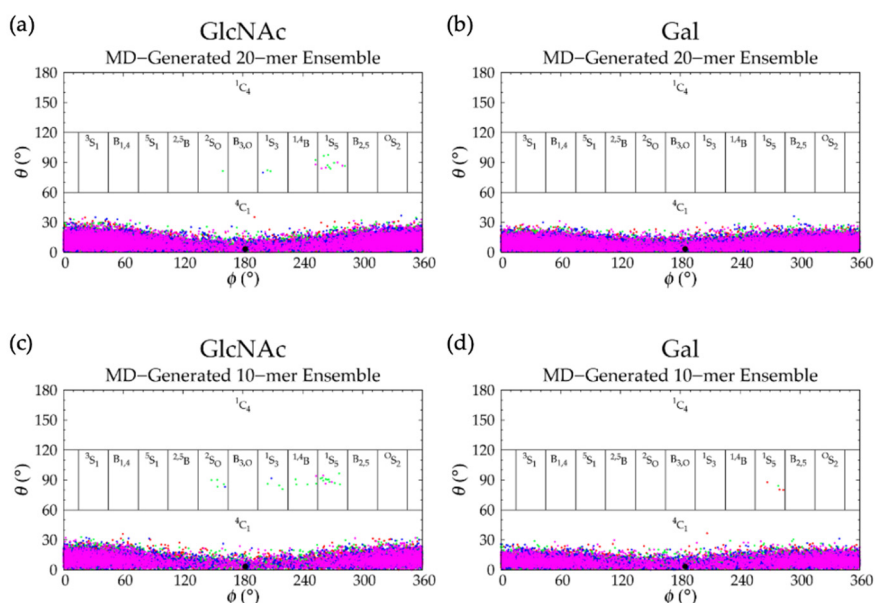


Figure 15. Cremer–Pople data for MD-generated nonsulfated keratan ensembles: (a) GlcNAc and (b) Gal in the 20-mer and (c) GlcNAc and (d) Gal in the 10-mer; geometries from the four sets of each ensemble are represented by red, green, blue, and magenta dots, respectively and the force-field geometry is represented by a single large black dot; Cremer–Pople parameters (ϕ , θ) for all rings in every tenth snapshot from each ensemble were plotted (i.e., 10 rings \times 1000 snapshots per run \times 4 runs = 40,000 parameter sets for the 20-mer and 20,000 parameter sets for the 10-mer).

Glycosidic linkage free energy $\Delta G(\phi, \psi)$ analysis of Gal β 1-4GlcNAc linkages revealed a global minimum in the basin with $-\phi$, $+\psi$ and a secondary minimum in the basin with $-\phi$, $-\psi$ (Figure 16a and Table 6), similar to hyaluronan GlcNAc β 1-4GlcA linkages (Figure 5b,d and Table 2), but with additional rare conformations in $+\phi$, $+\psi$. A molecular modeling study, which was shown to agree with experimental data, showed that β 1-4 linkages in keratan sulfate disaccharides took on the same conformations as β 1-4 linkages in hyaluronan disaccharides [117], confirming the primary and secondary basins sampled in our 20-mer MD simulations. Keratan 20-mer conformations with glycosidic linkage conformations in the tertiary basin (i.e., $+\phi$, $+\psi$) were visualized and end-to-end distances were analyzed (Figure S23). These conformations formed a slight bend in the polymer chain and were associated with more compact polymer backbone conformations. However, there did not appear to be any close contacts between atoms of adjacent monosaccharides in this conformation or in previous snapshots that could explain these linkage bond rotations. Furthermore, this conformation was not specific to any particular region of the polymer. Therefore, this behavior was likely random. As $-\phi$, $-\psi$ β 1-4 linkage conformations are associated with compact hyaluronan polymer conformations, and nonsulfated keratan 20-mer is relatively rigid and favors extended conformations in contrast to hyaluronan, we sought to determine the effects of Gal β 1-4GlcNAc linkage conformations in the secondary (i.e., $-\phi$, $-\psi$) basin on keratan 20-mer backbone flexibility in MD simulations. As in hyaluronan, these conformations caused a kink in the keratan 20-mer chain and were associated with more compact conformations (Figure S23 and Table S1). However, these conformations were slightly less stable in keratan 20-mer MD (Table 6) than in hyaluronan MD (Table 2), and were, therefore, less common.

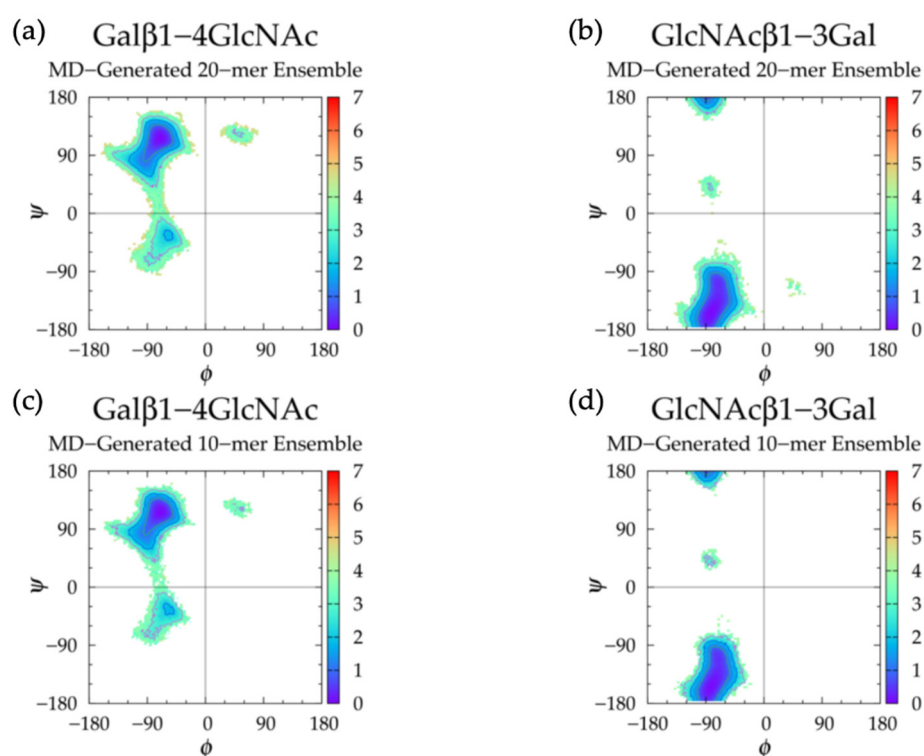


Figure 16. $\Delta G(\phi, \psi)$ in the MD-generated nonsulfated keratan ensembles for aggregated Gal β 1-4GlcNAc and GlcNAc β 1-3Gal glycosidic linkage data in the (a,b) 20-mer and (c,d) 10-mer, respectively; contour lines every 1 kcal/mol.

Table 6. Nonsulfated Keratan Glycosidic Linkage Dihedrals (ϕ, ψ)¹ and Free Energy ($\Delta G(\phi, \psi)$ kcal/mol) Minima.

	MD-Generated 20-mer Ensemble			MD-Generated 10-mer Ensemble			Constructed 20-mer Ensemble					
	Gal β 1-4 GlcNAc		GlcNAc β 1-3 Gal	Gal β 1-4 GlcNAc		GlcNAc β 1-3 Gal	Gal β 1-4 GlcNAc		GlcNAc β 1-3 Gal			
Min	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG		
I	-73.75°, 116.25°	0.00	-81.25°, -156.25°	0.00	-71.25°, 116.25°	0.00	-81.25°, -156.25°	0.00	-71.25°, 118.75°	0.00	-81.25°, -158.75°	0.00
II	-56.25°, -31.25°	1.75			-56.25°, -36.25°	1.67			-53.75°, -33.75°	1.75		

¹ ϕ, ψ dihedral angles were sorted into 2.5° bins.

Analysis of GlcNAc β 1-3Gal glycosidic linkages showed a global free energy minimum in the $-\phi, -\psi$ basin (Table 6) similar to that of IdoA β 1-3GalNAc linkages in nonsulfated dermatan MD (Figure 11a,c and Table 4), but with additional rare occurrences in a secondary basin in $-\phi, +\psi$, confirmed by the aforementioned molecular modeling study [117], and a small tertiary basin in $+\phi, -\psi$ (Figure 16b). To determine the effects of GlcNAc β 1-3Gal linkage conformations in these secondary and tertiary basins on backbone conformation, we visualized 20-mer conformations containing these glycosidic linkage conformations and analyzed end-to-end distances (Figure S24). Glycosidic linkage conformations in the tertiary basin caused a kink and were associated with more compact 20-mer conformations. In contrast to hyaluronan, nonsulfated keratan β 1-3 glycosidic linkage conformations in the secondary basin (i.e., $-\phi, +\psi$) caused only a slight bend in the polymer chain, but were still associated with compact 20-mer conformations (Table S1). Both secondary and tertiary GlcNAc β 1-3Gal linkage conformations were rare in keratan 20-mer MD, which helps explain why the keratan 20-mer chain was relatively rigid and favored extended conformations in MD simulation.

Backbone conformational analysis of nonsulfated keratan 10-mer MD revealed that the 10-mer chain was rigid and favored extended conformations (Figure S20c and Table 5), which is similar behavior to that of nonsulfated keratan 20-mer in MD simulation. This stands to reason, as conformations of monosaccharide rings (Figure 15) and glycosidic linkages (Figure 16 and Table 6) in keratan 10-mer MD matched those in keratan 20-mer MD, with the only differences being that (1) Gal briefly (i.e., for <5 ns) sampled a 1S_5 skew-boat conformation once in each of two of the 10-mer MD runs (Figure S25), and (2) GlcNAc β 1-3Gal linkages in the 10-mer did not sample conformations in the tertiary basin (i.e., $+\phi$, $-\psi$) sampled in the 20-mer. Boat/skew-boat conformations including 1S_5 were shown to have a high secondary free energy minimum (i.e., > 4 kcal/mol) in biased MD simulations of Gal monosaccharides [107], indicating that while possible, these conformations are not very stable in Gal monosaccharides. In both biased and unbiased MD simulations of a tetrasaccharide with a central Gal-GlcNAc disaccharide unit, the free energy minimum of Gal boat/skew-boat conformations was much higher (i.e., ~ 7 kcal/mol) [108], suggesting that non- 4C_1 conformations are less stable in Gal linked to other monosaccharides. This could explain why Gal does not sample non- 4C_1 conformations in keratan 20-mer MD, and suggests that these conformations would not be seen in long polymers. Although $+\phi$, $-\psi$ conformations in GlcNAc β 1-3Gal linkages are rare, they are physical conformations [117] that minorly contribute to polymer backbone flexibility. This further supports our belief that conformational landscapes from 20-mer MD are better-suited to construct conformational ensembles of GAGs with biologically-relevant chain lengths than those of short GAG oligosaccharides.

The higher degree of rigidity and probability of extended conformations in nonsulfated keratan MD compared to hyaluronan MD can be explained by the facts that: (1) for each glycosidic linkage type, the conformations that caused a kink in the polymer chain were less stable and, thus, rarer in nonsulfated keratan MD than in hyaluronan MD; and (2) GlcNAc and Gal are mostly rigid 4C_1 chair conformers in nonsulfated keratan whereas GlcNAc and GlcA took on more boat/skew-boat conformations that caused kinks in hyaluronan. Importantly, nonsulfated keratan 10- and 20-mers behaved randomly, and glycosidic linkages and monosaccharide rings behaved independently in MD simulations.

3.3.2. Construction Algorithm

Nonsulfated keratan 20-mer conformational ensembles (four sets of 10,000 conformations) were constructed using our algorithm and compared to MD-generated keratan 20-mer ensembles. The monosaccharide ring (Figure S26) and glycosidic linkage (Figure S27) conformations in the constructed keratan 20-mer ensemble were identical to the input data (i.e., MD-generated keratan 20-mer conformations; Figures 15a,b and 16a,b and Table 6), as expected. The end-to-end distances (Figure 17a and Table 5) and radii of gyration (Figure S20a,b) were similar in constructed and MD-generated ensembles, demonstrating that nonsulfated keratan 20-mer conformational ensembles provide an accurate representation of backbone flexibility in nonsulfated keratan 20-mer MD simulations.

Four sets of nonsulfated keratan 10-mer ensembles with 10,000 conformations each were constructed, and the results were compared to those of nonsulfated keratan 10-mer MD. The end-to-end distances (Figure 17b and Table 5) and radii of gyration (Figure S20c,d) in constructed 10-mer ensembles matched those of MD-generated 10-mer ensembles, demonstrating that MD-generated 20-mer conformations can be applied to construct ensembles of keratan polymers of different lengths that mimic backbone flexibility seen in MD simulation. A nonsulfated keratan 200-mer ensemble was constructed and, as expected, the end-to-end distance distribution (Figure 18) and radii of gyration (Figure S20e) showed that as the polymer length increased, conformations tended to be more compact. Notably, this shift in the end-to-end distance distribution curve was more subtle in keratan than in hyaluronan, which stands to reason, as monosaccharide rings and glycosidic linkages are more rigid and tend to be more extended in keratan 20-mer MD than in hyaluronan 20-mer MD.

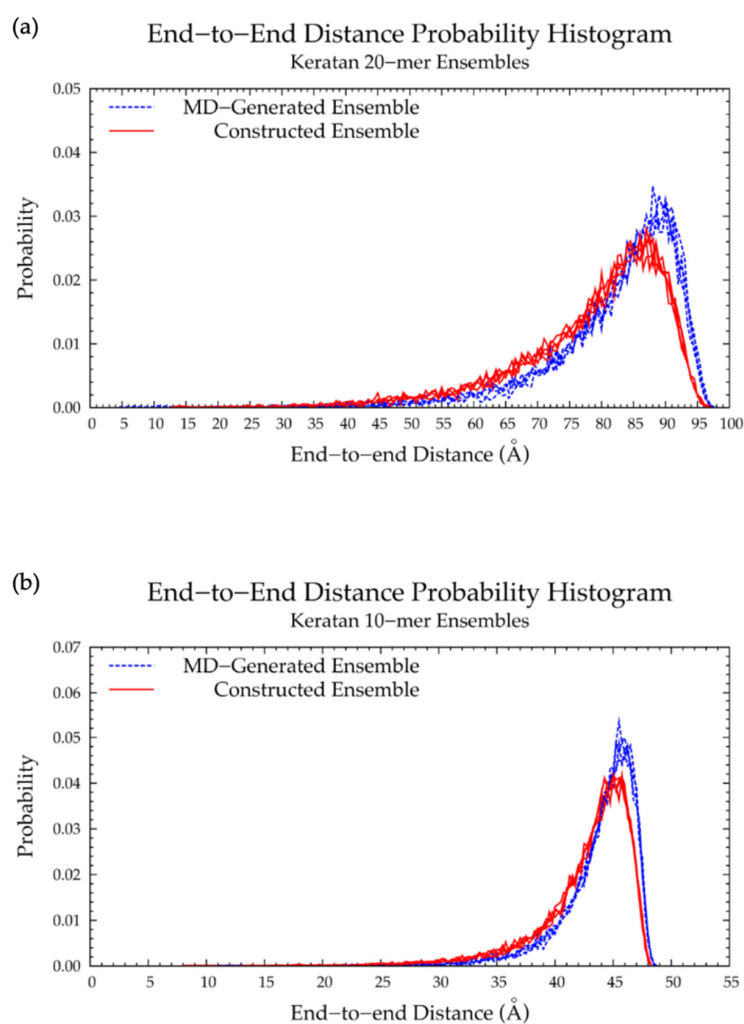


Figure 17. End-to-end distance probability distributions of MD-generated (blue dashed lines) and constructed (red solid lines) nonsulfated keratan ensembles: (a) 20-mer and (b) 10-mer; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer and 0.25 Å bins for the 10-mer; each ensemble includes four sets of 10,000 conformations.

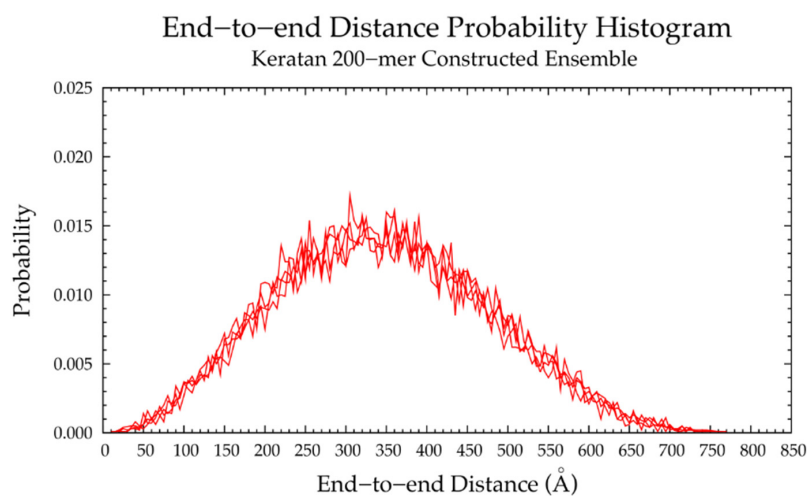


Figure 18. End-to-end distance probability distribution of constructed ensemble of nonsulfated keratan 200-mer; most probable end-to-end distance across all four sets is 305 Å; probabilities were calculated for end-to-end distances sorted into 5 Å bins; the ensemble contains four sets of 10,000 conformations.

3.4. Nonsulfated Heparan

3.4.1. Molecular Dynamics Simulations: Glycosidic Linkage and Monosaccharide Ring Geometry Effects on Polymer Backbone Flexibility

The backbone flexibility of nonsulfated heparan 20-mer, quantified by the end-to-end distances (Figure 19 and Table 7) and radii of gyration (Figure S28a) in MD simulations was analyzed. The wide end-to-end distance distribution curves and tendency toward lower end-to-end distances, relative to those of the extended 20-mer conformation, indicated that nonsulfated heparan 20-mer was highly flexible and tended toward compact conformations in MD simulations. To determine factors contributing to this conformational flexibility, monosaccharide ring and glycosidic linkage conformations were examined for patterns.

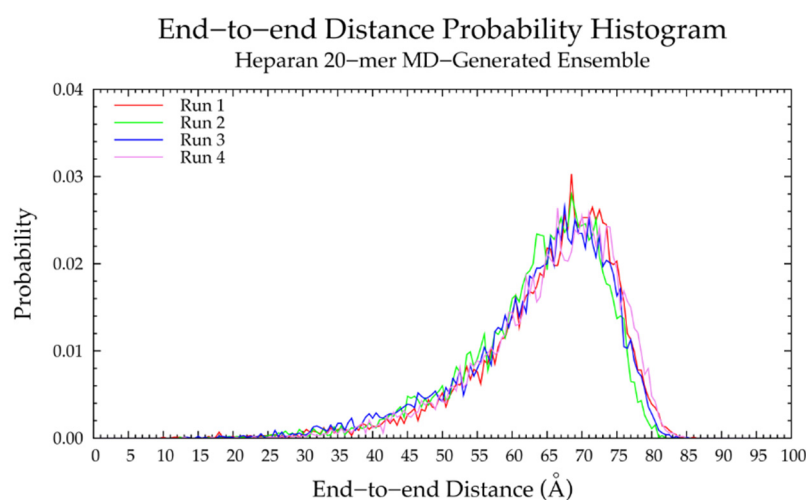


Figure 19. End-to-end distance probability distribution of MD-generated nonsulfated heparan 20-mer ensemble; each of the four runs includes 10,000 conformations; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins.

Table 7. Most Probable End-to-End Distances (d) in MD-Generated and Constructed Nonsulfated Heparan Ensembles ¹.

	20-mer Ensembles			10-mer Ensembles		
	MD-Generated d (Å)	Constructed d (Å)	% Difference	MD-Generated d (Å)	Constructed d (Å)	% Difference
Run 1	68.5	70.0		37.00	37.25	
Run 2	68.5	72.0		36.25	38.25	
Run 3	67.5	67.5		37.50	37.25	
Run 4	66.5	71.5		36.75	36.75	
All ²	68.5	72.0	4.98%	37.00	37.25	0.67%

¹ Probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer ensembles and 0.25 Å bins for the 10-mer ensembles. ² All = end-to-end distance distribution aggregated across all four runs.

C-P parameters of GlcNAc rings (Figure 20a) revealed all ⁴C₁ chair conformations, in line with NMR and force field data for α -GlcNAc [119]. This stands to reason, as we would not expect much difference in ring conformation between α - and β -GlcNAc, because the only structural difference is the orientation of the exocyclic oxygen atom on C₁ (i.e., the linker oxygen on nonterminal GlcNAc monosaccharides in GAG polysaccharides). This structural difference will impact glycosidic linkage conformation but is less likely to impact GlcNAc ring conformation. The IdoA monosaccharide C-P parameters (Figure 20b) from nonsulfated heparan 20-mer MD were similar to those of IdoA in nonsulfated dermatan 20-mer MD (Figure 9b): predominantly ¹C₄, ²S_O, and some ⁴C₁, which is also in

line with NMR and force field data for nonsulfated IdoA in heparan sulfate oligosaccharides [120,123] and nonsulfated heparan trisaccharides [126], and occasional boat and skew-boat conformations (Figure 4b,c; -4IdoA α 1- endocyclic ring and linker oxygen atoms in heparan are identical to those of -4GlcA β 1- in hyaluronan). As with GalNAc in dermatan, the presence of neighboring GlcNAc monosaccharides substantially decreased sampling of 4C_1 conformations [123], which is in line with our results. One NMR and force field study reported 19–49% 2S_0 conformations of nonsulfated IdoA in heparan sulfate hexasaccharides, and found that as the degree of GlcNAc sulfation increased, the percentage of 2S_0 conformations in adjacent nonsulfated IdoA decreased [120]. As we studied only nonsulfated heparan, we would expect to see a higher proportion of IdoA 2S_0 conformations (i.e., close to 49%) if simulated under the same conditions. However, the aforementioned study performed MD in aqueous solution with only neutralizing Na $^+$ ions, whereas our systems contained an additional 140 mM NaCl. Another NMR and molecular modeling study showed that increasing NaCl salt concentrations caused a shift in equilibrium of 1C_4 and 2S_0 toward 1C_4 conformations of 2-O-sulfated IdoA flanked by sulfated glucosamine [123]. This may explain the tendency toward 1C_4 IdoA conformations, with ~54% of IdoA rings across all four 20-mer MD runs in 1C_4 , ~25% in 2S_0 , and ~4% in 4C_1 (Table S3). Furthermore, very few IdoA rings sampled 4C_1 chair conformations (Figure S29), which is in contrast to the more random distribution of IdoA 4C_1 chair conformations in nonsulfated dermatan 20-mer MD (Figure S16). This is likely the result of random kinetic trapping during nonsulfated heparan 20-mer MD simulations, which means IdoA only rarely overcame energy barriers between boat/skew-boat and 4C_1 conformations. Importantly, it appears that all relevant IdoA conformations (i.e., those in line with the literature) were sampled in nonsulfated heparan 20-mer MD simulations. Therefore, we believe our database of MD-generated 20-mer conformations contains a full conformational landscape of IdoA rings in nonsulfated heparan.

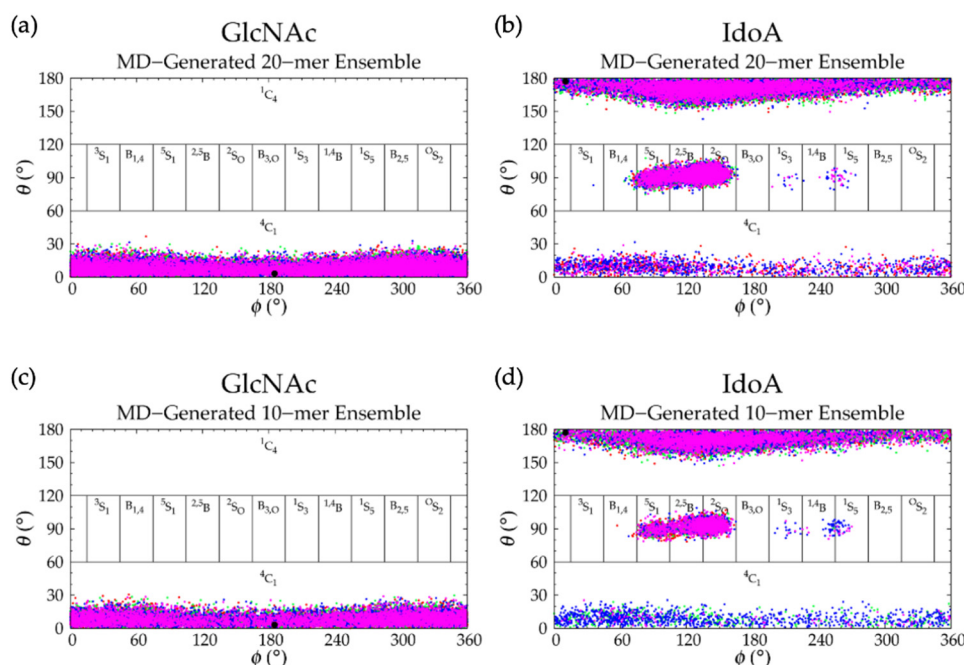


Figure 20. Cremer–Pople data for MD-generated nonsulfated heparan ensembles: (a) GlcNAc and (b) IdoA in the 20-mer and (c) GlcNAc and (d) IdoA in the 10-mer; geometries from the four sets of each ensemble are represented by red, green, blue, and magenta dots, respectively and the force-field geometry is represented by a single large black dot; Cremer–Pople parameters (ϕ , θ) for all rings in every tenth snapshot from each ensemble were plotted (i.e., 10 rings \times 1000 snapshots per run \times 4 runs = 40,000 parameter sets for the 20-mer and 20,000 parameter sets for the 10-mer).

The primary goal of our algorithm is to predict backbone conformations of long GAG chains, which requires that (1) monosaccharide rings in GAG 20-mers behave independently in MD simulation and (2) contributions of ring puckers to backbone flexibility match those expected in nonsulfated heparan polymers in aqueous solution. There did not appear to be any interdependency between adjacent rings, suggesting that monosaccharide rings behave independently in nonsulfated heparan 20-mer MD simulations. To determine the effects of IdoA ring puckering on polymer backbone conformations, end-to-end distances of 20-mer conformations containing monosaccharides in (1) boat and skew-boat conformations that caused a kink in the polymer chain and (2) 2S_0 conformations, were analyzed (Figure S30). The end-to-end distance distribution of 20-mer conformations with the ring puckers that caused a kink was similar to that of the average of the four MD runs and the most probable end-to-end distance was only 1 Å lower in conformations with ring puckers that caused a kink, suggesting that boat and skew-boat conformations that introduce a kink do not necessarily give less compact 20-mer conformations, and thus are not major factors contributing to backbone flexibility. The distribution curve of 20-mer conformations with 2S_0 IdoA ring puckers and that of 20-mer conformations with non- 2S_0 boat/skew-boat puckers that caused a kink were qualitatively similar to that of the full 20-mer MD ensemble, but had additional peaks in probability of higher end-to-end distances. These findings suggest that IdoA boat/skew-boat conformations are associated with the full range of 20-mer backbone conformations in MD simulations with only a slight tendency toward extended 20-mer conformations. Therefore, the proportion of IdoA 2S_0 conformations likely does not have a major effect on nonsulfated heparan 20-mer backbone flexibility.

Next, glycosidic linkage dihedral free energies $\Delta G(\phi, \psi)$ were examined (Figure 21 and Table 8). IdoA α 1-4GlcNAc glycosidic linkages sampled primarily $-\phi, +\psi$ with secondary basins in $-\phi, -\psi$ and GlcNAc α 1-4IdoA glycosidic linkages sampled conformations in a single basin ($+\phi, +\psi$). We compared our results to those of an NMR and force field study that performed two sets of MD on each of nonsulfated IdoA α 1-4GlcNAc and GlcNAc α 1-4IdoA disaccharides, one with IdoA restrained to 1C_4 and the other with IdoA restrained to 2S_0 , each with a biasing potential on glycosidic linkage dihedrals defined by $\phi = H_1-C_1-O-C_4$ and $\psi = C_1-O-C_4-H_4$ [30]. If we assume $O_5-C_1-O-C_4 = H_1-C_1-O-C_4 - 120^\circ$ in IdoA α 1-4GlcNAc linkages, $O_5-C_1-O-C_4 = H_1-C_1-O-C_4 + 120^\circ$ in GlcNAc α 1-4IdoA linkages, and $C_1-O-C_4-C_3 = C_1-O-C_4-H_4 + 120^\circ$ in both linkage types, our data are in close agreement with theirs.

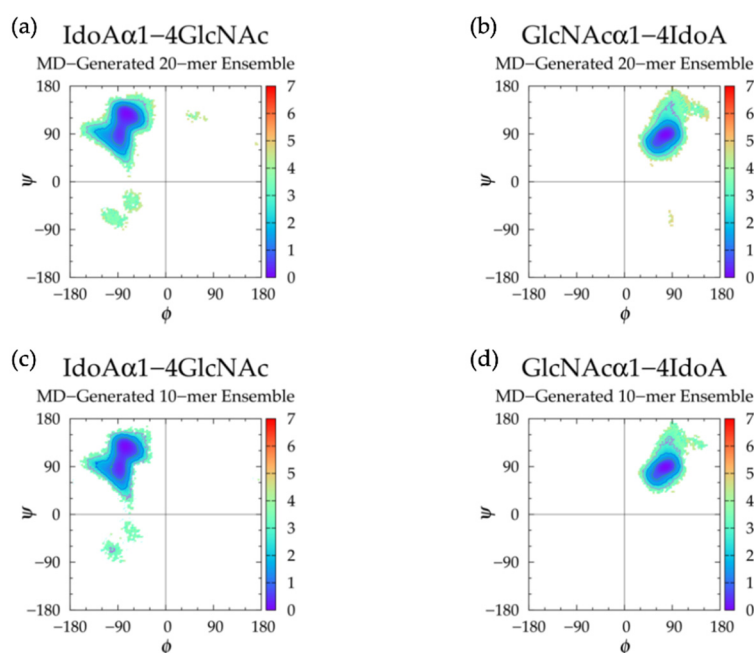


Figure 21. $\Delta G(\phi, \psi)$ in the MD-generated nonsulfated heparan ensembles for aggregated IdoA α 1-4GlcNAc and GlcNAc α 1-4IdoA glycosidic linkage data in the (a,b) 20-mer and (c,d) 10-mer, respectively; contour lines every 1 kcal/mol.

Table 8. Nonsulfated Heparan Glycosidic Linkage Dihedrals (ϕ, ψ)¹ and Free Energy ($\Delta G(\phi, \psi)$ kcal/mol) Minima.

	MD-Generated 20-mer Ensemble				MD-Generated 10-mer Ensemble				Constructed 20-mer Ensemble			
	IdoA α 1-4 GlcNAc		GlcNAc α 1-4 IdoA		IdoA α 1-4 GlcNAc		GlcNAc α 1-4 IdoA		IdoA α 1-4 GlcNAc		GlcNAc α 1-4 IdoA	
Min	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG	ϕ, ψ	ΔG
I	-73.75°, 128.75°	0.00	76.25°, 88.75°	0.00	-76.25°, 128.75°	0.00	73.75°, 88.75°	0.00	-76.25°, 128.75°	0.00	-71.25°, 116.25°	0.00
II	-103.75°, -68.75°	3.13			-103.75°, -63.75°	2.66			-98.75°, -68.75°	3.23		
II'	-66.25°, -31.25°	3.38			-66.25°, -33.75°	2.99			-61.25°, -36.25°	3.34		

¹ ϕ, ψ dihedral angles were sorted into 2.5° bins.

GlcNAc α 1-4IdoA glycosidic linkage conformations differ from any other GAG glycosidic linkage conformation because of the orientation of the linker oxygen atom with respect to GlcNAc. Specifically, the oxygen atom on C₁ of α -GlcNAc is in the opposite orientation as that on C₁ of any other GAG monosaccharide (Figure 1). The GlcNAc α 1-4IdoA linkages have a coiled conformation (Figure S31), which helps explain the high tendency toward compact conformations in nonsulfated heparan compared to other GAGs.

As IdoA α 1-4GlcNAc linkages have secondary conformations, we sought to determine if their behavior was random and independent. $\Delta G(\phi, \psi)$ was plotted for each individual IdoA α 1-4GlcNAc linkage in each MD run and there did not appear to be any connection between adjacent IdoA α 1-4GlcNAc linkages or any patterns across different MD runs. To determine the effects of IdoA α 1-4GlcNAc linkages with $-\phi, -\psi$ dihedrals on 20-mer backbone flexibility, end-to-end distances of 20-mer conformations with these glycosidic linkage conformations were analyzed. Although these linkage conformations caused a kink in the polymer chain (as with hyaluronan GlcNAc β 1-4GlcA linkages), they were not associated with more compact conformations (Figure S32).

In fact, the end-to-end distance distribution of heparan 20-mer conformations with IdoA α 1-4GlcNAc linkage $-\phi$, $-\psi$ dihedrals was similar to that of the full MD-generated 20-mer ensemble. This was likely because these linkage conformations occurred infrequently in nonsulfated heparan 20-mer MD, meaning a single heparan 20-mer conformation was not likely to have many kinks resulting from these $-\phi$, $-\psi$ linkage dihedrals. These findings are in line with the literature, which suggests that IdoA ring conformational flexibility in heparin/heparan sulfate oligo- and polysaccharides does not affect glycosidic linkage conformation or overall backbone shape [132].

To determine if there was any interdependency between different IdoA ring puckers and flanking IdoA α 1-4GlcNAc linkage geometries in 20-mer MD simulations, conformations of all IdoA α 1-4GlcNAc linkages flanking each of 1C_4 , 2S_0 , 4C_1 , and boat/skew-boat (non- 2S_0) conformations were analyzed separately (Figure S33). There did not appear to be strong associations between any particular IdoA ring conformation and flanking IdoA α 1-4GlcNAc linkage conformation, which is in line with IdoA α 1-4GlcNAc disaccharide data from NMR and MD with restrained 1C_4 and 2S_0 IdoA ring conformations and a biasing potential on glycosidic linkage dihedrals [30]. This supports our hypothesis that glycosidic linkages and monosaccharide rings behave independently in MD simulation of nonsulfated heparan 20-mer.

Backbone conformational analysis of nonsulfated heparan 10-mer in MD simulation revealed flexible, compact conformations (Figure S28c and Table 7), as seen in nonsulfated heparan 20-mer MD. Monosaccharide ring and glycosidic linkage conformations in nonsulfated heparan 10-mer MD were similar to those of nonsulfated heparan 20-mer MD. These findings indicate that in MD simulations of nonsulfated heparan 10- and 20-mers, (1) IdoA α 1-4GlcNAc glycosidic linkages with $-\phi$, $-\psi$ dihedrals cause a kink in the polymer chain but are rare and thus do not contribute to compact polymer conformations and (2) monosaccharide rings and glycosidic linkages behave randomly and independently.

3.4.2. Construction Algorithm

A nonsulfated heparan 20-mer ensemble with 40,000 conformations was constructed by the algorithm. Monosaccharide ring C-P parameters (Figure S34) and glycosidic linkage conformations (Figure S35) post-minimization in the constructed ensemble matched the input data (Figures 20a,b and 21a,b), as expected. Additionally, the end-to-end distance distribution and radii of gyration of the constructed ensemble closely resembled those of the MD-generated 20-mer ensemble (Figures 22a and S28a,b and Table 7), demonstrating that our algorithm generates nonsulfated heparan 20-mer ensembles with backbone conformations that mimic backbone flexibility seen in 20-mer MD.

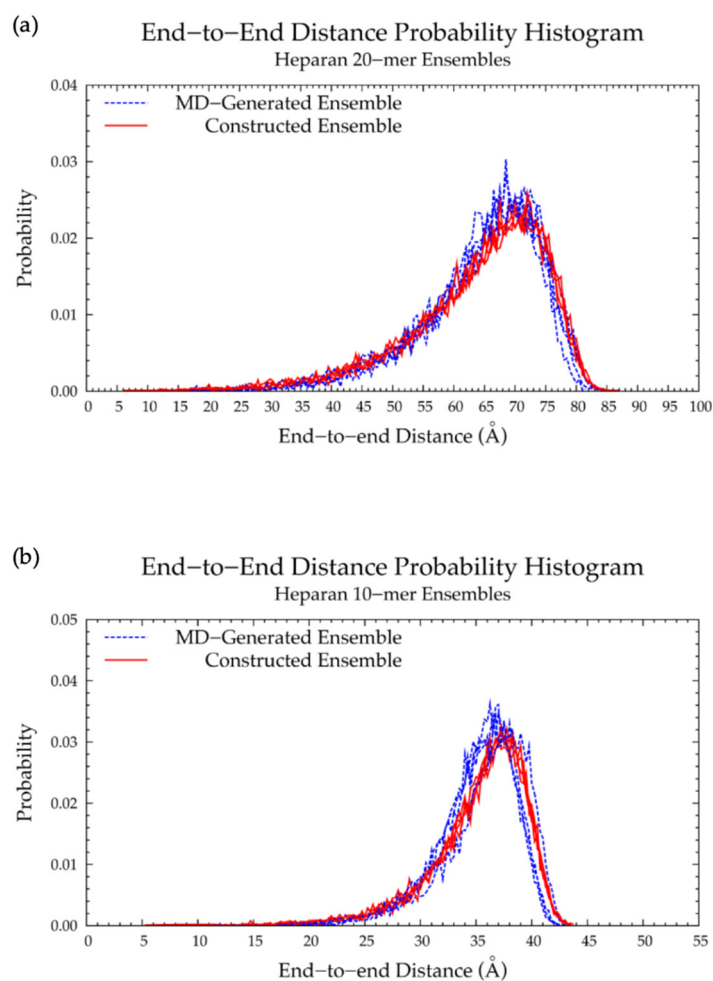


Figure 22. End-to-end distance probability distributions of MD-generated (blue dashed lines) and constructed (red solid lines) nonsulfated heparan ensembles: (a) 20-mer and (b) 10-mer; probabilities were calculated for end-to-end distances sorted into 0.5 Å bins for the 20-mer and 0.25 Å bins for the 10-mer; each ensemble includes four sets of 10,000 conformations.

A nonsulfated heparan 10-mer ensemble with 40,000 conformations was also constructed by the algorithm and had end-to-end distances and radii of gyration that matched those of MD-generated nonsulfated heparan 10-mer ensembles (Figures 22b and S28c,d and Table 7). This demonstrates that nonsulfated heparan 20-mer conformations from MD can be used to construct 10-mer ensembles that mimic the backbone flexibility seen in nonsulfated heparan 10-mer MD.

The constructed nonsulfated heparan 200-mer ensemble had reasonably expected end-to-end distance probability distributions (Figure 23), i.e., the skewness of the curves shifted toward the right, indicating more compact conformations with increasing polymer length. The heparan 200-mer end-to-end distance distribution curve was qualitatively similar to that of hyaluronan 200-mer (i.e., the probability peak was of similar magnitude and most probable end-to-end distances were similar), which stands to reason, as ring and linkage conformations in nonsulfated heparan 20-mer MD cause more compact backbone conformations, as in hyaluronan 20-mer MD.

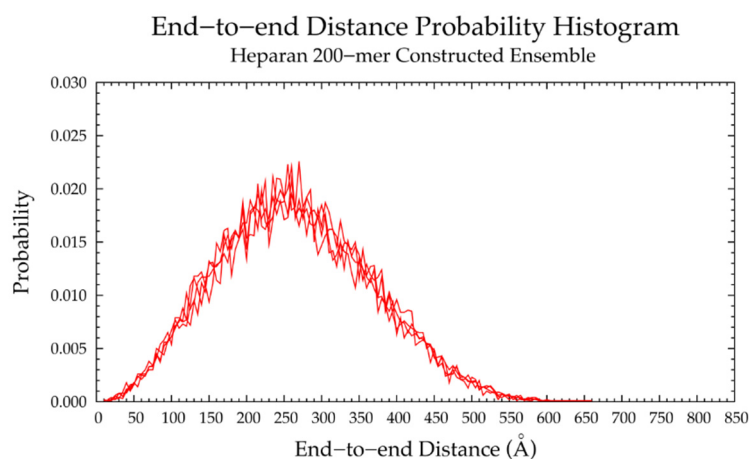


Figure 23. End-to-end distance probability distribution of constructed ensemble of nonsulfated heparan 200-mer; most probable end-to-end distance across all four sets is 260 Å; probabilities were calculated for end-to-end distances sorted into 5 Å bins; the ensemble contains four sets of 10,000 conformations.

4. Conclusions

Collectively, our findings support our hypotheses that (1) glycosidic linkages and monosaccharide rings in hyaluronan and nonsulfated dermatan, keratan, and heparan GAG 20-mers behave randomly and independently in MD simulation and (2) using a database of conformations from corresponding GAG 20-mer MD simulations and treating glycosidic linkages and monosaccharide rings independently, our algorithm can efficiently construct conformational ensembles of hyaluronan and nonsulfated dermatan, keratan, and heparan GAG 10- and 20-mers that mimic backbone flexibility observed in corresponding MD simulations. Additionally, our algorithm constructed sets of 10,000 molecular conformations of nonsulfated GAG 20-mers with backbone conformations that we would reasonably expect to see in MD simulation and did so within 12 h. This suggests that the algorithm can generate conformational ensembles of nonsulfated heterogeneous GAG polymers of arbitrary length in under a day. For perspective, 1- μ s MD simulations, each producing 10,000 snapshots (i.e., 3-D atomic coordinate sets), of GAG 10- and 20-mers were completed in about 1–2 weeks and 1–2 months, respectively, and using modern GPU-accelerated hardware and software. Furthermore, the algorithm's potential energy minimization and bond potential energy cutoff criterion exclude nonphysical conformations from constructed ensembles, leaving only conformations with reasonably expected bond energies (Figures S36–S39).

A comparison of the different GAG 20-mer MD simulations provided further insights into the relationship between GAG structure and conformation. For example, no associations between adjacent monosaccharide conformations were seen in any of the GAG 20-mer MD simulations, but differences in conformations of the same monosaccharide ring type in different GAGs were observed. This indicates that the structure of adjacent monosaccharides contributes to ring conformation. Furthermore, IdoA monosaccharides were much more conformationally flexible than GlcA monosaccharides in GAG 10- and 20-mer MD simulations, even though the structural difference between these two monosaccharide types (specifically, the orientation of the carboxylate group on C₅) is subtle. Interestingly, dermatan and heparan are the only GAGs with IdoA monosaccharides, yet hyaluronan and heparan showed the most conformational flexibility, and heparan showed the greatest tendency toward compact conformations compared to other GAGs. This was likely because of variability in glycosidic linkage conformation, which is independent of monosaccharide ring conformation. Independence of glycosidic linkages and monosaccharide rings was further evidenced by the fact that despite differences in flanking monosaccharide structure in different GAG types, there were similarities between all 1–3 linkage conformations and between all 1–4 linkage conformations, with the exception of GlcNAc α 1–4IdoA linkages in heparan.

A comparison of MD-generated GAG 20-mer conformations to those of GAG 10-mers and to existing experimentally-determined conformations of monosaccharides and GAG oligosaccharides supported the use of 20-mer data for the construction of longer GAG polymers. For example, certain IdoA conformations (namely 4C_1 chair) were found more in unbound monosaccharide rings and terminal rings of GAG oligosaccharides than in central rings. Similarly, galactose occasionally sampled non- 4C_1 conformations in monosaccharide rings and short oligosaccharides, including nonsulfated keratan 10-mer in MD simulation, but boat/skew-boat conformations of galactose were decreasingly common in central rings of polymers of increasing length. This is in line with our observation that only 4C_1 conformations of galactose were sampled in nonsulfated keratan 20-mer MD. Additionally, some nonhelical glycosidic linkage conformations that caused a kink in the polymer chain, and thus contributed to compact GAG backbone conformations, were found more in GAG 20-mers than in short GAG oligosaccharides. Therefore, conformational landscapes from GAG 20-mer MD simulations likely provide a better representation of conformations of long GAG polymers than existing conformational landscapes of monosaccharides and GAG oligosaccharides.

A comparison of backbone conformational analyses of 200-mers of different GAG types provided insights into structural features and conformational behaviors contributing to GAG polymer backbone flexibility. For example, 200-mer end-to-end distance distribution curves were qualitatively similar in nonsulfated dermatan (Figure 13), keratan (Figure 18), and chondroitin [33], whereas hyaluronan and nonsulfated heparan polymer end-to-end distance distributions (Figures 7 and 23, respectively) showed an even higher tendency toward compact conformations with increasing polymer length than other nonsulfated GAGs. To find possible explanations for this, we compared observations from MD analyses of all GAG types. In hyaluronan 10- and 20-mer MD simulations, (1) GlcNAc conformations were similar to those in nonsulfated keratan and heparan MD, and (2) GlcA conformations were similar to those in nonsulfated chondroitin MD [33]. Furthermore, IdoA rings showed more flexibility in MD simulation of nonsulfated dermatan and heparan than GlcA rings in hyaluronan and chondroitin MD [33]. However, boat/skew-boat conformations (except for 2S_0) in IdoA did not appear to cause more compact polymer backbone conformations than 1C_4 , 2S_0 , and 4C_1 IdoA conformations, while boat/skew-boat GlcA ring conformations were more highly associated with more compact polymer backbone conformations in hyaluronan and chondroitin [33]. Additionally, in hyaluronan 10- and 20-mer MD simulations, (1) GlcNAc β 1-4GlcA linkages took on the same conformations as Gal β 1-4GlcNAc linkages in nonsulfated keratan MD, IdoA α 1-4GlcNAc linkages in nonsulfated heparan MD, and GalNAc β 1-4GlcA in nonsulfated chondroitin MD [33] (i.e., primarily $-\phi$, $+\psi$ with a secondary basin at $-\phi$, $-\psi$) and (2) GlcA β 1-3GlcNAc linkages took on the same conformations as IdoA α 1-3GalNAc in nonsulfated dermatan MD, GlcNAc β 1-3Gal in nonsulfated keratan MD, and GlcA β 1-3GalNAc linkages in nonsulfated chondroitin MD [33] (i.e., primarily $-\phi$, $-\psi$), but with more conformations at $-\phi$, $+\psi$. All secondary and tertiary conformations (i.e., $-\phi$, $-\psi$ in 1–4 linkages and $-\phi$, $+\psi$ in 1–3 linkages) were nonhelical and caused a kink or slight bend in the polymer chain, and were, therefore, associated with more compact conformations. The differences in energetic stability of secondary and tertiary linkage conformations between different GAG types can be explained by the adjacent monosaccharide structure. The major observation that is unique to hyaluronan is that both glycosidic linkage types take on more of these nonhelical secondary conformations in 10- and 20-mer MD simulations than any of nonsulfated dermatan, keratan, or chondroitin [33] in 10- and 20-mer MD simulations. Heparan is unique, in that it has all 1–4 linkages, while all other GAGs have alternating 1–3 and 1–4 linkages, and it contains α -GlcNAc, which gives unique conformations in GlcNAc α 1-4IdoA linkages. These characteristics give nonsulfated heparan a tendency toward more coiled structures than other GAGs.

Based on these observations and the fact that nonsulfated dermatan, keratan, and chondroitin showed more extended polymer backbone conformations than hyaluronan and nonsulfated heparan in MD simulations, it is likely that: (1) monosaccharide structure and conformational flexibility determine adjacent monosaccharide conformational flexibility; (2) although there is a much higher degree of ring

flexibility in IdoA than in GlcA, the flexibility of GlcA rings is more highly associated with GAG polymer backbone flexibility than that of IdoA rings; and (3) nonhelical glycosidic linkage conformations, which cause kinks in GAG polymer chains, contribute to more compact GAG polymer backbone conformations and, thus, a higher degree of GAG polymer backbone flexibility. These are valuable insights that would be difficult to obtain from conformational analyses of only solid-state structures and MD-generated conformational ensembles of short GAG oligosaccharides. Other important insights that can be gained from 3-D atomic-resolution conformational ensembles of GAG polymers include potential binding properties, and thus, predictions of binding poses with other biomolecules of interest. Models of long-chain GAG polymers (Figure 24 and [33]) can help characterize complete PGs and GAG-mediated complexes between multiple biomolecules, and consequently, improve our understanding of the bioactivity and function of GAG biopolymers in animal tissue.

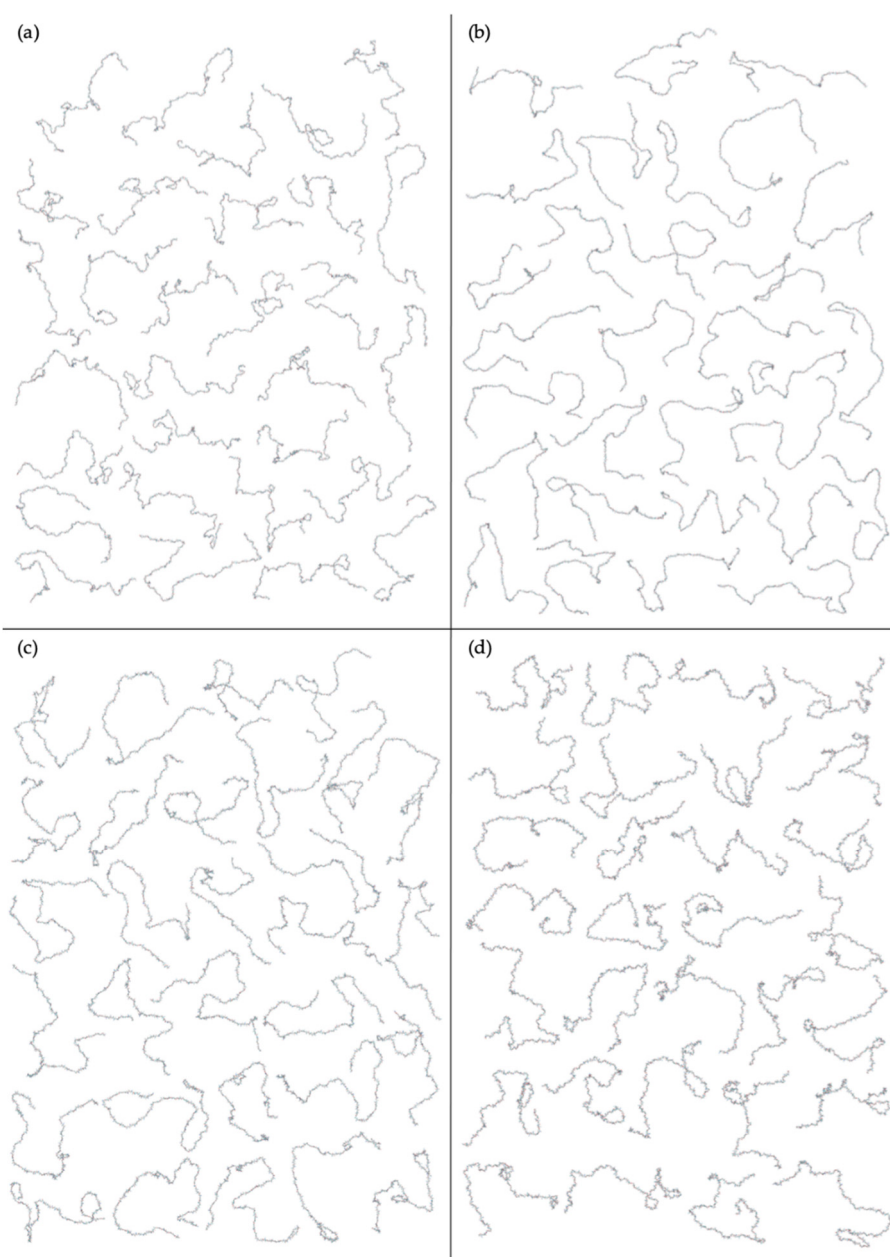


Figure 24. Constructed 200-mer conformations with probable end-to-end distances for each GAG type: (a) hyaluronan, (b) nonsulfated dermatan, (c) nonsulfated keratan, (d) nonsulfated heparan; n.b. depth is not shown, so some conformations appear to have overlap.

Supplementary Materials: The following are available online <http://www.mdpi.com/1422-0067/21/20/7699/s1>, Table S1: Most Probable End-to-End Distances (d) in MD-Generated 20-mer Conformations with Glycosidic Linkage Conformations in Secondary Basins, Table S2: Percent of Occurrences of Different IdoA Ring Puckers in Nonsulfated Dermatan MD Simulations, Table S3: Percent of Occurrences of Different IdoA Ring Puckers in Nonsulfated Heparan MD Simulations, Figure S1: Constructed hyaluronan 20-mer conformation with a pierced ring, Figure S2: Scatterplots of radius of gyration as a function of end-to-end distance in MD-generated and constructed hyaluronan ensembles, Figure S3: System potential energy probability distribution of the MD-generated hyaluronan 20-mer ensemble, Figure S4: C-P plots and C-P parameter θ timeseries for each GlcNAc monosaccharide ring in the MD-generated hyaluronan 20-mer ensemble, Figure S5: C-P plots and C-P parameter θ timeseries for each GlcA monosaccharide ring in the MD-generated hyaluronan 20-mer ensemble, Figure S6: Scatterplot of radius of gyration as a function of end-to-end distance in MD-generated hyaluronan 20-mer conformations with non- 4C_1 ring puckers, Figure S7: $\Delta G(\phi, \psi)$ plots for each glycosidic linkage in the MD-generated hyaluronan 20-mer ensemble, Figure S8: MD snapshots of hyaluronan 20-mer GlcA β 1-3GlcNAc glycosidic linkages with dihedrals in different basins and end-to-end distance distributions of 20-mer conformations with linkages in secondary basin, Figure S9: $\Delta G(\phi, \psi)$ plots for glycosidic linkages flanking non- 4C_1 ring puckers in the MD-generated hyaluronan 20-mer ensemble, Figure S10: MD snapshots of hyaluronan 20-mer GlcNAc β 1-4GlcA glycosidic linkages with dihedrals in different basins and end-to-end distance distributions of 20-mer conformations with linkages in secondary basin, Figure S11: C-P plots for GlcNAc and GlcA in the constructed hyaluronan 20-mer ensemble, Figure S12: $\Delta G(\phi, \psi)$ for aggregated glycosidic linkage data from the constructed hyaluronan 20-mer ensemble, Figure S13: Scatterplots of radius of gyration as a function of end-to-end distance in MD-generated and constructed nonsulfated dermatan ensembles, Figure S14: C-P parameter θ timeseries for each GalNAc monosaccharide ring in the MD-generated nonsulfated dermatan 20-mer ensemble, Figure S15: End-to-end distance distributions of MD-generated nonsulfated dermatan 20-mer conformations with boat/skew-boat ring puckers that cause a kink in the polymer chain and 2S_0 conformations, Figure S16: C-P plots and C-P parameter θ timeseries for each IdoA monosaccharide ring in the MD-generated nonsulfated dermatan 20-mer ensemble, Figure S17: C-P plots and C-P parameter θ timeseries for each IdoA monosaccharide ring in the MD-generated nonsulfated dermatan 10-mer ensemble, Figure S18: C-P plots for GalNAc and IdoA in the constructed nonsulfated dermatan 20-mer ensemble, Figure S19: $\Delta G(\phi, \psi)$ for aggregated glycosidic linkage data from the constructed nonsulfated dermatan 20-mer ensemble, Figure S20: Scatterplots of radius of gyration as a function of end-to-end distance in MD-generated and constructed nonsulfated keratan ensembles, Figure S21: C-P plots and C-P parameter θ timeseries for each GlcNAc monosaccharide ring in the MD-generated nonsulfated keratan 20-mer ensemble, Figure S22: Scatterplot of radius of gyration as a function of end-to-end distance in MD-generated nonsulfated keratan 20-mer conformations with non- 4C_1 ring puckers, Figure S23: MD snapshots of nonsulfated keratan 20-mer Gal β 1-4GlcNAc glycosidic linkages with dihedrals in different basins and end-to-end distance distributions of 20-mer conformations with linkages in secondary and tertiary basins, Figure S24: MD snapshots of nonsulfated keratan 20-mer GlcNAc β 1-3Gal glycosidic linkages with dihedrals in different basins and end-to-end distance distributions of 20-mer conformations with linkages in secondary and tertiary basins, Figure S25: C-P plots and C-P parameter θ timeseries for each Gal monosaccharide ring in the MD-generated nonsulfated keratan 10-mer ensemble, Figure S26: C-P plots for GlcNAc and Gal in the constructed nonsulfated keratan 20-mer ensemble, Figure S27: $\Delta G(\phi, \psi)$ for aggregated glycosidic linkage data from the constructed nonsulfated keratan 20-mer ensemble, Figure S28: Scatterplots of radius of gyration as a function of end-to-end distance in MD-generated and constructed nonsulfated heparan ensembles, Figure S29: C-P parameter θ timeseries for each IdoA monosaccharide ring in the MD-generated nonsulfated heparan 20-mer ensemble, Figure S30: End-to-end distance distributions of MD-generated nonsulfated heparan 20-mer conformations with boat/skew-boat ring puckers that cause a kink in the polymer chain and 2S_0 conformations, Figure S31: Snapshot of nonsulfated heparan 20-mer MD-generated ensemble highlighting GlcNAc α 1-4IdoA linkages, Figure S32: End-to-end distance distributions of MD-generated nonsulfated heparan 20-mer conformations with linkage dihedrals in secondary basins, Figure S33: $\Delta G(\phi, \psi)$ plots for IdoA α 1-4GlcNAc linkages flanking different IdoA conformations in the MD-generated nonsulfated heparan 20-mer ensemble, Figure S34: C-P plots for GlcNAc and IdoA in the constructed nonsulfated heparan 20-mer ensemble, Figure S35: $\Delta G(\phi, \psi)$ for aggregated glycosidic linkage data from the constructed nonsulfated heparan 20-mer ensemble, Figure S36: Bond potential energy probability distributions from constructed hyaluronan ensembles, Figure S37: Bond potential energy probability distributions from constructed nonsulfated dermatan ensembles, Figure S38: Bond potential energy probability distributions from constructed nonsulfated keratan ensembles, Figure S39: Bond potential energy probability distributions from constructed nonsulfated heparan ensembles.

Author Contributions: Conceptualization, E.K.W. and O.G.; methodology, E.K.W. and O.G.; software, E.K.W. and O.G.; validation E.K.W. and O.G.; formal analysis, E.K.W.; investigation E.K.W. and O.G.; resources, O.G.; data curation, E.K.W., D.M., and O.G.; writing—original draft preparation, E.K.W.; writing—review and editing, E.K.W., D.M., and O.G.; visualization, E.K.W.; supervision, O.G.; project administration, O.G.; funding acquisition, O.G. All authors have read and agree to the published version of the manuscript.

Funding: This research and the APC were funded by the National Science Foundation, grant number MCB-1453529 to O.G.

Acknowledgments: We would like to thank Richard Venable of the National Institutes of Health for sharing the Cremer-Pople parameter calculation code, modified from the original version by Tell Tuttle of the University of Strathclyde.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

PG	Proteoglycan
GAG	Glycosaminoglycan
MD	Molecular dynamics
GlcA	Glucuronate
IdoA	Iduronate
Gal	Galactose
GlcNAc	N-acetylglucosamine
GalNAc	N-acetylgalactosamine

References

1. Yoneda, A.; Couchman, J.R. Regulation of cytoskeletal organization by syndecan transmembrane proteoglycans. *Matrix Biol.* **2003**, *22*, 25–33. [[CrossRef](#)]
2. Kolset, S.O.; Prydz, K.; Pejler, G. Intracellular proteoglycans. *Biochem. J.* **2004**, *379*, 217–227. [[CrossRef](#)] [[PubMed](#)]
3. Mizumoto, S.; Fongmoon, D.; Sugahara, K. Interaction of chondroitin sulfate and dermatan sulfate from various biological sources with heparin-binding growth factors and cytokines. *Glycoconj. J.* **2012**, *30*, 619–632. [[CrossRef](#)] [[PubMed](#)]
4. Sugahara, K.; Mikami, T.; Uyama, T.; Mizuguchi, S.; Nomura, K.; Kitagawa, H. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* **2003**, *13*, 612–620. [[CrossRef](#)] [[PubMed](#)]
5. Sugahara, K.; Mikami, T. Chondroitin/dermatan sulfate in the central nervous system. *Curr. Opin. Struct. Biol.* **2007**, *17*, 536–545. [[CrossRef](#)]
6. Kawashima, H.; Atarashi, K.; Hirose, M.; Hirose, J.; Yamada, S.; Sugahara, K.; Miyasaka, M. Oversulfated chondroitin/dermatan sulfates containing GlcA β 1/IdoA α 1-3GalNAc(4,6-O-disulfate) interact with L- and P-selectin and chemokines. *J. Biol. Chem.* **2002**, *277*, 12921–12930. [[CrossRef](#)]
7. Lander, A.D.; Selleck, S.B. The elusive functions of proteoglycans: In vivo veritas. *J. Cell Biol.* **2000**, *148*, 227–232. [[CrossRef](#)]
8. Stallcup, W.B.; Dahlin, K.; Healy, P. Interaction of the NG2 chondroitin sulfate proteoglycan with type VI collagen. *J. Cell Biol.* **1990**, *111*, 3177–3188. [[CrossRef](#)]
9. Van Susante, J.L.C.; Pieper, J.; Buma, P.; van Kuppevelt, T.H.; van Beuningen, H.; van Der Kraan, P.M.; Veerkamp, J.H.; van den Berg, W.B.; Veth, R.P.H. Linkage of chondroitin-sulfate to type I collagen scaffolds stimulates the bioactivity of seeded chondrocytes in vitro. *Biomaterials* **2001**, *22*, 2359–2369. [[CrossRef](#)]
10. Streit, A.; Nolte, C.; Rásony, T.; Schachner, M. Interaction of astrochondrin with extracellular matrix components and its involvement in astrocyte process formation and cerebellar granule cell migration. *J. Cell Biol.* **1993**, *120*, 799–814. [[CrossRef](#)]
11. Djerbal, L.; Lortat-Jacob, H.; Kwok, J. Chondroitin sulfates and their binding molecules in the central nervous system. *Glycoconj. J.* **2017**, *34*, 363–376. [[CrossRef](#)] [[PubMed](#)]
12. Forsten-Williams, K.; Chu, C.L.; Fannon, M.; Buczek-Thomas, J.A.; Nugent, M.A. Control of growth factor networks by heparan sulfate proteoglycans. *Ann. Biomed. Eng.* **2008**, *36*, 2134–2148. [[CrossRef](#)] [[PubMed](#)]
13. Bhavanandan, V.P.; Davidson, E.A. Mucopolysaccharides associated with nuclei of cultured mammalian cells. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 2032–2036. [[CrossRef](#)]
14. Fedarko, N.S.; Conrad, H.E. A unique heparan sulfate in the nuclei of hepatocytes: Structural changes with the growth state of the cells. *J. Cell Biol.* **1986**, *102*, 587–599. [[CrossRef](#)] [[PubMed](#)]
15. Ishihara, M.; Fedarko, N.S.; Conrad, H.E. Transport of heparan sulfate into the nuclei of hepatocytes. *J. Biol. Chem.* **1986**, *261*, 13575–13580.

16. Tumova, S.; Hatch, B.A.; Law, D.J.; Bame, K.J. Basic fibroblast growth factor does not prevent heparan sulphate proteoglycan catabolism in intact cells, but it alters the distribution of the glycosaminoglycan degradation products. *Biochem. J.* **1999**, *337*, 471–481. [[CrossRef](#)] [[PubMed](#)]
17. Kjellén, L.; Lindahl, U. Specificity of glycosaminoglycan-protein interactions. *Curr. Opin. Struct. Biol.* **2018**, *50*, 101–108. [[CrossRef](#)] [[PubMed](#)]
18. Sharif, M.; George, E.; Shepstone, L.; Knudson, W.; Thonar, E.J.M.; Cushnaghan, J.; Dieppe, P. Serum hyaluronic acid level as a predictor of disease progression in osteoarthritis of the knee. *Arthritis Rheum.* **1995**, *38*, 760–767. [[CrossRef](#)]
19. Ghosh, P. The role of hyaluronic acid (hyaluronan) in health and disease: Interactions with cells, cartilage and components of synovial fluid. *Clin. Exp. Rheum.* **1994**, *12*, 75–82.
20. Mizumoto, S.; Kosho, T.; Hatamochi, A.; Honda, T.; Yamaguchi, T.; Okamoto, N.; Miyake, N.; Yamada, S.; Sugahara, K. Defect in dermatan sulfate in urine of patients with Ehlers-Danlos syndrome caused by a CHST14/D4ST1 deficiency. *Clin. Biochem.* **2017**, *50*, 670–677. [[CrossRef](#)]
21. Kosho, T. CHST14/D4ST1 deficiency: New form of Ehlers–Danlos syndrome. *Pediatr. Int.* **2016**, *58*, 88–99. [[CrossRef](#)]
22. Lindahl, B.; Eriksson, L.; Spillmann, D.; Caterson, B.; Lindahl, U. Selective loss of cerebral keratan sulfate in Alzheimer’s disease. *J. Biol. Chem.* **1996**, *271*, 16991–16994. [[CrossRef](#)] [[PubMed](#)]
23. Martin, L.; Schmitz, S.; De Santis, R.; Doemming, S.; Haase, H.; Hoeger, J.; Heinbockel, L.; Brandenburg, K.; Marx, G.; Schuerholz, T. Peptide 19-2.5 inhibits heparan sulfate-triggered inflammation in murine cardiomyocytes stimulated with human sepsis serum. *PLoS ONE* **2015**, *10*, e0127584. [[CrossRef](#)] [[PubMed](#)]
24. Schmidt, E.P.; Overdier, K.H.; Sun, X.; Lin, L.; Liu, X.; Yang, Y.; Ammons, L.A.; Hiller, T.D.; Suflita, M.A.; Yu, Y. Urinary glycosaminoglycans predict outcomes in septic shock and acute respiratory distress syndrome. *Am. J. Resp. Crit. Care* **2016**, *194*, 439–449. [[CrossRef](#)] [[PubMed](#)]
25. Gallagher, J. Fell-Muir Lecture: Heparan sulphate and the art of cell regulation: A polymer chain conducts the protein orchestra. *Int. J. Exp. Pathol.* **2015**, *96*, 203–231. [[CrossRef](#)] [[PubMed](#)]
26. Shriver, Z.; Raguram, S.; Sasisekharan, R. Glycomics: A pathway to a class of new and improved therapeutics. *Nat. Rev. Drug Discov.* **2004**, *3*, 863–873. [[CrossRef](#)] [[PubMed](#)]
27. Trottein, F.; Schaffer, L.; Ivanov, S.; Paget, C.; Vendeville, C.; Cazet, A.; Groux-Degroote, S.; Lee, S.; Krzewinski-Recchi, M.-A.; Faveeuw, C.; et al. Glycosyltransferase and sulfotransferase gene expression profiles in human monocytes, dendritic cells and macrophages. *Glycoconj. J.* **2009**, *26*, 1259–1274. [[CrossRef](#)]
28. Gandhi, N.S.; Mancera, R.L. The Structure of glycosaminoglycans and their interactions with proteins. *Chem. Biol. Drug Des.* **2008**, *72*, 455–482. [[CrossRef](#)]
29. Casu, B.; Petitou, M.; Provasoli, M.; Sinay, P. Conformational flexibility: A new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans. *Trends Biochem. Sci.* **1988**, *13*, 221–225. [[CrossRef](#)]
30. Mulloy, B.; Forster, M.J.; Jones, C.; Drake, A.F.; Johnson, E.A.; Davies, D.B. The effect of variation of substitution on the solution conformation of heparin: A spectroscopic and molecular modelling study. *Carbohydr. Res.* **1994**, *255*, 1–26. [[CrossRef](#)]
31. Zamparo, O.; Comper, W.D. The hydrodynamic frictional coefficient of polysaccharides: The role of the glycosidic linkage. *Carbohydr. Res.* **1991**, *212*, 193–200. [[CrossRef](#)]
32. Samsonov, S.A.; Gehrcke, J.-P.; Pisabarro, M.T. Flexibility and explicit solvent in molecular-dynamics-based docking of protein–glycosaminoglycan systems. *J. Chem. Inf. Model.* **2014**, *54*, 582–592. [[CrossRef](#)] [[PubMed](#)]
33. Whitmore, E.K.; Vesenka, G.; Sihler, H.; Guvench, O. Efficient Construction of Atomic-Resolution Models of Non-Sulfated Chondroitin Glycosaminoglycan Using Molecular Dynamics Data. *Biomolecules* **2020**, *10*, 537. [[CrossRef](#)] [[PubMed](#)]
34. Toyoda, H.; Kinoshita-Toyoda, A.; Selleck, S.B. Structural analysis of glycosaminoglycans in *Drosophila* and *Caenorhabditis elegans* and demonstration that tout-velu, a drosophila gene related to EXT tumor suppressors, affects heparan sulfate in vivo. *J. Biol. Chem.* **2000**, *275*, 2269–2275. [[CrossRef](#)] [[PubMed](#)]
35. Yang, H.O.; Gunay, N.S.; Toida, T.; Kuberan, B.; Yu, G.; Kim, Y.S.; Linhardt, R.J. Preparation and structural determination of dermatan sulfate-derived oligosaccharides. *Glycobiology* **2000**, *10*, 1033–1039. [[CrossRef](#)] [[PubMed](#)]
36. Zaia, J.; Costello, C.E. Compositional analysis of glycosaminoglycans by electrospray mass spectrometry. *Anal. Chem.* **2001**, *73*, 233–239. [[CrossRef](#)] [[PubMed](#)]

37. Shao, C.; Zhang, F.; Kemp, M.M.; Linhardt, R.J.; Waisman, D.M.; Head, J.F.; Seaton, B.A. Crystallographic analysis of calcium-dependent heparin binding to annexin A2. *J. Biol. Chem.* **2006**, *281*, 31689–31695. [[CrossRef](#)]
38. Capila, I.; Hernáiz, M.a.J.; Mo, Y.D.; Mealy, T.R.; Campos, B.; Dedman, J.R.; Linhardt, R.J.; Seaton, B.A. Annexin V–heparin oligosaccharide complex suggests heparan sulfate–mediated assembly on cell surfaces. *Structure* **2001**, *9*, 57–64. [[CrossRef](#)]
39. Dementiev, A.; Petitou, M.; Herbert, J.-M.; Gettins, P.G.W. The ternary complex of antithrombin–anhydrothrombin–heparin reveals the basis of inhibitor specificity. *Nat. Struct. Mol. Biol.* **2004**, *11*, 863–867. [[CrossRef](#)] [[PubMed](#)]
40. Li, W.; Johnson, D.J.D.; Esmon, C.T.; Huntington, J.A. Structure of the antithrombin–thrombin–heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat. Struct. Mol. Biol.* **2004**, *11*, 857–862. [[CrossRef](#)] [[PubMed](#)]
41. Li, Z.; Kienetz, M.; Cherney, M.M.; James, M.N.G.; Brömme, D. The crystal and molecular structures of a cathepsin K:chondroitin sulfate complex. *J. Mol. Biol.* **2008**, *383*, 78–91. [[CrossRef](#)]
42. Imberty, A.; Lortat-Jacob, H.; Pérez, S. Structural view of glycosaminoglycan–protein interactions. *Carbohydr. Res.* **2007**, *342*, 430–439. [[CrossRef](#)] [[PubMed](#)]
43. Ferro, D.R.; Provasoli, A.; Ragazzi, M.; Torri, G.; Casu, B.; Gatti, G.; Jacquinet, J.C.; Sinay, P.; Petitou, M.; Choay, J. Evidence for conformational equilibrium of the sulfated L-iduronate residue in heparin and in synthetic heparin mono- and oligo-saccharides: NMR and force-field studies. *J. Am. Chem. Soc.* **1986**, *108*, 6773–6778. [[CrossRef](#)]
44. Bossennec, V.; Petitou, M.; Perly, B. ¹H-n.m.r. investigation of naturally occurring and chemically oversulphated dermatan sulphates. Identification of minor monosaccharide residues. *Biochem. J.* **1990**, *267*, 625–630. [[CrossRef](#)] [[PubMed](#)]
45. Ferro, D.R.; Provasoli, A.; Ragazzi, M.; Casu, B.; Torri, G.; Bossennec, V.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. Conformer populations of l-iduronic acid residues in glycosaminoglycan sequences. *Carbohydr. Res.* **1990**, *195*, 157–167. [[CrossRef](#)]
46. Yamada, S.; Yamane, Y.; Sakamoto, K.; Tsuda, H.; Sugahara, K. Structural determination of sulfated tetrasaccharides and hexasaccharides containing a rare disaccharide sequence, -3GalNAc(4,6-disulfate)beta1-4IdoAalpha1-, isolated from porcine intestinal dermatan sulfate. *Eur. J. Biochem.* **1998**, *258*, 775–783. [[CrossRef](#)]
47. Silipo, A.; Zhang, Z.; Cañada, F.J.; Molinaro, A.; Linhardt, R.J.; Jiménez-Barbero, J. Conformational analysis of a dermatan sulfate-derived tetrasaccharide by NMR, molecular modeling, and residual dipolar couplings. *Chembiochem* **2008**, *9*, 240–252. [[CrossRef](#)]
48. Sattelle, B.M.; Shakeri, J.; Cliff, M.J.; Almond, A. Proteoglycans and their heterogeneous glycosaminoglycans at the atomic scale. *Biomacromolecules* **2015**, *16*, 951–961. [[CrossRef](#)]
49. Almond, A.; Sheehan, J.K. Glycosaminoglycan conformation: Do aqueous molecular dynamics simulations agree with x-ray fiber diffraction? *Glycobiology* **2000**, *10*, 329–338. [[CrossRef](#)]
50. Gandhi, N.S.; Mancera, R.L. Can current force fields reproduce ring puckering in 2-O-sulfo- α -l-iduronic acid? A molecular dynamics simulation study. *Carbohydr. Res.* **2010**, *345*, 689–695. [[CrossRef](#)]
51. Sattelle, B.M.; Hansen, S.U.; Gardiner, J.; Almond, A. Free energy landscapes of iduronic acid and related monosaccharides. *J. Am. Chem. Soc.* **2010**, *132*, 13132–13134. [[CrossRef](#)]
52. Balogh, G.; Gyöngyösi, T.S.; Timári, I.; Herczeg, M.; Borbás, A.; Fehér, K.; Kövér, K.E. Comparison of carbohydrate force fields using Gaussian Accelerated Molecular Dynamics simulations and development of force field parameters for heparin-analogue pentasaccharides. *J. Chem. Inf. Model.* **2019**, *59*, 4855–4867. [[CrossRef](#)] [[PubMed](#)]
53. Ng, C.; Nandha Premnath, P.; Guvench, O. Rigidity and flexibility in the tetrasaccharide linker of proteoglycans from atomic-resolution molecular simulation. *J. Comput. Chem.* **2017**, *38*, 1438–1446. [[CrossRef](#)] [[PubMed](#)]
54. Guvench, O. Revealing the mechanisms of protein disorder and N-glycosylation in CD44-hyaluronan binding using molecular simulation. *Front. Immunol.* **2015**, *6*, 305. [[CrossRef](#)] [[PubMed](#)]
55. Faller, C.E.; Guvench, O. Sulfation and cation effects on the conformational properties of the glycan backbone of chondroitin sulfate disaccharides. *J. Phys. Chem. B* **2015**, *119*, 6063–6073. [[CrossRef](#)] [[PubMed](#)]
56. Favreau, A.J.; Faller, C.E.; Guvench, O. CD44 receptor unfolding enhances binding by freeing basic amino acids to contact carbohydrate ligand. *Biophys. J.* **2013**, *105*, 1217–1226. [[CrossRef](#)] [[PubMed](#)]

57. Guvench, O.; Mallajosyula, S.S.; Raman, E.P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T.J.; Jamison, F.W., II; Mackerell, A.D., 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*, 3162–3180. [[CrossRef](#)]
58. Brannigan, G.; Lin, L.C.; Brown, F.L. Implicit solvent simulation models for biomembranes. *Eur. Biophys. J.* **2006**, *35*, 104–124. [[CrossRef](#)]
59. Srinivas, G.; Cheng, X.; Smith, J.C. A solvent-free coarse grain model for crystalline and amorphous cellulose fibrils. *J. Chem. Theory Comput.* **2011**, *7*, 2539–2548. [[CrossRef](#)]
60. Ingólfsson, H.I.; Lopez, C.A.; Uusitalo, J.J.; de Jong, D.H.; Gopal, S.M.; Periole, X.; Marrink, S.J. The power of coarse graining in biomolecular simulations. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 225–248. [[CrossRef](#)]
61. Guvench, O.; Greene, S.N.; Kamath, G.; Brady, J.W.; Venable, R.M.; Pastor, R.W.; Mackerell, A.D., Jr. Additive empirical force field for hexopyranose monosaccharides. *J. Comput. Chem.* **2008**, *29*, 2543–2564. [[CrossRef](#)] [[PubMed](#)]
62. Guvench, O.; Hatcher, E.R.; Venable, R.M.; Pastor, R.W.; Mackerell, A.D. CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J. Chem. Theory Comput.* **2009**, *5*, 2353–2370. [[CrossRef](#)] [[PubMed](#)]
63. Mallajosyula, S.S.; Guvench, O.; Hatcher, E.; Mackerell, A.D., Jr. CHARMM additive all-atom force field for phosphate and sulfate linked to carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 759–776. [[CrossRef](#)] [[PubMed](#)]
64. Hascall, V.C.; Heinegård, D. Aggregation of cartilage proteoglycans II. Oligosaccharide competitors of the proteoglycan-hyaluronic acid interaction. *J. Biol. Chem.* **1974**, *249*, 4242–4249. [[PubMed](#)]
65. Winter, W.; Smith, P.; Arnott, S. Hyaluronic acid: Structure of a fully extended 3-fold helical sodium salt and comparison with the less extended 4-fold helical forms. *J. Mol. Biol.* **1975**, *99*, 219–235. [[CrossRef](#)]
66. Sheehan, J.; Gardner, K.; Atkins, E. Hyaluronic acid: A double-helical structure in the presence of potassium at low pH and found also with the cations ammonium, rubidium and caesium. *J. Mol. Biol.* **1977**, *117*, 113–135. [[CrossRef](#)]
67. Almond, A.; Brass, A.; Sheehan, J.K. Dynamic exchange between stabilized conformations predicted for hyaluronan tetrasaccharides: Comparison of molecular dynamics simulations with available NMR data. *Glycobiology* **1998**, *8*, 973–980. [[CrossRef](#)]
68. Almond, A.; Sheehan, J.K.; Brass, A. Molecular dynamics simulations of the two disaccharides of hyaluronan in aqueous solution. *Glycobiology* **1997**, *7*, 597–604. [[CrossRef](#)]
69. Gribbon, P.; Heng, B.C.; Hardingham, T.E. The molecular basis of the solution properties of hyaluronan investigated by confocal fluorescence recovery after photobleaching. *Biophys. J.* **1999**, *77*, 2210–2216. [[CrossRef](#)]
70. Atkins, E.; Meader, D.; Scott, J. Model for hyaluronic acid incorporating four intramolecular hydrogen bonds. *Int. J. Biol. Macromol.* **1980**, *2*, 318–319. [[CrossRef](#)]
71. Scott, J.E. Supramolecular organization of extracellular matrix glycosaminoglycans, in vitro and in the tissues. *FASEB J.* **1992**, *6*, 2639–2645. [[CrossRef](#)]
72. Scott, J.E.; Cummings, C.; Brass, A.; Chen, Y. Secondary and tertiary structures of hyaluronan in aqueous solution, investigated by rotary shadowing-electron microscopy and computer simulation. *Biochem. J.* **1991**, *274*, 699–705. [[CrossRef](#)] [[PubMed](#)]
73. Scott, J.; Heatley, F.; Moorcroft, D.; Olavesen, A. Secondary structures of hyaluronate and chondroitin sulphates. A 1H nmr study of NH signals in dimethyl sulphoxide solution. *Biochem. J.* **1981**, *199*, 829. [[CrossRef](#)] [[PubMed](#)]
74. Singh, A.; Montgomery, D.; Xue, X.; Foley, B.L.; Woods, R.J. GAG Builder: A web-tool for modeling 3D structures of glycosaminoglycans. *Glycobiology* **2019**, *29*, 515–518. [[CrossRef](#)] [[PubMed](#)]
75. Engelsen, S.B.; Hansen, P.I.; Perez, S. POLYS 2.0: An open source software package for building three-dimensional structures of polysaccharides. *Biopolymers* **2014**, *101*, 733–743. [[CrossRef](#)]
76. Kuttel, M.M.; Stähle, J.; Widmalm, G. CarbBuilder: Software for building molecular models of complex oligo- and polysaccharide structures. *J. Comput. Chem.* **2016**, *37*, 2098–2105. [[CrossRef](#)]
77. Clerc, O.; Deniaud, M.; Vallet, S.D.; Naba, A.; Rivet, A.; Perez, S.; Thierry-Mieg, N.; Ricard-Blum, S. MatrixDB: Integration of new data with a focus on glycosaminoglycan interactions. *Nucleic Acids Res.* **2018**, *47*, D376–D381. [[CrossRef](#)] [[PubMed](#)]

78. Clerc, O.; Mariethoz, J.; Rivet, A.; Lisacek, F.; Pérez, S.; Ricard-Blum, S. A pipeline to translate glycosaminoglycan sequences into 3D models. Application to the exploration of glycosaminoglycan conformational space. *Glycobiology* **2018**, *29*, 36–44. [[CrossRef](#)]
79. Jo, S.; Kim, T.; Iyer, V.G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **2008**, *29*, 1859–1865. [[CrossRef](#)] [[PubMed](#)]
80. Jo, S.; Song, K.C.; Desaire, H.; MacKerell, A.D., Jr.; Im, W. Glycan Reader: Automated sugar identification and simulation preparation for carbohydrates and glycoproteins. *J. Comput. Chem.* **2011**, *32*, 3135–3141. [[CrossRef](#)]
81. Park, S.-J.; Lee, J.; Patel, D.S.; Ma, H.; Lee, H.S.; Jo, S.; Im, W. Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank. *Bioinformatics* **2017**, *33*, 3051–3057. [[CrossRef](#)] [[PubMed](#)]
82. Park, S.-J.; Lee, J.; Qi, Y.; Kern, N.R.; Lee, H.S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im, W. CHARMM-GUI glycan modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **2019**, *29*, 320–331. [[CrossRef](#)] [[PubMed](#)]
83. Brooks, B.R.; Bruccoleri, R.E.; Olafson, B.D.; States, D.J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187–217. [[CrossRef](#)]
84. Brooks, B.R.; Brooks, C.L., III; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [[CrossRef](#)]
85. MacKerell, A.D., Jr.; Brooks, B.; Brooks, C.L., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The energy function and its parameterization with an overview of the program. In *Encyclopedia of Computational Chemistry*; John Wiley and Sons, Ltd.: New York, NY, USA, 1998; Volume 1, pp. 271–277.
86. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [[CrossRef](#)]
87. Durell, S.R.; Brooks, B.R.; Ben-Naim, A. Solvent-induced forces between two hydrophilic groups. *J. Phys. Chem.* **1994**, *98*, 2198–2202. [[CrossRef](#)]
88. Beglov, D.; Roux, B. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. *J. Chem. Phys.* **1994**, *100*, 9050–9063. [[CrossRef](#)]
89. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781–1802. [[CrossRef](#)] [[PubMed](#)]
90. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38. [[CrossRef](#)]
91. Cremer, D.; Pople, J. General definition of ring puckering coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358. [[CrossRef](#)]
92. Banerji, S.; Wright, A.J.; Noble, M.; Mahoney, D.J.; Campbell, I.D.; Day, A.J.; Jackson, D.G. Structures of the CD44–hyaluronan complex provide insight into a fundamental carbohydrate-protein interaction. *Nat. Struct. Mol. Biol.* **2007**, *14*, 234. [[CrossRef](#)]
93. Jedrzejewski, M.J.; Mello, L.V.; de Groot, B.L.; Li, S. Mechanism of hyaluronan degradation by *Streptococcus pneumoniae* Hyaluronate Lyase. *J. Biol. Chem.* **2002**, *277*, 28287–28297. [[CrossRef](#)] [[PubMed](#)]
94. Mello, L.V.; de Groot, B.L.; Li, S.; Jedrzejewski, M.J. Structure and flexibility of *Streptococcus agalactiae* hyaluronate lyase complex with its substrate. *J. Biol. Chem.* **2002**, *277*, 36678–36688. [[CrossRef](#)] [[PubMed](#)]
95. Marković-Housley, Z.; Miglierini, G.; Soldatova, L.; Rizkallah, P.J.; Müller, U.; Schirmer, T. Crystal structure of hyaluronidase, a major allergen of bee venom. *Structure* **2000**, *8*, 1025–1035. [[CrossRef](#)]
96. Tashima, T.; Nagatoishi, S.; Caaveiro, J.M.; Nakakido, M.; Sagara, H.; Kusano-Arai, O.; Iwanari, H.; Mimuro, H.; Hamakubo, T.; Ohnuma, S.-i. Molecular basis for governing the morphology of type-I collagen fibrils by Osteomodulin. *Commun. Biol.* **2018**, *1*, 1–10. [[CrossRef](#)]
97. Paracuellos, P.; Kalamajski, S.; Bonna, A.; Bihan, D.; Farndale, R.W.; Hohenester, E. Structural and functional analysis of two small leucine-rich repeat proteoglycans, fibromodulin and chondroadherin. *Matrix Biol.* **2017**, *63*, 106–116. [[CrossRef](#)] [[PubMed](#)]
98. Ohto, U.; Usui, K.; Ochi, T.; Yuki, K.; Satow, Y.; Shimizu, T. Crystal structure of human β -galactosidase structural basis of gm1 gangliosidosis and morquio b diseases. *J. Biol. Chem.* **2012**, *287*, 1801–1812. [[CrossRef](#)]

99. Scally, S.W.; Petersen, J.; Law, S.C.; Dudek, N.L.; Nel, H.J.; Loh, K.L.; Wijeyewickrema, L.C.; Eckle, S.B.; van Heemst, J.; Pike, R.N. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *J. Exp. Med.* **2013**, *210*, 2569–2582. [[CrossRef](#)] [[PubMed](#)]
100. Sattelle, B.M.; Almond, A. Is N-acetyl-d-glucosamine a rigid 4C1 chair? *Glycobiology* **2011**, *21*, 1651–1662. [[CrossRef](#)]
101. Sicińska, W.; Adams, B.; Lerner, L. A detailed ¹H and ¹³C NMR study of a repeating disaccharide of hyaluronan: The effects of temperature and counterion type. *Carbohydr. Res.* **1993**, *242*, 29–51. [[CrossRef](#)]
102. Gargiulo, V.; Morando, M.A.; Silipo, A.; Nurisso, A.; Pérez, S.; Imberty, A.; Cañada, F.J.; Parrilli, M.; Jiménez-Barbero, J.; De Castro, C. Insights on the conformational properties of hyaluronic acid by using NMR residual dipolar couplings and MD simulations. *Glycobiology* **2010**, *20*, 1208–1216. [[CrossRef](#)] [[PubMed](#)]
103. Ikemizu, S.; Sparks, L.M.; van der Merwe, P.A.; Harlos, K.; Stuart, D.I.; Jones, E.Y.; Davis, S.J. Crystal structure of the CD2-binding domain of CD58 (lymphocyte function-associated antigen 3) at 1.8-Å resolution. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4289–4294. [[CrossRef](#)] [[PubMed](#)]
104. Chang, C.; Magracheva, E.; Kozlov, S.; Fong, S.; Tobin, G.; Kottenko, S.; Wlodawer, A.; Zdanov, A. Crystal structure of interleukin-19 defines a new subfamily of helical cytokines. *J. Biol. Chem.* **2003**, *278*, 3308–3313. [[CrossRef](#)] [[PubMed](#)]
105. Xu, R.; McBride, R.; Paulson, J.C.; Basler, C.F.; Wilson, I.A. Structure, receptor binding, and antigenicity of influenza virus hemagglutinins from the 1957 H2N2 pandemic. *J. Virol.* **2010**, *84*, 1715–1721. [[CrossRef](#)] [[PubMed](#)]
106. Maveyraud, L.; Niwa, H.; Guillet, V.; Svergun, D.I.; Konarev, P.V.; Palmer, R.A.; Peumans, W.J.; Rougé, P.; Van Damme, E.J.; Reynolds, C.D. Structural basis for sugar recognition, including the Tn carcinoma antigen, by the lectin SNA-II from *Sambucus nigra*. *Proteins Struct. Funct. Bioinf.* **2009**, *75*, 89–103. [[CrossRef](#)] [[PubMed](#)]
107. Alibay, I.; Bryce, R.A. Ring puckering landscapes of glycosaminoglycan-related monosaccharides from molecular dynamics simulations. *J. Chem. Inf. Model.* **2019**, *59*, 4729–4741. [[CrossRef](#)]
108. Alibay, I.; Burusco, K.K.; Bruce, N.J.; Bryce, R.A. Identification of rare Lewis oligosaccharide conformers in aqueous solution using enhanced sampling molecular dynamics. *J. Phys. Chem. B* **2018**, *122*, 2462–2474. [[CrossRef](#)]
109. Nyerges, B.; Kovacs, A. Density functional study of the conformational space of 4C1 D-glucuronic acid. *J. Phys. Chem. A* **2005**, *109*, 892–897. [[CrossRef](#)]
110. Blundell, C.D.; Roberts, I.S.; Sheehan, J.K.; Almond, A. Investigating the molecular basis for the virulence of *Escherichia coli* K5 by nuclear magnetic resonance analysis of the capsule polysaccharide. *J. Mol. Microbiol. Biotechnol.* **2009**, *17*, 71–82. [[CrossRef](#)]
111. Furlan, S.; La Penna, G.; Perico, A.; Cesàro, A. Hyaluronan chain conformation and dynamics. *Carbohydr. Res.* **2005**, *340*, 959–970. [[CrossRef](#)]
112. Almond, A.; DeAngelis, P.L.; Blundell, C.D. Hyaluronan: The local solution conformation determined by NMR and computer modeling is close to a contracted left-handed 4-fold helix. *J. Mol. Biol.* **2006**, *358*, 1256–1269. [[CrossRef](#)] [[PubMed](#)]
113. Holmbeck, S.M.; Petillo, P.A.; Lerner, L.E. The solution conformation of hyaluronan: A combined NMR and molecular dynamics study. *Biochemistry* **1994**, *33*, 14246–14255. [[CrossRef](#)] [[PubMed](#)]
114. Adrian-Scotto, M.; Abdallah, K.B.; Mallet, G.; Vasilescu, D. Quantum molecular modeling of free radical saccharides from hyaluronan. *J. Molec. Struct.* **2003**, *636*, 89–113. [[CrossRef](#)]
115. Moulabbi, M.; Broch, H.; Robert, L.; Vasilescu, D. Quantum molecular modeling of hyaluronan. *J. Mol. Struct.* **1997**, *395*, 477–508. [[CrossRef](#)]
116. Guss, J.; Hukins, D.; Smith, P.; Winter, W.; Arnott, S.; Moorhouse, R.; Rees, D. Hyaluronic acid: Molecular conformations and interactions in two sodium salts. *J. Mol. Biol.* **1975**, *95*, 359–384. [[CrossRef](#)]
117. Rees, D. Conformational analysis of polysaccharides. Part II. Alternating copolymers of the agar-carrageenan-chondroitin type by model building in the computer with calculation of helical parameters. *J. Chem. Soc. B* **1969**, 217–226. [[CrossRef](#)]
118. Mitra, A.; Arnott, S.; Atkins, E.; Isaac, D. Dermatan sulfate: Molecular conformations and interactions in the condensed state. *J. Mol. Biol.* **1983**, *169*, 873–901. [[CrossRef](#)]
119. Virudachalam, R.; Rao, V.S. Theoretical investigations on 2-acetamido-2-deoxy aldohexopyranoses: Conformation and the anomeric effect. *Carbohydr. Res.* **1976**, *51*, 135–139. [[CrossRef](#)]

120. Hsieh, P.-H.; Thieker, D.F.; Guerrini, M.; Woods, R.J.; Liu, J. Uncovering the relationship between sulphation patterns and conformation of iduronic acid in heparan sulphate. *Nature* **2016**, *6*, 29602. [[CrossRef](#)]
121. Gatti, G.; Casu, B.; Torri, G.; Vercellotti, J.R. Resolution-enhanced ¹H-nmr spectra of dermatan sulfate and chondroitin sulfates: Conformation of the uronic acid residues. *Carbohydr. Res.* **1979**, *68*, C3–C7. [[CrossRef](#)]
122. Ernst, S.; Venkataraman, G.; Sasisekharan, V.; Langer, R.; Cooney, C.L.; Sasisekharan, R. Pyranose ring flexibility. Mapping of physical data for iduronate in continuous conformational space. *J. Am. Chem. Soc.* **1998**, *120*, 2099–2107. [[CrossRef](#)]
123. Van Boeckel, C.; Van Aelst, S.; Wagenaars, G.; Mellema, J.R.; Paulsen, H.; Peters, T.; Pollex, A.; Sinnwell, V. Conformational analysis of synthetic heparin-like oligosaccharides containing α -L-idopyranosyluronic acid. *Recl. Trav. Chim. Pays Bas.* **1987**, *106*, 19–29. [[CrossRef](#)]
124. Inoue, Y.; Inouye, Y.; Nagasawa, K. Conformational equilibria of the L-iduronate residue in non-sulphated di-, tetra- and hexa-saccharides and their alditols derived from dermatan sulphate. *Biochem. J.* **1990**, *265*, 533–538. [[CrossRef](#)] [[PubMed](#)]
125. Ragazzi, M.; Provasoli, A.; Ferro, D. Molecular mechanics and the structure of iduronate-containing carbohydrates. In *Computer Modeling of Carbohydrate Molecules*; French, A., Ed.; ACS Publications: Washington, DC, USA, 1990; pp. 332–344.
126. Forster, M.J.; Mulloy, B. Molecular dynamics study of iduronate ring conformation. *Biopolymers* **1993**, *33*, 575–588. [[CrossRef](#)]
127. Venkataraman, G.; Sasisekharan, V.; Cooney, C.L.; Langer, R.; Sasisekharan, R. A stereochemical approach to pyranose ring flexibility: Its implications for the conformation of dermatan sulfate. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 6171–6175. [[CrossRef](#)]
128. Arnott, S.; Guss, J.; Hukins, D.; Dea, I.; Rees, D. Conformation of keratan sulphate. *J. Mol. Biol.* **1974**, *88*, 175–184. [[CrossRef](#)]
129. Rudrum, M.; Shaw, D. 10. The structure and conformation of some monosaccharides in solution. *J. Chem. Soc.* **1965**, 52–57. [[CrossRef](#)]
130. Angyal, S.J. Conformational analysis in carbohydrate chemistry: I. Conformational free energies. The conformations and α : β ratios of aldopyranoses in aqueous solution. *Aust. J. Chem.* **1968**, *21*, 2737–2746. [[CrossRef](#)]
131. Bush, C.A.; Feeney, R.E. Conformation of the glycotriptide repeating unit of antifreeze glycoprotein of polar fish as determined from the fully assigned proton nmr spectrum. *Int. J. Pept. Prot. Res.* **1986**, *28*, 386–397. [[CrossRef](#)] [[PubMed](#)]
132. Mulloy, B.; Forster, M.J. Conformation and dynamics of heparin and heparan sulfate. *Glycobiology* **2000**, *10*, 1147–1156. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).