HHS Public Access

Author manuscript

Can J Chem. Author manuscript; available in PMC 2017 June 08.

Published in final edited form as:

Can J Chem. 2016 November ; 94(11): 927–935. doi:10.1139/cjc-2015-0606.

Extension and validation of the GLYCAM force field parameters for modeling glycosaminoglycans

Arunima Singh, **Matthew B. Tessier**, **Kari Pederson**, **Xiaocong Wang**, **Andre P. Venot**, **Geert-Jan Boons**, **James H. Prestegard**, and **Robert J. Woods**

Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Road, Athens, GA 30602, USA

Abstract

Glycosaminoglycans (GAGs) are an important class of carbohydrates that serve critical roles in blood clotting, tissue repair, cell migration and adhesion, and lubrication. The variable sulfation pattern and iduronate ring conformations in GAGs influence their polymeric structure and nature of interaction. This study characterizes several heparin-like GAG disaccharides and tetrasaccharides using NMR and molecular dynamics simulations to assist in the development of parameters for GAGs within the GLYCAM06 force field. The force field additions include parameters and charges for a transferable sulfate group for O- and N-sulfation, neutral (COOH) forms of iduronic and glucuronic acid, and Δ4,5-unsaturated uronate (ΔUA) residues. ΔUA residues frequently arise from the enzymatic digestion of heparin and heparin sulfate. Simulations of disaccharides containing ΔUA reveal that the presence of sulfation on this residue alters the relative populations of $^{1}H_{2}$ and $^{2}H_{1}$ ring conformations. Simulations of heparin tetrasaccharides containing N-sulfation in place of N-acetylation on glucosamine residues influence the ring conformations of adjacent iduronate residues.

Résumé

Les glycosaminoglycanes (GAG) sont une classe importante d'hydrates de carbone qui jouent un rôle crucial dans la coagulation sanguine, la réparation des tissus, la migration et l'adhérence cellulaires, et la lubrification. La disposition variable des groupes sulfate et la conformation du cycle de l'iduronate des GAG influent sur leur structure polymérique et sur la nature des interactions. Dans la présente étude, nous caractérisons divers GAG disaccharidiques et tétrasaccharidiques semblables à l'héparine par RMN et modélisation de dynamique moléculaire en vue de contribuer à la détermination de paramètres pour les GAG dans le champ de force GLYCAM06. Les éléments additionnels au champ de force comprennent les paramètres et les charges associés à un groupe sulfate transférable lors de la O-sulfatation et de la N-sulfatation, les formes neutres (COOH) des acides iduronique et glucuronique et les résidus uronate Δ4,5-insaturés (ΔUA). Des résidus ΔUA sont souvent formés lors de la digestion enzymatique de

Corresponding author: Robert J. Woods (rwoods@ccrc.uga.edu).

This article is part of a Special Issue dedicated to Professor David Bundle in recognition of his seminal contributions and lifetime achievements in the fields of carbohydrate chemistry and glycobiology.

Supplementary material

Supplementary material is available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjc-2015-0606.

l'héparine et du sulfate d'héparine. Des modélisations de disaccharides contenant des ΔUA révèlent que la présence de groupes sulfate sur ces résidus modifie les populations relatives des conformations de cycle $^{1}H_{2}$ et $^{2}H_{1}$. Les modelisations de tétrasaccharides à base d'héparine présentant une N-sulfatation au lieu d'une N-acétylation des résidus glucosamine influent sur les conformations de cycle des résidus iduronate adjacents. [Traduit par la Rédaction]

Keywords

force field development; GLYCAM; glycosaminoglycans; heparin

Mots-clés

création d'un champ de force; GLYCAM; glycosaminoglycanes; héparine

Introduction

Glycosaminoglycans (GAGs) are linear polysaccharides, generally found covalently attached to proteins, forming a protein class called proteoglycans that are widely present on the plasma membrane, in the extracellular matrix, and in secretory granules of all animal cells.[1] GAGs can be classified into five main categories based on the unique composition of the polysaccharide: hyaluronan (HA), heparin/heparan sulfate (HS), chondroitin sulfate (CS), dermatan sulfate (DS), and keratan sulfate (KS). The polysaccharides are typically composed of repeating units of a hexosamine – uronic acid disaccharide. The hexosamine may be an N-sulfated or an N-acetylated glucosamine (HS, HS, and KS) or galactosamine (CA and DS), variably O-sulfated at the 3, 4, and (or) 6 positions. The uronic acid may be a glucuronic acid or an iduronic acid formed as a result of enzymatic epimerization of a glucuronic acid at the C-5 position. These uronic acid moieties may also be 2-O-sulfated. KS lacks uronic acids and instead contains variably sulfated galactose residues, and while most GAGs have a heterogeneous pattern of sulfation, HA is an unsulfated GAG.[2] GAG–protein interactions are critical in biological processes such as cell adhesion, anticoagulation, regulation of cell growth and proliferation, immobilization of proteins, maintenance of protein concentration gradient in regions of inflammation, viral invasion, and tumor metastasis.[2–8] Each tissue produces a distinctive repertoire of GAGs that interact with proteins in a tissue-specific manner. Most GAG-binding proteins interact with heparin[3] because due to structural similarity, it mimics the interaction of these proteins with the widely abundant cell surface HS chains.

GAG sulfation patterns have been demonstrated to modulate biological function, for example in the cases of heparan sulfate in growth factor activation and cellular defense,[5,9] CS growth factor recognition,[10] and synthetic HS in anticoagulant activity.[11] Differences in the sulfation pattern also alter the mode of interaction of heparin oligosaccharides with proteins, such as CCL5, where they have been shown to interact selectively with certain residues depending on the degree and pattern of sulfation.[12] This property consequently alters their ability to inhibit the interaction of CCL5 to its receptor CCR1. These sulfation patterns, in addition to altering the charge, also impact the 3D structure of GAG fragments.

Nuclear magnetic resonance (NMR) spectroscopy of GAG fragments has shown that sulfation patterns can alter the ring conformations of IdoA[13–15] and nonreducing terminal Δ4,5-unsaturated uronates (ΔUA)[16–18] that typically result from use of bacterial heparin lyase to cleave GAGs. Unsaturated uronate residues are often present in GAG–protein crystal structures and in GAG fragments employed in experimental binding studies. Recently published simulations of HS GAGs have shown that IdoA ring flipping can have a significant impact on the 3D shape of the GAG polymer;[19] however, as noted earlier,[20] not all differences in ring puckering lead to an altered overall shape.[21]

The variable levels and patterns of sulfation make the structural analysis of GAGs a challenge. This often limits experimental characterization of GAG structures to composition-based analyses of digested fragments of native GAGs. More detailed analysis, such as by NMR or crystallography, usually employ short, isolated, or synthetic oligomers, where the sulfation patterns are well controlled. Theoretical methods like molecular dynamics (MD) simulations have been widely used[22–24] to augment experimental methods in studying the conformational and binding properties of biomolecules, and the development of an accurate force field is key to the accuracy of these simulations. Molecular simulations employing a consistent and validated force field not only provide a basis for interpreting experimental NMR data but also enable structural analysis of polymers that are either too large for NMR analysis or too complex for routine synthetic preparation. Previously, existing carbohydrate force fields have been augmented in an ad hoc manner for examining specific sulfation patterns[25] and only recently have parameters been developed for transferable sulfate moieties.[26] In this work, we add two key features to the GLYCAM force field to enable accurate MD simulations of sulfated GAG sequences with AMBER.[27] The first addition is the creation of a generalizable sulfate parameter set to model N- and O-sulfation, including new bond, angle, and torsion terms as well as partial atomic charges, consistent with existing GLYCAM partial atomic charges.[28–30] The second is the development of force field parameters for ΔUA residues, which will permit simulation of this nonnaturally occurring residue. In addition, parameters for neutral (NH_2) and protonated (NH_3^+) glucosamine and protonated glucuronic and iduronic acids have been included.

To test the performance of the new parameters set, MD simulations were performed on variably sulfated GAG disaccharides containing ΔUA residues, and NMR scalar coupling and NOE measurements were collected for comparison with the theoretical data. MD and NMR data were collected for two synthetic GAG tetrasaccharides, with the aim of confirming the accuracy of the MD simulations and examining any influence of sulfation pattern on GAG conformation. The analysis presented here demonstrates that the new force field parameters reproduce the NMR data for a number of GAG fragments, both with and without terminal ΔUA. The simulations confirm the previous observation[13] that the IdoA ring populates two conformations ($^1C_4$ 63% and ^2S_O 37%) and surprisingly indicate that the terminal GlcA ring does not exclusively adopt the expected $^4C_1$ conformation.

Methods

NMR

NMR spectroscopy was carried out on a spectrometer operating at 18.8 T for disaccharides and 14.0 T for tetrasaccharides, equipped with a Varian Inova console and a 5 mm cryogenically cooled probe. 2,2-Dimethyl-2-silapentane-5-sulfonate (DSS) was included as an internal reference in each sample. NMR samples consisted of 0.5 mg of disaccharide in 100% D_2O buffer containing 20 mmol/L sodium phosphate and 1 mmol/L DSS, pH 6.5. The sample was shimmed to a DSS linewidth of <1 Hz. Proton resonances were assigned using a standard COSY experiment (Varian ChemPack), processed with NMRpipe[31] and assigned in Sparky.[32] 3J-coupling measurements were made from a 1D proton experiment with presaturation to suppress the signal due to any residual H_2O, collected with a spectral width of 9000 Hz and 32k points, processed, and analyzed in MestReNova.

Nuclear Overhauser effects (NOEs) were measured using a standard 2D NOESY experiment (Varian ChemPack) with a mixing time of 0.4 s, 512 increments, and 9000 points for disaccharides and 0.3 s, 512 increments, and 6000 points for tetrasaccharides, processed with NMRpipe. NOE peaks were integrated in NMRViewJ[33] and the distance was calibrated using the distance from the MD simulations between either the $H1^B$ and $H2^B$ or the $H2^B$ and $H3^B$ protons on the disaccharide reducing terminal residue (residue B) for **1**–**5**. Tetrasaccharides **6** and **7** were calibrated using the distance from the MD simulations between the H1 and H5 protons on the glucuronate (residue C), which was shown to be insensitive to fluctuations in ring conformations during the simulations.

Calculation of theoretical NMR properties

Theoretical NOEs were calculated using the isolated spin-pair approximation[34] in which NOE intensity is assumed to be proportional to $1/R^6$, where R is the distance between the two spin pairs. Based on previous NMR studies of GAG fragments, it can be assumed that the tetrasaccharides tumble isotropically.[35,36] Three-bond proton–proton scalar couplings ($^3J_{HH}$) were calculated using a Karplus-like equation developed by Haasnoot et al.[37] using the electronegativity values identified by Altona et al.[38] (eq. S4.1 and Table S4.3) (see Supplementary material section). Where relevant, experimental $^3J_{HH}$ couplings were decomposed into populations by least-squares fitting of the contributions from theoretical J values computed for each individual state.[39]

Molecular mechanics calculations

The SANDER program from the AMBER11 software package was used to compute the molecular mechanic energies associated with the parameter development. None of the one to four nonbonded interactions were scaled, and torsions were restrained at their desired values with a restraint weight of 5000 kcal/mol·rad^2. A 12 Å cutoff for nonbonded interactions was applied.

MD simulations

Initial structures for performing the MD simulations of the methyl glycosides for ensemble-averaged charge calculation were obtained from quantum mechanics (QM) optimized

models. Solutes were solvated with explicit TIP3P waters[40] with at least a 12 Å buffer between the glycan solute and each edge of the solvated cubic box using the LEaP module of AMER12.[27] Counterions were used to neutralize the net charge of each system. Energy minimization was performed under nVT conditions (500 steps steepest descent followed by 24 500 steps of conjugate-gradient minimization). Each system was then heated under nPT conditions for 50 ps, raising the temperature from 0 to 300 K, followed by 100 ps of equilibration while the temperature was maintained at 300 K. All simulations used periodic boundary conditions where the pressure was maintained at 1.0 atm, the external dielectric was set to 1.0, and the system compressibility was set to that of water. The Berendsen thermostat[41] was used for all temperature controls and the SHAKE algorithm[42] was used to constrain bonds with hydrogens, allowing a 2 fs timestep to be used. Nonbonded scaling factors were set to unity, and a 10.0 Å nonbonded cutoff was employed in all steps. Minimization and equilibration were performed using the PMEMD[43] implementation for CPU in AMBER12. Subsequently, production simulations were performed with the PMEMD-Cuda[43] implementation for GPUs.

The simulations of heparin disaccharides **1**–**5** and tetrasaccharides **6** and **7** were performed using a similar protocol, except for the minimization steps. The first minimization step was performed in generalized Born implicit solvent[44] with an infinite nonbonded cutoff, prior to addition of counterions and explicit solvent. A second minimization step was performed after each system was explicitly solvated and neutralized.

QM calculations

All QM calculations were performed using the Gaussian 09[45] software package.

Parameter development (partial charges)

Partial atomic charges were derived from the restrained electrostatic potential (RESP) charge fitting methodology.[46] The ESPs for the small molecules employed in parameter development were computed from the lowest energy conformational state at the HF/cc-pVTZ level of theory with a RESP weight of 0.0005. For anionic monosaccharides, ESPs were computed with diffuse functions at the HF/6-31++G**//HF/6-31++G** level, whereas for neutral and cationic monosaccharides, calculations were performed at the HF/6-31G*//HF/6-31G* level; in each case, a RESP weight of 0.01 was employed to be compatible with GLYCAM06.

Charge models for *N*- and *O*-sulfates, glucosamine, and ΔUA were developed using the standard GLYCAM ensemble-averaged charge method.[28] The charges were developed for sulfates using 4-O- and 6-O-sulfated β-D-GalNAc and both anomers of N-sulfated α- and β-D-glucosamine (D-GlcNS) using initial glycan geometries extracted from co-crystallized protein–sugar complexes. For the ensemble-averaged charge calculation, an initial QM-optimized structure was used to derive single-point RESP charges and employed for 10–50 ns of MD simulations, as required, for adequate sampling of exocyclic rotamers. From the simulations, 100 evenly spaced snapshots were extracted as a representative ensemble of the 3D structures. Each of these geometries was subjected to QM optimization with all torsion

angles frozen in their MD conformation. RESP charges were calculated for each frame and averaged to get the ensemble-averaged charge set for each particular molecule.

The computed charges for the sulfate moieties (SO_3^-) in both *N*- and *O*-sulfates were within statistical variance of each other, allowing the creation of an interchangeable sulfate residue. Examination of the sulfated sugars revealed similar atomic charges on the sulfated and nonsulfated atoms in GLYCAM06.[29] The most significant deviation between them was associated with the oxygen or nitrogen atom at the point of sulfate attachment. Consequently, for transferability, the charge on the linking heteroatom was adjusted as necessary to achieve a net integer charge on each sulfated sugar (Table S4.4).

Charges for protonated α- and β-D-glucosamine ($GlcNH_3^+$) were similarly developed and found to significantly vary from the GLYCAM charges for α- and β-glucose and *N*-acetyl-glucosamine, particularly for the ring carbon atoms (Table S4.4). This variation suggests that such analogs require unique charge sets for each monosaccharide, which is not surprising, as the positively charged site is directly adjacent to the sugar ring.

Charges for the ΔUA monomers were obtained by averaging the charges for each of the low-energy half-chair states $^1H_2$ and $^2H_1$ (Fig. 1*a*; Table S4.4).

Parameter development (atom, bond, angle, and torsion parameters)

The GLYCAM06 force field for carbohydrates[29] and lipids[30] was adapted to include new terms required to model the double bonds found in unsaturated uronic acids. The only new atom type added in this work was the sulfate sulfur atom for which the van der Waals parameters were transferred from the sulfate atom type (S1) found in parm99.[47] All valence and torsion terms were developed using the hierarchical development procedure outlined in two prior GLYCAM06 publications[29,30] wherein bond terms are developed first followed by angle and torsion terms (Tables S4.1 and S4.2). Small molecules were selected for parameter development such that each contained as few new terms as possible while maintaining an electronic environment relevant to the carbohydrate.[29] Equilibrium values for bonds and angles were obtained from the averages of crystal structures found in the Cambridge Structural Database[48] with molecule IDs HEMKEP, KOCOJ, SRHXGU, MIZFUX, GUVFOS, GUVFEI, GUVFAE, and ZULPIF (Table S4.1); force constants were derived by fitting to QM data computed at the B3LYP/6-31++g(2d,2p)//HF/6-31++g(2d,2p) level.

Torsion potentials were generated for the relevant bonds in the molecules found in Table S4.2. Rotations were sampled in 30° increments with the exception of terms describing double bond rotations in which only 0°, 90°, and 180° orientations were used to characterize the *cis*/*trans* relative energies and the barrier height between them. All torsion terms were developed without the use of a phase shift adjustment. Exocyclic torsion terms were developed using tetrahydropyran or its unsaturated analog of ΔUA. In the case of *N*-sulfate parameters, planarity of the nitrogen atom was maintained during the QM torsion rotation to reflect solution conformations of an sp^2-hybridized nitrogen.

In the development of the torsion terms associated with the unsaturated bond, a better fit to experimentally observed rotamer preferences required the use of a higher level of QM theory

(second-order Moller–Plesset, MP2) during geometry optimization. This was the case for rotation about the central Cg–Os bond in the Os–Cg–Os–Ck sequence (Fig. 1*b*), where the terminal Os–Cg bond did not favor the experimentally observed rotamers at the HF level. This behavior was corrected by performing geometry optimizations at the MP2 level. The need for a higher QM level that includes electron correlation may reflect the presence of hyperconjugation between the oxygen atoms (Os) and the unsaturated carbon center (Ck). Having observed a dependence of rotamer preference on the level of QM theory for this term, all other terms were reexamined and found not to show any notable dependence on the QM level. The energy contributions to the barrier for *cis*/*trans* rotation in double bonds were distributed equally between heavy and light atom terms, Cg–Ck–Ck–C and Ha–Ck–Ck–C, avoiding the need for improper torsions.[30]

Tuning torsion terms to reproduce solution populations for ΔUA

MD simulations (100 ns) with the preliminary parameters were collected for disaccharides **1** and **2** (Fig. 2), which contain ΔUA residues. An analysis of the populations of the $^{1}H_{2}$ and $^{2}H_{1}$ ring states ($^{1}H_{2}$:$^{2}H_{1}$ = 70:30 and 35:65, respectively) showed poor agreement with the NMR-derived populations, 40:60 and 69:31, respectively. As the partial atomic charges in the ΔUA residue had been derived under the assumption of an equal population of half-chair states, this appeared to be a potential source of error. However, the populations from MD simulations, in which the contribution of the partial charges from each half-chair was varied from 0% $^{1}H_{2}$ to 100% $^{1}H_{2}$, were relatively insensitive to the atomic charges.

Subsequently, the ring torsion terms (Oh–Cg–Cg–Ck (**1**) and Os–Cg–Cg–Ck (**2**)) were iteratively adjusted so as to obtain optimal agreement with the NMR populations. MD simulations with the optimized torsion terms yielded average population ratios ($^{1}H_{2}$: $^{2}H_{1}$) for **1** and **2** of 42:58 and 67:33, respectively, that were then in good agreement with experimental values (Table 1). All subsequent simulations employed these parameters.

Results and discussion

Conformational analysis of ΔUA disaccharides (3–5)

GAG disaccharides **3**–**5** (Fig. 3) were analyzed using NMR and MD simulations to validate ring conformational populations and glycosidic linkage geometry profiles obtained using the new parameters. Ring conformations and populations were determined from homonuclear $^{3}J_{HH}$ couplings, while NOEs were collected to characterize the global 3D shape of these GAG fragments.

Ring state populations for the ΔUA residue from MD simulations showed a preference for the $^{1}H_{2}$ state in all cases, consistent with the populations derived from NMR *J* couplings (Table 2). Conformational analysis of the ^{3}J couplings for the GlcNx residues (data not shown) was consistent with the $^{4}C_{1}$ conformation exclusively.

NMR characterization of the glycosidic linkages was provided by an analysis of $H1^{A}$–$H3^{B}$ and $H1^{A}$–$H5^{B}$ proton–proton NOE contacts observed for *trans*-glycosidic interactions. Comparison of the theoretical- and NMR-derived NOE distances for these protons shows agreement within 0.6 Å in **4** and **5** (Table 2).

The φ and ψ population distribution heat maps presented in Fig. 4 show that the global minimum from the MD simulations is approximately φ = 50° and ψ = 0° for all of the linkages. Each disaccharide also shows a second stable anti-ψ state near φ = 50° and ψ = 180°.

Conformational analysis of heparin tetrasaccharides

The conformational properties of two synthetic GAG tetrasaccharides, a naturally occurring sequence, **7**, and its non-naturally occurring variant, **6**, that differ only in the presence (**7**) or absence (**6**) of N-sulfation (Fig. 5) were characterized by NMR. To deconvolute the NMR data, MD simulations were performed on **6** and **7** for each of the three common IdoA ring conformations ($^1C_4$, ^2S_O and $^4C_1$) in each tetrasaccharide.

Ring conformational analysis

During the 1 μs MD simulations, residues C, D, and F each populated only the $^4C_1$ ring conformation, and back-calculation of the *J* couplings led to agreements with the experimental values within 0.5 Hz for D and F (Table 3). However, the theoretical H1–H2 *J* values for the terminal C residue (9.8 Hz) in both **6** and **7** were larger than those observed experimentally by almost 2 Hz. Given the otherwise close agreements, this suggests that the MD simulation may not have detected all of the conformations adopted by the C-rings, despite the relatively long simulation time. In the $^4C_1$ conformation, protons H1 and H2 in the β-anomer of the GlcA ring are *anti* to each other, leading to a large *J* coupling, whereas in the $^1C_4$ conformation, they would be *gauche*, leading to minimal coupling, and a mixture of approximately 80% $^4C_1$ and 20% $^1C_4$ would explain the observed *J* value in the C residue. However, in the absence of further experimental data, this is not necessarily a unique solution. It is notable that, at least in the case of a fully sulfated GlcA residue, NMR data indicated that the uronate preferred to adopt the $^1C_4$ conformation rather than the $^4C_1$.[49]

Optimal agreement between the experimental and theoretical *J* values for the IdoA (residue E) ring structure in **7** was achieved using least-squares fitting analysis of the contributions from multiple ring forms, resulting in a population distribution ($^1C_4$:^2S_O) of 63:37 with no contribution from $^4C_1$. The absence of the $^4C_1$ state is supported by NMR data for similar GAGs,[13,20,21,50,51] which indicate this state to be the least populated of the three, if present at all. A search of the Protein Databank[52] revealed the $^1C_4$ (73%) and ^2S_O (24%) states to be the dominant forms of IdoA. For **6**, only the $^3J_{H1H2}$ and $^3J_{H4H5}$ couplings were experimentally observed (0.9 and 2.4 Hz, respectively), and both were below the theoretical values computed from any of the ring conformations. Nevertheless, the small value of the experimental H1–H2 coupling indicates that there cannot be significant amounts of either the ^2S_O or $^4C_1$ conformations present.

Interresidue conformational analysis

The theoretical interproton distances (Table 4) showed agreement to within 0.5 Å of the NMR-derived values for all but the distance between protons H1 and H4 (0.8 Å) in residues D and E of **6**. The theoretical distances were very similar between the $^1C_4$ and ^2S_O conformations for the NOEs between residues D and E and E and F, indicating that these IdoA ring conformations do not have a significant influence on the overall shape of the

tetrasaccharide. This has previously been determined to be the case in NMR structures of heparin dodecasaccharides.[21]

To understand the role of N-sulfation on the glycosidic linkages, heat maps were plotted for the φ versus ψ values for the three glycosidic linkages in each trajectory (Fig. 6). The glycosidic linkages between residues E and F showed very similar distributions for both **6** and **7**. In addition to the observed major conformation (φ ≈ 0°−60° and ψ ≈ −60°−60°), this linkage also sampled both the anti-*exo* (φ ≈ −90°−0° and ψ ≈ −60°−0°) and the anti-ψ (ψ ≈ −150°−+150°) states. Only the simulation of **6** restrained in the $^{2}S_{O}$ conformation did not sample the anti-ψ state for this linkage. The percent distribution of each state is presented in Table 5.

The linkage between residues C and D in both **6** and **7** also showed very similar distribution, with an additional sparsely populated anti-φ state (φ ≈ −150° to +150°) for the $^{1}C_{4}$ conformation. This state was also observed in **7** in the $^{4}C_{1}$ simulation, but not in **6**. Additionally, the anti-ψ orientation was missing for **6** in $^{2}S_{O}$ and anti-φ was missing for **7** in $^{2}S_{O}$. Overall, N-sulfation had little impact on the preferences of the glycosidic linkages, with the possible exception of the E–F linkage in **7**, where regardless of the conformation of the IdoA ring, there appeared to be a modest increase in the population of the anti-ψ conformation (bold entries in Table 5).

For the GlcNx (D) and IdoA (E) linkage (Fig. 6*b*), each ring shape showed a slightly different distribution of glycosidic angles. While the $^{2}S_{O}$ conformation showed the tightest distribution around φ ≈ −50° and ψ ≈ −50°, $^{1}C_{4}$ showed a wider spread of the φ angle and $^{4}C_{1}$ a wider spread of the ψ angle. Each of the three also showed a sparsely populated distribution around φ ≈ 40° and ψ ≈ 0°. For the simulation with $^{2}S_{O}$ conformation, the D–E linkage also sampled the anti-psi state for both 6 and 7, albeit differing in the percent distribution.

Conclusions

A new parameter set for GAGs containing iduronic acid, ΔUA, sulfate, and protonated glucuronic and iduronic acids has been added to GLYCAM. The development of a transferable sulfate model allows it to be used for multiple attachment points without a need for development of separate charge sets. In addition, development of a single model for ΔUA that reproduces solution conformations permits more accurate modeling of these residues.

The performance of the new parameter set was tested by performing MD simulations on variably sulfated GAG disaccharides containing ΔUA residues and two synthetic GAG tetrasaccharides. NMR scalar coupling and NOE measurements were collected for comparison with the theoretical data, with the aim of verifying the accuracy of the MD simulations and examining any influence of sulfation pattern on GAG conformation.

Unrestrained simulations of ΔUA on timescales that allowed direct parameterization of the ring populations were performed. Analysis of NMR J_{HH} couplings showed that the conformation populations of the ΔUA ring are largely insensitive to the adjacent sulfation

patterns and the N-substituent; however, presence (**2**–**5**) or absence (**1**) of 2-O-sulfation on ΔUA altered the favored geometry.

The most notable effect of sulfation, in the case of GAG tetrasaccharides, was observed on the ring geometries for IdoA. Examination of the NMR data showed that tetrasaccharide **6**, which contained 2-O-sulfated IdoA but no N-sulfated glucosamine residues, exclusively favored the $^{1}C_{4}$ conformation, while **7**, which contained two N-sulfated glucosamine residues adjacent to the 2-O-sulfated IdoA, sampled a substantial $^{2}S_{O}$ population (37%). While the ring flip dynamics were not captured by this work, long timescale simulations of IdoA using GLYCAM have previously shown experimentally consistent ring populations.[53] The $^{3}J_{HH}$ coupling analysis also suggested that the terminal GlcA ring may not exclusively adopt the expected $^{4}C_{1}$ conformation.

These parameters and related structure files are available for download from the GLYCAM website (www.glycam.org).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the National Institutes of Health for support (R01 GM100058, P41 GM103390).

References

1. Kjellén L, Lindahl U, et al. Annu. Rev. Biochem. 1991; 60:443. [PubMed: 1883201]
2. Esko, JD., Kimata, K., Lindahl, U. Essentials of Glycobiology. 2nd. Varki, A.Cummings, R., Esko, J., editors. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2009.
3. Esko, JD., Linhardt, RJ. Essentials of Glycobiology. 2nd. Varki, A.Cummings, R., Esko, J., editors. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2009.
4. San Antonio JD, Iozzo RV. eLS. 2001
5. Bernfield M, Götte M, Park PW, Reizes O, Fitzgerald ML, Lincecum J, Zako M. Annu. Rev. Biochem. 1999; 68:729. [PubMed: 10872465]
6. Raman R, Sasisekharan V, Sasisekharan R. Chem. Biol. 2005; 12(3):267. [PubMed: 15797210]
7. Rostand KS, Esko JD. Infect. Immun. 1997; 65(1):1. [PubMed: 8975885]
8. Schmidtchen A, Frick IM, Bjorck L. Mol. Microbiol. 2001; 39(3):708. [PubMed: 11169110]
9. Kreuger J, Spillmann D, Li J, Lindahl U. J. Cell Biol. 2006; 174(3):323. [PubMed: 16880267]
10. Gama CI, Tully SE, Sotogaku N, Clark PM, Rawat M, Vaidehi N, Goddard WA, Nishi A, Hsieh-Wilson LC. Nat. Chem. Biol. 2006; 2(9):467. [PubMed: 16878128]
11. van Boeckel CAA, Petitou M. Angew. Chem. Int. Ed. 1993; 32(12):1671.
12. Singh A, Kett WC, Severin IC, Agyekum I, Duan J, Amster IJ, Proudfoot AEI, Coombe DR, Woods RJ. J. Biol. Chem. 2015; 290(25):15421. [PubMed: 25907556]
13. Ferro DR, Provasoli A, Ragazzi M, Casu B, Torri G, Bossennec V, Perly B, Sinaÿ P, Petitou M, Choay J. Carbohydr. Res. 1990; 195(2):157. [PubMed: 2331699]
14. Blanchard V, Chevalier F, Imberty A, Leeflang BR, Sugahara K, Kamerling JP. Biochemistry. 2007; 46(5):1167. [PubMed: 17260946]
15. Sattelle BM, Shakeri J, Roberts IS, Almond A. Carbohydr. Res. 2010; 345(2):291. [PubMed: 20022001]
16. Bazin HG, Capila I, Linhardt RJ. Carbohydr. Res. 1998; 309(2):135. [PubMed: 9741074]

17. Jin L, Hricovíni M, Deakin JA, Lyon M, Uhrín D. Glycobiology. 2009; 19(11):1185. [PubMed: 19648354]
18. Mikhailov D, Linhardt RJ, Mayo KH. Biochem. J. 1997; 328(1):51. [PubMed: 9359833]
19. Sattelle BM, Shakeri J, Almond A. Biomacromolecules. 2013; 14(4):1149. [PubMed: 23439078]
20. Mulloy B, Forster MJ, Jones C, Drake AF, Johnson EA, Davies DB. Carbohydr. Res. 1994; 255:1. [PubMed: 8181000]
21. Mulloy B, Forster MJ, Jones C, Davies DB. Biochem. J. 1993; 293:849. [PubMed: 8352752]
22. Cheatham TEI. Curr. Opin. Struct. Biol. 2004; 14(3):360. [PubMed: 15193317]
23. Fadda E, Woods RJ. Drug Discov. Today. 2010; 15(15–16):596. [PubMed: 20594934]
24. Feller SE. Curr. Opin. Colloid Interface Sci. 2000; 5(3–4):217.
25. Gandhi NS, Mancera RL. Carbohydr. Res. 2010; 345(5):689. [PubMed: 20097328]
26. Mallajosyula SS, Guvench O, Hatcher E, Mackerell AD. J. Chem. Theory Comput. 2012; 8(2):759. [PubMed: 22685386]
27. Case, DA., Babin, V., Berryman, JT., Betz, RM., Cai, Q., Cerutti, DS., Cheatham, TE., III, Darden, TA., Duke, RE., Gohlke, H., Goetz, AW., Gusarov, S., Homeyer, N., Janowski, P., Kaus, J., Kolossváry, I., Kovalenko, A., Lee, TS., LeGrand, S., Luchko, T., Luo, R., Madej, B., Merz, KM., Monard, G., Needham, P., Nguyen, H., Nguyen, HT., Omelyan, I., Onufriev, A., Roe, DR., Roitberg, A., Salomon-Ferrer, R., Simmerling, CL., Smith, W., Swails, J., Walker, RC., Wang, J., Wolf, RM., Wu, X., York, DM., Kollman, PA. AMBER 2015. San Francisco, CA: University of California; 2015.
28. Basma M, Sundara S, Calgan D, Vernali T, Woods RJ. J. Comput. Chem. 2001; 22(11):1125. [PubMed: 17882310]
29. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL, Woods RJ. J. Comput. Chem. 2008; 29(4):622. [PubMed: 17849372]
30. Tessier MB, Demarco ML, Yongye AB, Woods RJ. Mol. Simul. 2008; 34(4):349. [PubMed: 22247593]
31. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. J. Biomol. NMR. 1995; 6(3):277. [PubMed: 8520220]
32. Goddard, TD., Kneller, DG. San Francisco, CA: University of California;
33. Johnson BA. Methods Mol. Biol. 2004; 278:313. [PubMed: 15318002]
34. Macura S, Ernst RR. Mol. Phys. 1980; 100(1):135.
35. Hricovíni M, Torri G. Carbohydr. Res. 1995; 268(2):159. [PubMed: 7736468]
36. Lucas R, Angulo J, Nieto PM, Martín-Lomas M. Org. Biomol. Chem. 2003; 1(13):2253. [PubMed: 12945695]
37. Haasnoot CAG, de Leeuw FAAM, Altona C. Tetrahedron. 1980; 36(19):2783.
38. Altona C, Francke R, de Haan R, Ippel JH, Daalmans GJ, Hoekzema AJAW, van Wijk J. Magn. Reson. Chem. 1994; 32(11):670.
39. Yongye AB, Foley BL, Woods RJ. J. Phys. Chem. A. 2008; 112(12):2634. [PubMed: 18311953]
40. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. J. Chem. Phys. 1983; 79:926.
41. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. J. Chem. Phys. 1984; 81(8): 3684.
42. Ryckaert J-P, Ciccotti G, Berendsen HJ. J. Comput. Phys. 1977; 23(3):327.
43. Götz AW, Williamson MJ, Xu D, Poole D, Le Grand S, Walker RC. J. Chem. Theory Comput. 2012; 8(5):1542. [PubMed: 22582031]
44. Onufriev A, Bashford D, Case DA. Proteins. 2004; 55(2):383. [PubMed: 15048829]
45. Frisch, MJ., Trucks, GW., Schlegel, HB., Scuseria, GE., Robb, MA., Cheeseman, JR., Scalmani, G., Barone, V., Mennucci, B., Petersson, GA., Nakatsuji, H., Caricato, M., Li, X., Hratchian, HP., Izmaylov, AF., Bloino, J., Zheng, G., Sonnenberg, JL., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, JA., Jr, Peralta, JE., Ogliaro, F., Bearpark, M., Heyd, JJ., Brothers, E., Kudin, KN., Staroverov, VN., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, JC., Iyengar, SS., Tomasi, J., Cossi, M., Rega, N., Millam, JM., Klene, M., Knox, JE., Cross, JB., Bakken, V.,

Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, RE., Yazyev, O., Austin, AJ., Cammi, R., Pomelli, C., Ochterski, JW., Martin, RL., Morokuma, K., Zakrzewski, VG., Voth, GA., Salvador, P., Dannenberg, JJ., Dapprich, S., Daniels, AD., Farkas, Ö., Foresman, JB., Ortiz, JV., Cioslowski, J., Fox, DJ. Gaussian 09, Revision E.01. Wallingford, CT: Gaussian, Inc; 2009.

46. Cornell WD, Cieplak P, Bayly CI, Kollmann PA. J. Am. Chem. Soc. 1993; 115(21):9620.
47. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Proteins. 2010; 78(8):1950. [PubMed: 20408171]
48. Allen FH. Acta Cryst. 2002; B58:380.
49. Maruyama T, Toida T, Imanari T, Yu G, Linhardt RJ. Carbohydr. Res. 1998; 306(1–2):35. [PubMed: 9691438]
50. Ferro DR, Provasoli A, Ragazzi M, Torri G, Casu B, Gatti G, Jacquinet JC, Sinay P, Petitou M, Choay J. J. Am. Chem. Soc. 1986; 108(21):6773.
51. Gatti G, Casu B, Hamer GK, Perlin AS. Macromolecules. 1979; 12(5):1001.
52. Berman HM. Nucleic Acids Res. 2000; 28(1):235. [PubMed: 10592235]
53. Sattelle BM, Bose-Basu B, Tessier M, Woods RJ, Serianni AS, Almond A. J. Phys. Chem. B. 2012; 116(22):6380. [PubMed: 22577942]

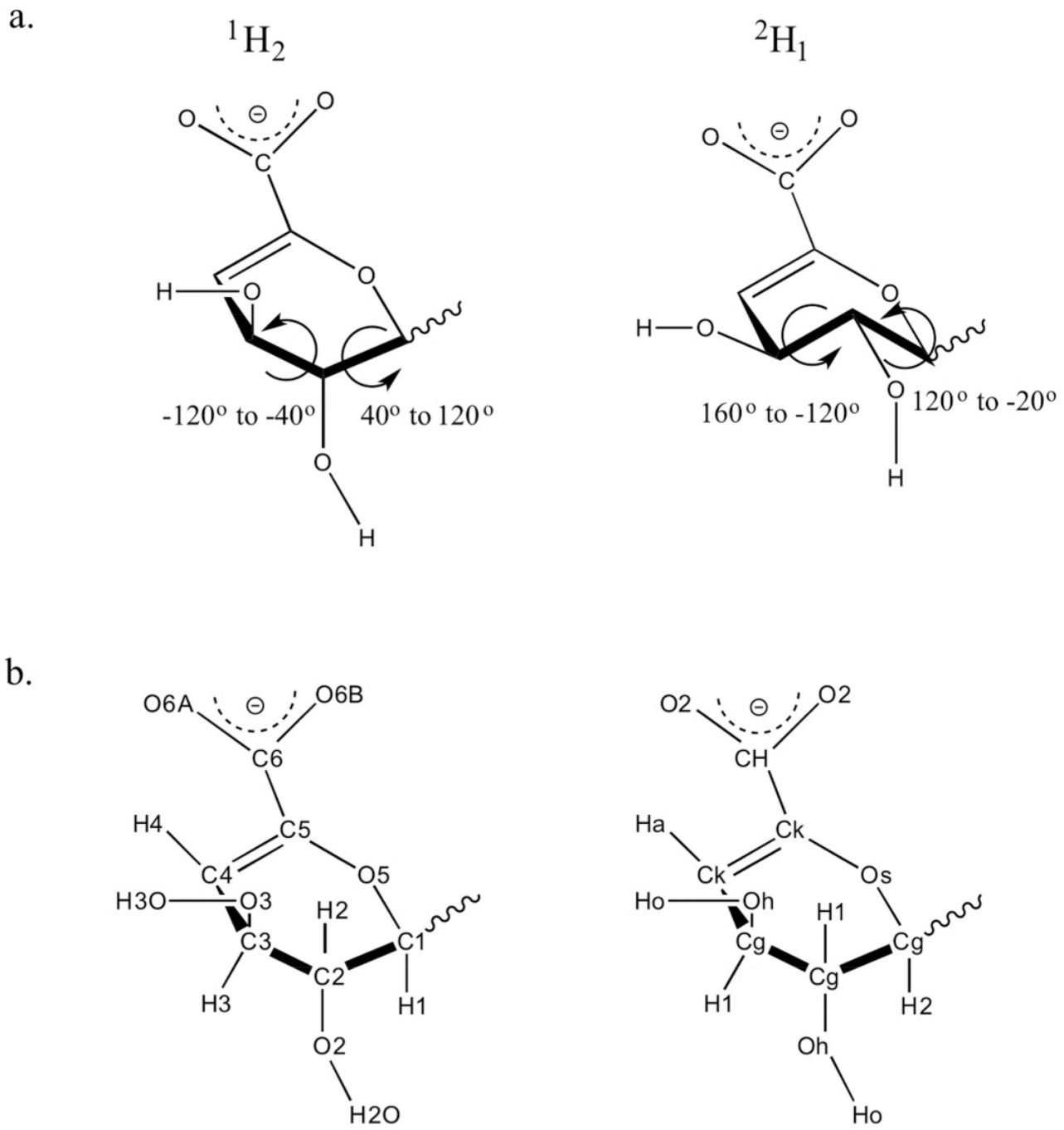

Fig. 1.

(*a*) ΔUA ring conformations with the torsion angle ranges typically associated with H_1–C_1–C_2–H_2 and H_2–C_2-C_3–H_3 atomic sequences. (*b*) Atom names (left) and atom types (right) employed in GLYCAM for ΔUA.

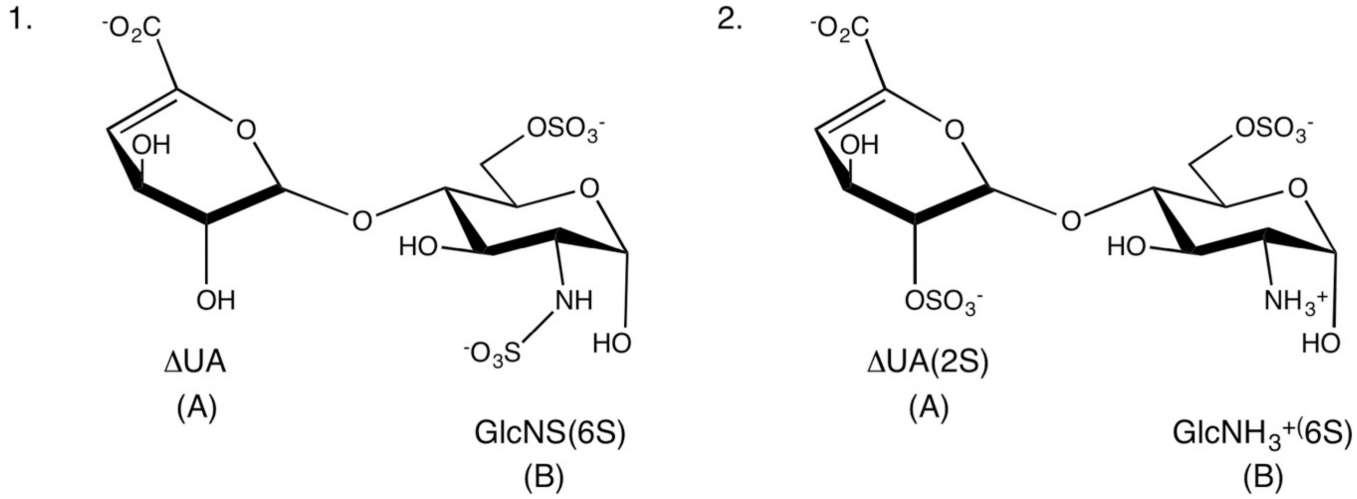

Fig. 2.

Schematic structures of GAG disaccharides (**1** and **2**), ΔUA monosaccharide (A), and glucosamine residue (B).

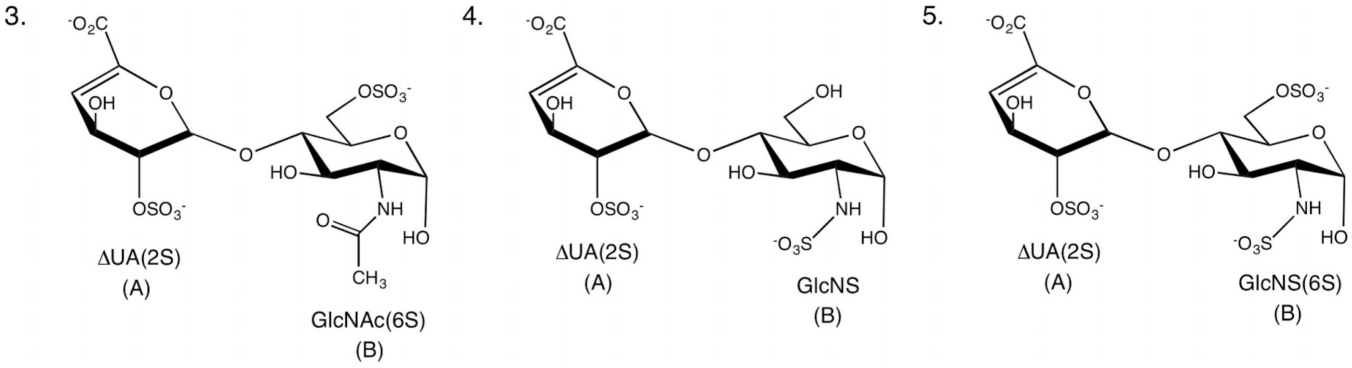

Fig. 3.
Schematic structures of GAG disaccharides (**3**–**5**).

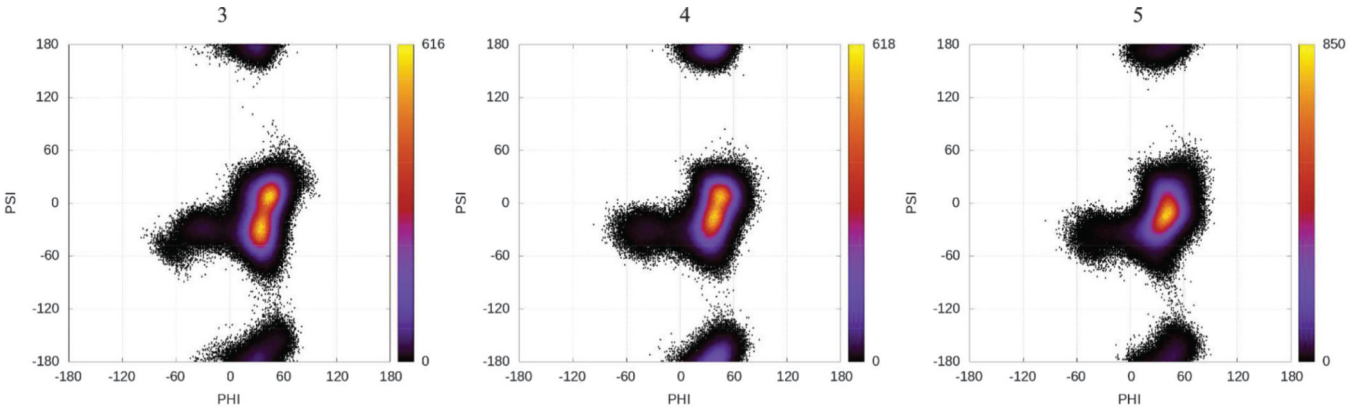

Fig. 4.

Heat maps for φ versus ψ angles for the glycosidic linkage between A and B for **3**, **4**, and **5**.

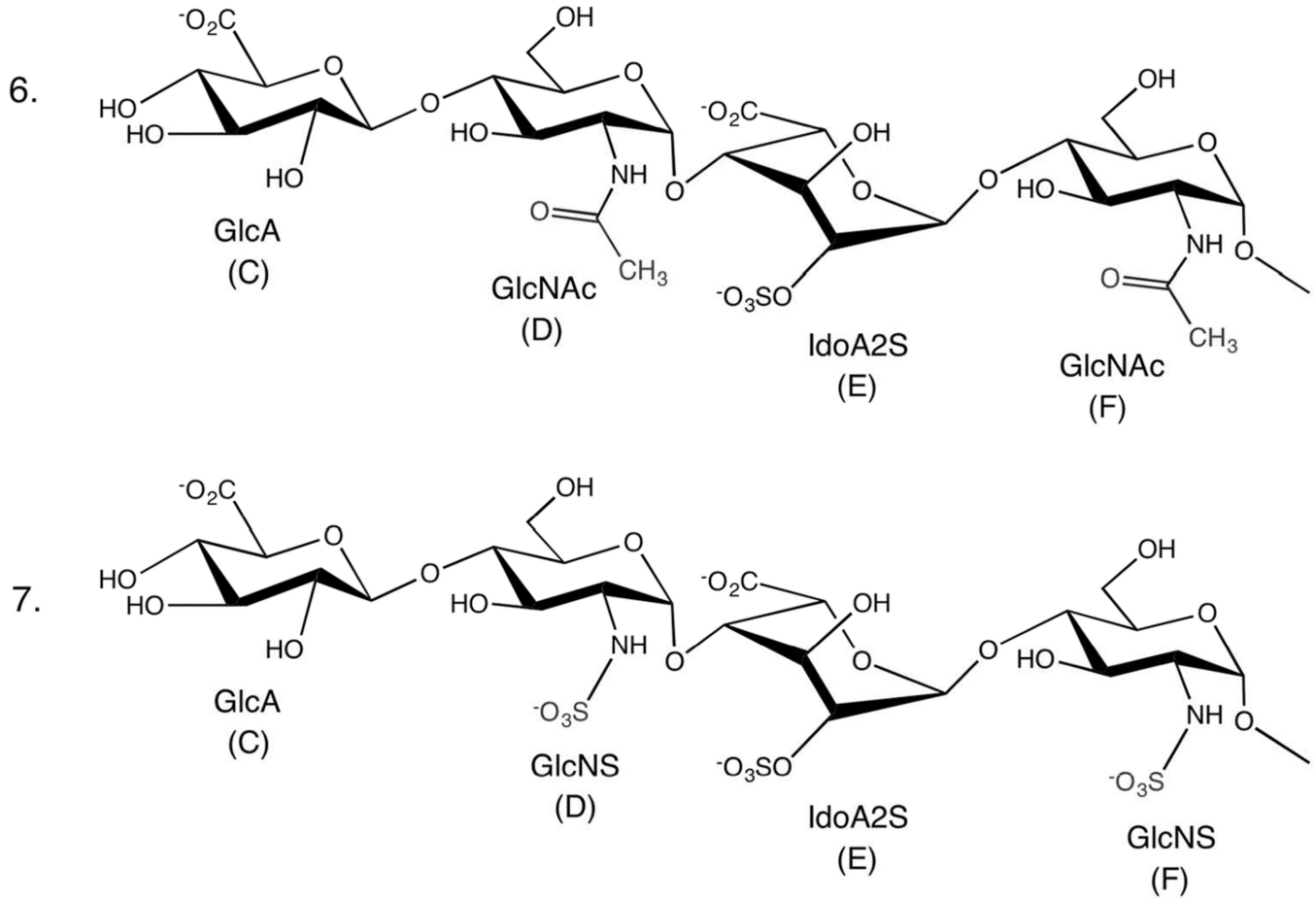

Fig. 5.

Schematic structures of GAG tetrasaccharides (**6** and **7**). Labels C–F are used to identify the monosaccharide residues.

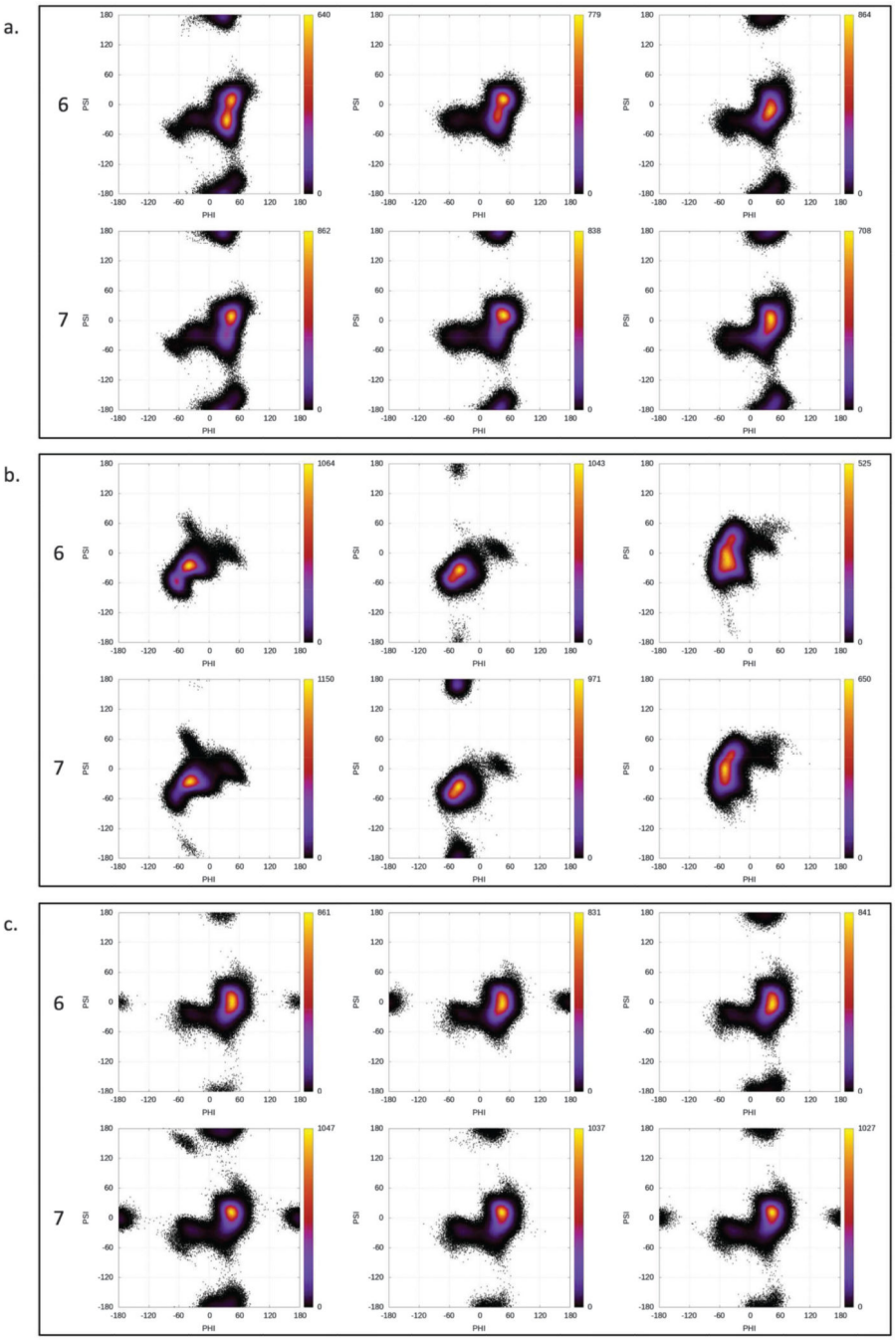

Fig. 6.

Heat maps for φ versus ψ angles for the glycosidic linkages between (*a*) C and D, (*b*) D and E, and (*c*) E and F for **6** and **7**.

Table 1

NMR *J* couplings and ring state populations for the ΔUA residue in **1** and **2**.

	1		**2**	
	NMR	**Optimized theoretical**	**NMR**	**Optimized theoretical**
^3J coupling (Hz)[a]				
H_1–C_1–C_2–H_2	5.5	4.6	3.4	3.4
H_2–C_2–C_3–H_3	4.9	5.0	2.8	3.4
H_3–C_3–C_4–H_4	3.8	3.4	4.4	4.4
Population ($^1H_2$:$^2H_1$)	40:60	42:58	69:31	67:33
NOEs (Å)				
H_1^A–H_3^B	3.0	4.2	3.1	4.4
H_1^A–H_5^B		3.9	2.8	3.9

[a] *J* couplings for H1–H2, H2–H3, and H3–H4 in the $^1H_2$ ring form are 1.7, 1.2, and 5.8 Hz, respectively, and for the $^2H_1$ form are 6.7, 7.8, and 1.6 Hz, respectively. The *J* values were independent of the anomeric configuration (α or β) at the reducing terminus.

Table 2

Experimental observables for residue A of the ΔUA containing disaccharides **3–5**.

	3		**4**		**5**	
Residue A	**NMR**	**Theoretical**	**NMR**	**Theoretical**	**NMR**	**Theoretical**
^3J coupling (Hz)[a]						
H1–H2	3.4	2.7	3.0	2.7	3.7	2.9
H2–H3		2.5	2.5	2.5	2.8	2.8
H3–H4	4.7	5.0	4.7	5.0	4.4	4.8
Population ($^1H_2$:$^2H_1$)	70:30	80:20	76:24	80:20	67:33	76:24
NOEs (Å)						
$H1^A$–$H3^B$		3.5	2.7	3.3	2.4	2.9
$H1^A$–$H5^B$		3.5	3.1	3.6		3.3

[a] J couplings for H1–H2, H2–H3, and H3–H4 in the $^1H_2$ ring form are 1.7, 1.2, and 5.8 Hz, respectively, and for the 2H1 form are 6.7, 7.8, and 1.6 Hz, respectively. The J values were independent of the anomeric configuration (α or β) at the reducing terminus.

Table 3

Theoretical and experimental ^{3}J couplings for **6** and **7**.

		6				**7**			
Residue	^{3}J coupling	Theoretical ($^{4}C_{1}$)[a]			NMR	Theoretical ($^{4}C_{1}$)[a]			NMR
C	H1–H2	9.8			7.9	9.8			7.9
D	H1–H2	3.5			3.6	3.4			3.6
D	H2–H3	10.1			10.6	10.1			10.2
F	H1–H2	3.5			3.6	3.4			3.5
F	H2–H3	10.0			na	10.0			10.2
F	H4–H5	10.0			na	10.0			9.5
Residue	^{3}J coupling	$^{1}C_{4}$	$^{2}S_{O}$	$^{4}C_{1}$	NMR	$^{1}C_{4}$	$^{2}S_{O}$	$^{4}C_{1}$	NMR
E	H1–H2	1.7	8.0	9.9	0.9	1.7	7.9	9.9	3.2
E	H2–H3	1.8	10.0	9.7	na	1.8	10.0	9.7	5.9
E	H3–H4	2.0	6.6	10.0	na	2.0	6.6	10.0	3.7
E	H4–H5	3.4	4.5	4.0	2.4	3.4	4.5	4.0	2.7

[a]Only the $^{4}C_{1}$ conformation was sampled during the MD simulations.

Table 4

Theoretical and experimental interring NOE distances (Å) for **6** and **7**.

		6			**7**		
C	D	Theoretical		NMR	Theoretical		NMR
H1	H4	2.3		2.8	2.4		2.9
H1	H61/2	2.9		2.8	2.7		2.8
D	E	$^1C_4$	^2S_O	NMR	$^1C_4$	^2S_O	NMR
H1	H2	4.6	4.8	4.2	4.9	4.8	N/A
H1	H3	2.4	2.7	2.5	2.5	2.7	2.7
H1	H4	2.5	2.5	3.3	2.3	2.6	2.7
H5	H4	3.1	3.2	2.7	3.4	3.2	na
E	F	$^1C_4$	^2S_O	NMR	$^1C_4$	^2S_O	NMR
H1	H3	3.2	4.3	na	3.1	3.2	2.6
H1	H4	2.3	2.3	2.7	2.3	2.4	2.6
H1	H61/2	3.0	3.0	2.7	2.9	2.9	2.7

Table 5

Percent distribution of the φ versus ψ states sampled by the glycosidic linkages during of **6** and **7** during the MD simulations performed with each conformation of IdoA.

Tetrasaccharide	Conformation	Linkage	Exo	Anti-exo	Anti-ψ	Anti-φ
6	$^{1}C_{4}$	C–D	97.6	2.2	0.1	0.1
		D–E	99.0	1.0	0.0	0.0
		E–F	92.4	2.2	5.4	0.0
	$^{2}S_{O}$	C–D	97.2	1.9	0.0	0.9
		D–E	99.3	0.7	0.1	0.0
		E–F	97.5	2.5	0.0	0.0
	$^{4}C_{1}$	C–D	95.9	2.5	1.7	0.0
		D–E	99.3	0.7	0.0	0.0
		E–F	94.8	2.1	3.1	0.0
7	$^{1}C_{4}$	C–D	87.3	4.1	6.3	2.3
		D–E	94.7	5.3	0.0	0.0
		E–F	87.9	3.9	8.2	0.0
	$^{2}S_{O}$	C–D	95.0	4.5	0.5	0.0
		D–E	92.6	0.4	7.1	0.0
		E–F	88.1	4.5	7.4	0.0
	$^{4}C_{1}$	C–D	94.8	4.4	0.5	0.2
		D–E	98.9	1.1	0.0	0.0
		E–F	87.9	3.3	8.8	0.0