Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 281–290
Carbohydrate RESEARCH

# Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution

Hugo Verli and Jorge A. Guimarães*

*Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Av. Bento Gonçalves, 9500, CP 15005, Porto Alegre 91500-970, RS, Brazil*

Received 22 July 2003; accepted 16 September 2003

**Abstract**—Molecular dynamics (MD) simulations on heparin–water–sodium systems were carried out in order to establish a simulation protocol able to represent heparin solution conformation under physiological conditions. Atomic charges suitable for heparin oligosaccharides were obtained from ab initio quantum-mechanical computations, at the 6-31G** level. The GROMACS forcefield, the SPC, and SPC/E water models were employed. Also heparin was simulated with IdoA residues in $^1C_4$ or $^2S_0$ conformational states. The results of the performed MD simulations are in agreement with the available experimental data, suggesting that this approach can be applied for the study of heparin interactions with its target proteins and thus play a role in the development of new antithrombotic agents.


*Keywords:* Heparin; Molecular dynamics; Polysaccharide; Structure; GROMACS

## 1. Introduction

Heparin was identified and then isolated in 1916 from a preparation of dog liver, being the first compound used clinically as anticoagulant and antithrombotic agent.[1] It is mainly composed of hexasaccharide units containing iduronic acid (IdoA) 2-sulfate, glucosamine (GlcN) 2,6-disulfate, and non-sulfated glucuronic acid (GlcA),[2] although small variations occur among heparins from different sources.[1] This polysaccharide is capable of forming a ternary complex with antithrombin III and different serine proteases of the blood clotting cascade.[3] Once complexed, heparin potentiate antithrombin inhibitory activity over serine proteases of the coagulation cascade.

The heparin conformational profile presents an unusual mobility due to the presence of IdoA residues.[4] The internal iduronate residues can adopt an equilibrium between a chair $^1C_4$ and skew-boat $^2S_0$ forms[5] without causing the whole polysaccharide chain to bend.[4] This flexibility within IdoA residues is proposed to contribute to the unique heparin binding properties as compared to the lower antithrombotic activity of glycosaminoglycans presenting more rigid uronate residues.[6]

Efforts to elucidate the three-dimensional structure and the dynamic properties of oligosaccharides is a prerequisite for a better understanding of the molecular basis of their recognition by protein targets, which represents the main challenges for structural glycobiology.[7] Experimental methods such as NMR spectroscopy and X-ray crystallography have a long history of application to carbohydrates. However, these methods generally result in a single three-dimensional model for the oligosaccharide, which may fail to adequately describe its dynamic properties. So these time-honored methods can be complemented with molecular modeling techniques such as molecular dynamics (MD) simulations. This methodology allows the simulation of carbohydrates in their natural environment, solvated with counterions or complexed with target proteins.

* Corresponding author. Tel.: +55-51-3316-6062; fax: +55-51-3316-7309; e-mail: guimar@dna.cbiot.ufrgs.br

0008-6215/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.09.026

Unfortunately, sulfated glycosaminoglycans like heparin and similar carbohydrates are thought to be difficult to model because of their highly polar functionality, their flexibility, and conformational/configurational changes, for example, anomeric, exo-anomeric, and gauche effects.[7] Several contributions have been made to set up some relevant parameterizations that would account for these specific features of carbohydrate moieties. However, the only agreement concerning these contributions appears to be restricted to the disaccharide level, and contributed only with individual answers to the problem.[8]

In this work we have studied the conformational profile of a decasaccharide fragment of heparin in aqueous solution using molecular modeling techniques. An MD simulation protocol was established, as well as some simulation conditions that characterize the physiological medium where heparin acts (e.g., salt concentration and explicit water representation). Our results indicate that heparin can be simulated using MD techniques with reasonable confidence and under inexpensive conditions. Indeed the polysaccharide can be simulated with IdoA residues in either the $^1C_4$ or $^2S_0$ conformation. Thus it is possible to represent the conformational profile of heparin in aqueous solution, which opens the perspective of studying the interactions of heparins with their biological targets.

## 2. Experimental

### 2.1. Computational methods

**2.1.1. Nomenclature and software.** The recommendations and symbols of nomenclature as proposed by IUPAC[9] are used. The relative orientation of a pair of contiguous sugar residues (e.g., iduronic acid, glucosamine, and/or glucuronic acid) is described by two torsional angles at the glycosidic linkage, denoted $\phi$ and $\psi$. For a $(1\rightarrow4)$ linkage the definitions become those shown in Eqs. 1 and 2:

$$\phi = \text{O-5–C-1–O-1–C-4}', \quad (1)$$

$$\psi = \text{C-1–O-1–C-4}'\text{–C-5}'. \quad (2)$$

The decasaccharide topologies were generated with the PRODRG program,[10] the ab initio calculations were performed using GAMESS,[11] manipulation of structures was performed with MOLDEN program,[12] and all the MD calculations and analysis were performed using the GROMACS simulation suite and forcefield.[13,14]

**2.1.2. Topology construction.** The heparin fragment under the 1HPN PDB code includes two NMR models of a heparin fragment consisting of six IdoA–GlcN disaccharides.[15] In one model all IdoA residues are in the $^2S_0$ conformation, and in the other model all IdoA residues lie in the $^1C_4$ conformation. These structures were submitted to the PRODRG site,[10] and the initial geometries and crude topologies were retrieved. We reduced the dodecasaccharide to a decasaccharide fragment of heparin composed by five disaccharide units of IdoA–GlcN (Fig. 1). The PRODRG topology was modified to include some refinements such as the reference value for the S–N bond in the sulfonamide groups in residues B, D, F, H, J (Fig. 1), and the atomic charges.

The two conformational states ($^2S_0$ and $^1C_4$) were defined by addition of improper dihedral angles (i.e., a torsion angle in which the four atoms are not bonded in sequence) in the respective topology files. This torsion angle is meant to keep planar groups in a plane or to prevent molecules from flipping over to their mirror images.[14] In our case the improper dihedral angles were intended to constrain the two conformational states of IdoA because of difficulties of the performed simulations in reproducing the transitions between the $^2S_0$ and $^1C_4$ structures. Two topology files were used, one for the $^2S_0$ conformational state of the IdoA residues and one for the $^1C_4$ conformational state of the IdoA residue. The reference angle for the added improper dihedral angles was the only difference between them.

**2.1.3. Atomic charge calculation.** The atomic charges were generated using the sulfated sugars GlcA, IdoA, and GlcN as molecular probes. The sulfate and carboxylate groups were kept in the negatively charged forms. In addition, we used a disaccharide structure without sulfate groups to obtain atomic charges for the glycosidic linkage. These four structures were submitted

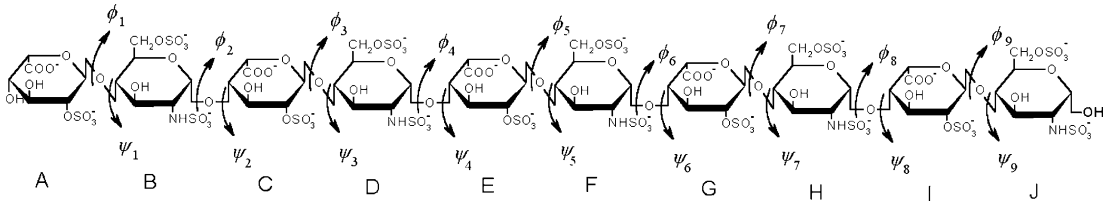



**Figure 1.** The decasaccharide A–J fragment of heparin. This structure comprehends five disaccharide repeating units of the GlcN and IdoA residues. The dihedral angles $\phi$ and $\psi$ are indicated.

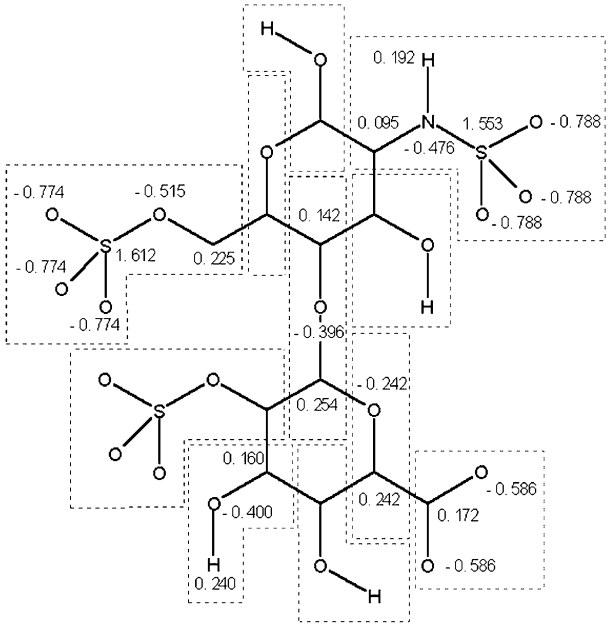



**Figure 2.** Schematic representation of the atomic charges calculated for heparin uronic acids and glucosamine residues. The charges are presented within each charge group of a disaccharide unit.

to full-geometry optimization using ab initio quantum-mechanical computations at the 3-21G level with GAMESS.[11] Hessian matrix analyses were employed to unequivocally characterize the optimized structures as true minima on the potential energy surface. These minimal energy conformations were submitted to single-point ab initio calculations at the 6-31G** level to calculate the Löwdin atomic charges.

In order to make the calculated atomic charges easily applicable to different sugar residues and forcefields, the uronic acids and glucosamine residues were divided in charge groups (Fig. 2). Another advantage of this approach is that each charge group in each sugar residue along the polysaccharide chain is identical, and so are their interactions with the same molecule. The charges of these groups were averaged over both the uronic acid and glucosamine structures. The disaccharide structure was used only to calculate the charges of the glycosidic linkage.

**2.1.4. Molecular dynamics calculations.** The decasaccharide structure obtained from NMR data[15] (PDB entry 1HPN) was the starting point for all simulations performed. This structure was solvated in a rectangular box using periodic boundary conditions and SPC[16] or SPC/E[17] water models. The solvation procedure was done by varying the box size in a manner such that we could control the sodium concentration (~0, ~145, and ~404 mM). The sodium ions were added as counterions following two approaches, one by the *Genion* Program of the GROMACS simulation suite, and other by changing manually the water molecule closest to each negatively charged group. The total system size comprised 18,561 atoms for runs 1, 2, 4, 5, 6, and 7, and 6963 atoms for run 3, respectively (the parameters of runs 1–6 are depicted in Table 1). The MD protocol used has already been described.[18] Briefly, the Lincs and Settle methods[19–21] were applied to constrain covalent bond lengths, allowing an integration time step of 2 fs after an initial energy minimization using the Steepest Descents algorithm. Electrostatic interactions were calculated with Particle–Mesh Ewald method.[22] Temperature and pressure were kept constant separately by coupling the carbohydrate, ions, and solvent to external temperature and pressure baths with coupling constants of $\tau = 0.1$ ps and $\tau = 0.5$ ps,[19] respectively. The dielectric constant was treated as $\varepsilon = 1$, and the reference temperature was adjusted to 310 K. The simulation lengths were 3 ns for each of the conditions studied. The system was heated slowly from 50 to 310 K, in steps of 5 ps, each one increasing the reference temperature by 50 K. The

**Table 1.** Average dihedral angles and NMR reference values for $^2S_0$ MD simulations varying the water model (SPC and SPC/E) and the $Na^+$ insertion methodology

| Dihedral angle[a] | Average[a] dihedral (°) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $^2S_0$ NMR[b] | Steepest Descents | Run # | | | | | |
| | | | SPC | | | SPC/E | | |
| | | | 1[c,e] | 2[d,e] | 3[d,f] | 4[c,e] | 5[d,e] | 6[g] |
| IdoA → GlcN $\phi$ | −55.4 | −60.0 | −77.9 ± 10.8 | −77.4 ± 11.2 | −77.9 ± 10.6 | −77.5 ± 10.6 | −75.0 ± 10.7 | −82.2 ± 13.1 |
| IdoA → GlcN $\psi$ | −107.4 | −102.8 | −119.3 ± 11.3 | −118.8 ± 11.5 | −119.2 ± 10.4 | −118.8 ± 10.4 | −116.6 ± 10.6 | −121.7 ± 12.2 |
| GlcN → IdoA $\phi$ | 108.6 | 100.3 | 92.8 ± 15.3 | 94.1 ± 12.7 | 95.8 ± 15.7 | 96.6 ± 13.9 | 96.6 ± 14.1 | 88.4 ± 13.4 |
| GlcN → IdoA $\psi$ | −157.5 | −153.2 | −130.4 ± 16.3 | −127.6 ± 13.8 | −125.9 ± 16.7 | −125.4 ± 15.6 | −125.0 ± 15.1 | −133.8 ± 15.0 |

[a] See Experimental section for details.
[b] Data obtained from PDB code 1HPN, Ref. 15.
[c] $Na^+$ inserted using *Genion* module of GROMACS.
[d] $Na^+$ inserted manually at each negative charged group of heparin (see Experimental section for details).
[e] Box size adjusted to give $[Na^+] = \sim 145$ mM (see Experimental section for details).
[f] Box size adjusted to give $[Na^+] = \sim 404$ mM (see Experimental section for details).
[g] No $Na^+$ inserted.

**Table 2.** Average MD dihedral angles and NMR reference values for $^1C_4$ and $^2S_0$ conformations

| Dihedral angle[a] | Average[a] dihedral (°) | | | | | |
|---|---|---|---|---|---|---|
| | $^2S_0$ Steepest Descents | $^2S_0$ NMR[b] | Run 5 ($^2S_0$) | $^1C_4$ Steepest Descents | $^1C_4$ NMR[b] | Run 7 ($^1C_4$) |
| IdoA → GlcN $\phi$ | −60.0 | −55.4 | −75.0 ± 10.7 | −78.8 | −77.1 | −124.0 ± 18.3 |
| IdoA → GlcN $\psi$ | −102.8 | −107.4 | −116.6 ± 10.6 | −116.1 | −110.1 | −144.5 ± 11.4 |
| GlcN → IdoA $\phi$ | 100.3 | 108.6 | 96.6 ± 14.1 | 75.8 | 78.7 | 89.5 ± 11.9 |
| GlcN → IdoA $\psi$ | −153.2 | −157.5 | −125.0 ± 15.1 | −141.8 | −149.9 | −141.3 ± 11.0 |

[a]See Experimental section for details.
[b]Data obtained from PDB code 1HPN, Ref. 15.

thermalization, together with the first nanosecond, was considered as equilibration, and all analyses were performed over the remaining part of the trajectory. The final structure obtained in run 5 was further minimized with Steepest Descents using the $^1C_4$ improper dihedral angles for the IdoA residues in the topology file, giving the $^1C_4$-minimized structure presented in Table 2. This structure was used as input for run 7. Data concerning the decasaccharide structure, for example, glycosidic linkage and intramolecular hydrogen bonds, were obtained by averaging over four or five values for the same linkage located in five different positions in the decasaccharide. The average of these property values indicates not only deviations within a single linkage but also contributions from different positions in the polysaccharide, as reported before.[23] It should be noted that when a hydrogen bond occurs between a donor and multiple acceptors, for example, carboxylate and sulfate groups, only the lowest distance of two or three possibilities was retrieved at each frame of MD. As a result we have an average distance representative of multiple interactions. A reference value of 3.5 Å between heavy atoms was considered for a hydrogen bond[24] and a cutoff angle of 60° between donor–hydrogen–acceptor.[14]

## 3. Results and discussion

### 3.1. Atomic charges

Despite the lack of carbohydrate parameterization for polysulfated sugars as heparin, we decided to generate charges applicable to heparin uronic acids (glucuronic and iduronic acids) and glucosamine residues. Previous workers have attempted to develop parameters for sulfated monosaccharide salts.[25] Considering that it was based on Allinger's MM2 forcefield and that no parameter for the glycosidic linkage was presented, we conclude that there is a need to develop atomic charges suitable to be used in heparin MD simulation with the GROMACS forcefield. The atomic charges at the Hartree–Fock HF/6-31G** level were chosen for this purpose.

Instead of the usual Mulliken population analysis, we adopted the Löwdin atomic charges due to the tendency of Mulliken analysis to put all of the charge on the oxygen atom.[26] Electrostatic potential (ESP) based charges were reported as more accurate than Mulliken charges even for high-level ab initio theory.[27] However, due to differences in electrostatic properties along the heparin structure, the use of ESP-based charges would give different atomic charges for the atoms located in the middle from those located at the extremes of the structure. Figure 2 shows the charges obtained for the disaccharide units of the heparin decasaccharide fragment.

### 3.2. Simulation conditions

Some simulation conditions capable of modifying the heparin dynamics and conformation in the MD trajectory, that is, the sodium insertion methodology, the sodium concentration, and the chosen water model were initially evaluated. These simulations were carried out using only the $^2S_0$ conformer of iduronic acid in heparin (runs 1–6), and the best conditions identified were further applied to the $^1C_4$ conformer (run 7, see Experimental section for details and Table 1 for definitions of runs).

To neutralize the simulated system (water plus the decasaccharide), 20 $Na^+$ ions were added. We use two procedures to add sodium ions to the decasaccharide solution: the *Genion* Program of the GROMACS simulation suite (runs 1 and 4) and the manual insertion of sodium at each negatively charged group (runs 2, 3, 5, and 7—e.g., carboxylate and sulfate groups). The *Genion* Program replaces solvent molecules by monoatomic ions based on the most favorable electrostatic site. However, we noticed that for the simulated system the sodium ions were mainly placed at the corners of the box, therefore distant from the decasaccharide molecule. Considering that this distribution of sodium ions could create some instability in the simulation, we decided to locate the ions close to each charged group, as already reported.[23] The decasaccharide was also simulated under different sodium concentrations (~0, ~145, and ~404 mM) and two water models (SPC and SPC/E).

The analysis of the MD simulation of the decasaccharide was based on the glycosidic linkage conformational profile over the MD trajectory. The $\phi_n$ and $\psi_n$

angles were averaged during the last 2 ns of MD simulation and over the decasaccharide chain. We also used the values of the $\phi$ and $\psi$ angles obtained by the Steepest Descents minimization of the system comprising water, ions, and the decasaccharide itself in order to discern the quality of the generated topology. The average dihedral angles from the MD simulations of runs 1–6, from NMR data and from the Steepest Descents minimization are presented in Table 1. It should be noted that no end effects were observed in the polysaccharide chain, i.e., greater conformational flexibility of the terminal glycosidic linkages is indicated. Only small variations were noted, which were always smaller than the standard deviation of the angles. So there is no loss of information about the chain position on the linkage fluctuations as the data are presented here.

Comparing the average angles shown in Table 1, we observe that most angle variations are smaller than the standard deviation ($\sim$13°). Moreover run 5 conditions present the closest values to the NMR experimental data. These conditions (a manually replacement of water by sodium ions and the SPC/E water model, with IdoA residues in the $^2S_0$ conformation) were thus used in a decasaccharide structure with all IdoA residues in a $^1C_4$ conformation (Table 2, run 7).

A problem frequently found in molecular modeling studies is the choice of the atomic charges for the system. The reliability of the atomic charges becomes more difficult to achieve when the molecules presents hypervalent groups, for example, sulfonamide and sulfate. The atomic charges for these groups are predicted by semi-empirical methods to be very high as a consequence of the non-association of d orbitals with the sulfur atom.[27] This problem can be increased in polysulfated compounds such as heparin, creating an abnormally strong electrostatic field around the carbohydrate molecule. Actually initial simulation conditions swere tested using semi-empirical atomic charges, but as great distortions of the decasaccharide structure were observed (data not shown), we avoided this methodology.

In order to test the reliability of the calculated heparin residue charges (Fig. 2), we compared them to the atomic charges described previously by Ferro and co-workers.[25] From this comparison a correlation between the two sets of charges can be made, albeit with some important differences. A two-fold variation was observed in carbon, hydrogen, and nitrogen atoms of sulfonamide group. This discrepancy seems to be due to the 6-31G** Löwdin atomic charges used in the present work (compared to the 6-31+G** level with the post-SCF Møller–Plesset electronic correlation treatment at the second order) and the use of molecular probes to calculate the charges. The work of Ferro used a methyl-*O*-sulfate, while we used the entire sulfated monosaccharide (except for calculation of the atomic charges of the glycosidic linkage, when a disaccharide probe was used—see Experimental section for details). Another relevant difference observed in comparison with the work of Ferro and co-workers is related to the oxygen charges of the sulfate group. Distinctly from those authors, we made the atomic charges of these oxygen atoms uniform. We choose this based on the difficulty in differentiate the hydrogen bond formed by each sulfate oxygen atom in solution as well as the effect of the entire polysaccharide chain on these charges. So we attempted not to create abnormal interactions using different charges on each oxygen atom of the sulfate groups that could modify the 3D structure of heparin as well as modify the interaction profile of heparin with target proteins.

The charges thus obtained were also compared to those described in the GROMACS forcefield for galactose and glucose. An important difference was noted in the hydroxyl group where the oxygen atoms have a charge of −0.548 e and the hydrogen atoms have a charge of 0.398 e (contrasting with the charges of −0.400 e and 0.240 e, respectively, described in the present work). We thus also tested these charges in MD simulations. However, the reproduction of the heparin NMR structure was not so good (data not shown) as that obtained with the charges presented as in Figure 2, which reinforces the fitness of the approach presented here.

Together with the atomic charges from heparin, the amount and initial positions of sodium ions are important properties to the heparin tri-dimensional structure and dynamics due to their interaction with the charged groups of heparin. The lack as well as the location of sodium in unfavorable regions could give rise to strong electrostatic interactions not found under the natural heparin solution conditions. In fact it has been reported that the interactions between the polysaccharide chain and water molecules determine carbohydrate structure and functionality,[28] being furthermore influenced by the water model used. In this context the simulation of the decasaccharide at different sodium concentrations may give insights into the influence of this ion in heparin conformation.

Concerning the sodium placement approach, differences were not observed between the MDs of decasaccharides in runs 1 and 2 (Table 1). These two simulation conditions used the SPC water model, but the picture is very similar considering runs 4 and 5 that used the SPC/E water model. Also the comparison between runs 2 and 5 show that the system simulated with the SPC/E water model tends to better represent the heparin conformational profile when compared with the system simulated with the SPC water model, with the exception of the GlcN → IdoA $\psi$ angle (see further). Globally, run 5 presented the best conditions found for MD simulation of heparin.

Together with the localization of ions prior to MD, the performed simulations also give insights about the

influence of sodium concentration on the theoretical studies of heparin. The comparison between runs 2 and 3, with a four-fold variation in sodium concentration (Table 1) shows no significant modification in the dynamics profile of the glycosidic linkage, suggesting that is possible to simulate heparin with a lower amount of water and a greatly reduced computational cost. In the case of runs 2 and 3, there was a three-fold reduction in water amount and an increase in speed of calculation by almost a factor of four. Considering the great amount of water necessary to solvate the sodium salt of heparin under physiological conditions (~145 mM), the use of hypertonic solutions can be a good alternative to reduce the computational cost of heparin simulation without great loss in precision of the results. However, the above results could give the impression that the presence of sodium ions do not make any difference in the conformational profile of the heparin fragment simulated. As can be seen (Table 1, run 6), when the simulation was made in absence of sodium ions but with the same amount of water used for the runs at physiological concentration of the ion (runs 1, 2, 4, and 5), a considerable distortion of the glycosidic bond geometry was observed. These data thus indicate the relevance of positive ions in the MD description of the heparin conformational profile.

Considering that the overall conformation of oligosaccharides can be essentially determined by the torsion angles of the glycosidic linkage, we also analyzed the time-dependent fluctuations of such angles. The values of dihedral angles $\phi_n$ and $\psi_n$ are presented in Figure 3 over the 3-ns trajectory of run 5. All the angles show a stable conformational behavior around the average angles presented in Table 1, indicating that the system is quite equilibrated. These data can also suggest a good conformational sampling of the performed simulations. However, a considerable difference in the representation of heparin dynamics in solution was observed over the different types of heparin glycosidic linkages (Fig. 3). The IdoA → GlcN angles are closer to the NMR reference values and also present less fluctuation over the MD when compared to the GlcN → IdoA angles. This profile appears to be due to intramolecular hydrogen bonds between the sulfonamide group of the GlcN residue and the hydroxyl group of the IdoA residue. The hydrogen bond between these two groups (interaction C of Table 3) presents a deviation of about 1 Å from the NMR data, giving a weaker intramolecular interaction and thus increasing the flexibility of the dihedral angle.

It is also possible that the GlcN → IdoA angle be more flexible than the IdoA → GlcN angle as an intrinsic property of heparin. In agreement with this observation and reinforcing the MD representation of heparin conformation in solution is the fact that the experimental value of the IdoA → GlcN $\psi$ angle is also represented during the MD trajectory. This dynamic look in the conformational behavior of heparin observed here has been lost in the one-structure NMR heparin under PDB-ID 1HPN and in forcefield adjustments of the polysaccharide structure to fit the NOESY data, thus indicating

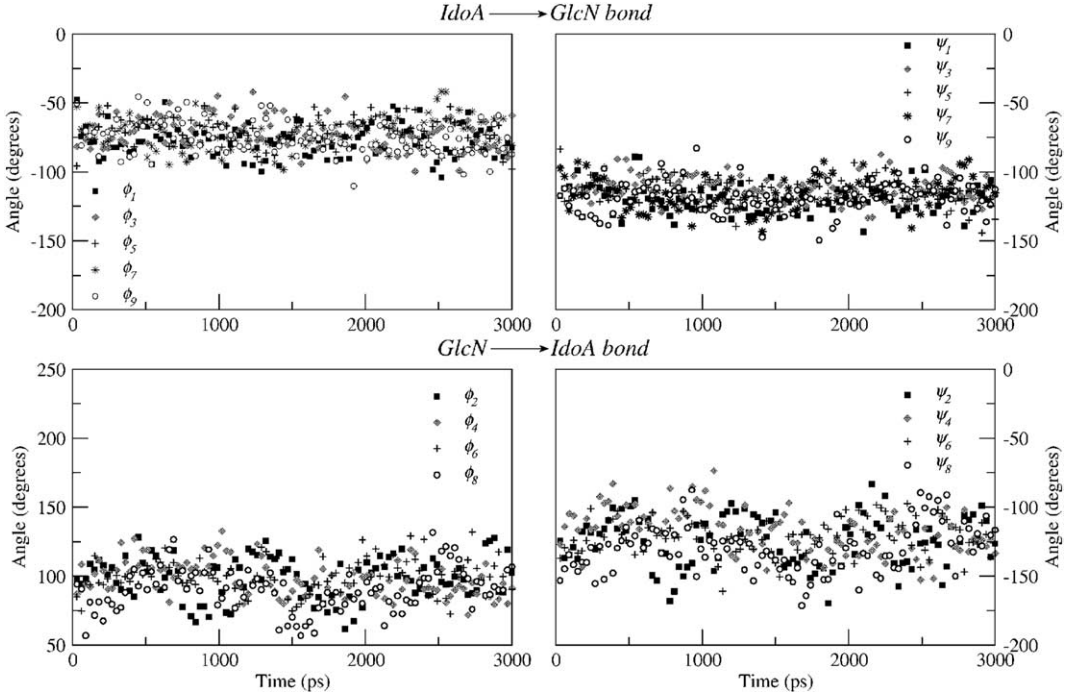
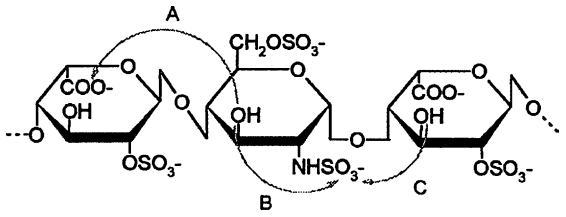



**Figure 3.** Time dependence of the glycosidic linkage torsion angles $\phi_n$ and $\psi_n$ in the decasaccharide over the 3 ns of run 5. The curves were smoothed in 100 ps windows for the sake of clarity.

**Table 3.** Average MD intramolecular hydrogen bond distance and NMR reference values for $^1C_4$ and $^2S_0$ conformations




| Hydrogen bond[a] | Average[a] distance (Å) | | | |
|---|---|---|---|---|
| | $^2S_0$ NMR[b] | Run 5 ($^2S_0$) | $^1C_4$ NMR[b] | Run 7 ($^1C_4$) |
| A | 3.4 | 3.8 | 3.4 | 5.0 |
| B | 3.7 | 3.5 | 3.0 | 3.6 |
| C | 2.8 | 3.7 | 6.1 | 5.0 |

[a]See Experimental section for details.
[b]Data obtained from PDB code 1HPN, Ref. 15.

MD as an adequate and complementary procedure for one-structure NMR studies.

The application of the MD conditions of run 5 (IdoA residues in the $^2S_0$ conformation) to simulate a decasaccharide with IdoA residues in a $^1C_4$ form gives rise to intriguing aspects of the representation of heparin conformation by MD. While in the $^2S_0$ form of IdoA residues the error in glycosidic linkage compared to NMR values ranged from ~10° (IdoA → GlcN $\psi$ and GlcN → IdoA $\phi$) to ~20° (IdoA → GlcN $\phi$) and ~30° (GlcN → IdoA $\psi$) in the $^1C_4$ form of IdoA residues the error ranged from ~5° (GlcN → IdoA $\psi$) to ~10° (GlcN → IdoA $\phi$), ~30° (IdoA → GlcN $\psi$), and ~50° (IdoA → GlcN $\phi$), indicating that the skew-boat conformation is better described by MD techniques than the chair conformation in the heparin IdoA residues. Also the dihedral angle that is better described in the $^1C_4$ conformation of the IdoA residues is the worst in the $^2S_0$ conformation (GlcN → IdoA $\psi$).

The analysis of the intramolecular hydrogen bonds of the decasaccharide suggest that an interaction between a hydroxyl group of the IdoA residue and the sulfonamide group of the GlcN (interaction C, Table 3) residue can be essential for the geometry of GlcN → IdoA $\psi$ angle. In run 5 (IdoA residues in a $^2S_0$ form) a deviation of 0.9 Å in this interaction from the NMR structure implies the existence of a weaker interaction between the two residues, thus producing the distortion of the angle. In run 7, however, this interaction is not important for the heparin conformation since the distance C is already too far from a hydrogen bond reference value in the NMR structure with the IdoA residues in a $^1C_4$ form. This explains why the GlcN → IdoA $\psi$ angle is better described when the IdoA residues lie in a chair conformation compared to the skew-boat conformation. Also these results indicate that the change in IdoA conformation alters significantly the hydrogen-bond network within heparin and can have important implications in the interaction of the polysaccharide with its target proteins.

As shown in Table 2, there is a greater deviation of angle IdoA → GlcN $\psi$ (and maybe of angle IdoA → GlcN $\phi$) in run 7 as compared to run 5. This find is probably due to a hydrogen bond between the hydroxyl group of the GlcN residue and the carboxylate group of the IdoA residue (interaction A, Table 3) observed in the NMR structure but not reproduced in the MD simulations for a decasaccharide with IdoA residues in a $^1C_4$ conformation. In fact the MD was capable of representing this hydrogen bond only with IdoA residues in the $^2S_0$ form (3.8 Å in the $^2S_0$ form vs 5.0 Å in the $^1C_4$ form), justifying the deviations from the NMR data. Also these changes in heparin structures are not a consequence of simple polysaccharide denaturation since the length of the decasaccharide is almost the same in both runs 5 and 7 (39.9 and 38.8 Å, respectively, compared to the experimental value of 38.4 Å). These results indicate clearly that the MD simulations performed better with IdoA residues in the $^2S_0$ conformation than in the $^1C_4$ conformation. Even with these limitations and with the differences obtained between the simulated heparin systems and the NMR reference value, the overall geometry of the molecule was maintained in MD simulations as shown in Figure 4.

The preference of IdoA residues to stay in a skew-boat form to the detriment of a chair form has been described in solution[29] and when interacting with antithrombin.[30] This is in agreement with the preference of sugars to keep their hydroxyl groups in equatorial orientation due to the interference with the surrounding solvent: equatorial groups would be more stable (less interference) than axial groups (more interference).[31] So the mutual conformational influence between sugars and solvent reinforce the relevance of the solvent model to

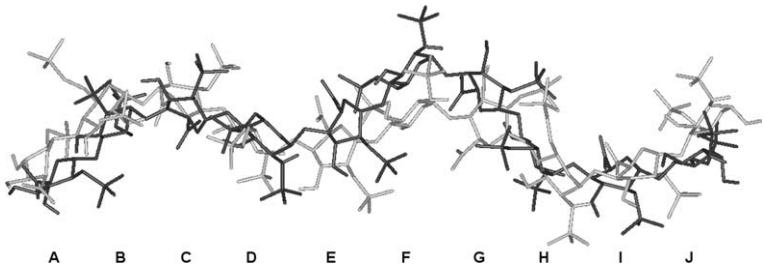



**Figure 4.** All atom superimposition of minimized decasaccharide (light gray), and run 5 (dark gray).

the overall dynamics of the polysaccharide, ring conformation, and glycosidic linkage. In agreement with this assumption is the low value of the stretching constant associated with a glycosidic linkage in available forcefields (~3.8 kJ), in a way that the interactions with solvent will give enough energy to modify the glycosidic linkage geometries over the polysaccharide chain. Also our simulations indicate that the interaction energy between decasaccharide and solvent is about 550 kJ more intense when IdoA residues lie in the $^2S_0$ conformation (−6601.7 kJ/mol) than in the $^1C_4$ conformation (−6056.9 kJ/mol).

In order to verify the reliability of the performed MD simulation we have compared the data obtained with other works reported by MD, NMR, and crystallographic studies of heparin.[6,15,23,32] These comparisons are represented in Table 4.

As we can see in Table 4, there is a great variability in geometry through the previously determined heparin structures. For example, heparin tetrasaccharide in solution[6] has a variation in its glycosidic linkages relative to crystal structure[32] of ~20° (GlcN → IdoA $\phi$) to ~50° (GlcN → IdoA $\psi$). This variation could be due to both the packing effects in the crystal or to conformational modification of heparin induced by FGFb. However the heparin hexasaccharide has a variation of ~20° (IdoA → GlcN $\psi$ and GlcN → IdoA $\phi$) to 35° (IdoA → GlcN $\phi$) from MD data relative to the NMR structure.[23] In this case we have no interaction with other proteins capable of conformational induction. One possibility is the limitation of the forcefield used in the MD simulations. However, the comparison between MD results of the hexasaccharide and the decasaccharide (run 5) shows a remarkable similarity around the IdoA → GlcN dihedral angle (Table 4). Considering that the hexasaccharide simulation used the AMBER forcefield with Homans' additions for saccharides,[33] and in the decasaccharide simulation we used the GROMACS forcefield,[14] a GROMOS87-based forcefield, which are two basis sets with great differences,[8] we can reduce the relevance of the forcefield in the differences between the NMR and MD data.[23] Another possibility, also applicable to the differences between the $^2S_0$ NMR structure and run 5, lies in the refinement of the heparin structure. In this process there is a loss of dynamic information about the glycosidic linkage (e.g., the Mulloy heparin structure shows only two conformations, one with IdoA in the $^2S_0$ conformation and the other in the $^1C_4$ conformation) that appears with more detail in molecular dynamics calculations.[34]

The differences in dihedral angle average values for the glycosidic linkages presented in Table 4 for both theoretical and experimental data can also be analyzed in the polysaccharide hydration context. The carbohydrate's polar functionality creates a hydrogen-bond network that affects the surrounding water and, in return, the water affects the structure of the sugar. For oligosaccharides containing glycosidic linkages at the

**Table 4.** Average MD dihedral angles and NMR, crystallographic, and solution reference values for $^2S_0$ conformations

| Dihedral angle | Average dihedral (°) | | | | | |
|---|---|---|---|---|---|---|
| | $^2S_0$ NMR[a] | Run 5[b] | Tetrasaccharide | | Hexasaccharide[e] | |
| | | | Solution[c] | Crystal[d] | NMR | MD |
| IdoA → GlcN $\phi$ | −55.4 | −75.0 | −72 | −78.8 | −40 | −75 |
| IdoA → GlcN $\psi$ | −107.4 | −116.5 | −102 | −107.1 | −99 | −121 |
| GlcN → IdoA $\phi$ | 108.6 | 96.6 | 74 | 94.5 | 73 | 51 |
| GlcN → IdoA $\psi$ | −157.5 | −125.0 | −75 | −129.8 | −82 | −83 |

[a]PDB code 1HPN, Ref. 15.
[b]The present work (see Table 2).
[c]Ref. 6.
[d]PDB code 1BFB, Ref. 32.
[e]Ref. 23.

6-position, the correct experimental rotamer distribution about the $\omega$-angle is obtained only with the inclusion of explicit water molecules in the MD simulation.[34] Once heparin has highly charged characteristics, its interaction with the surrounding water should be stronger than the interactions of other non-sulfated carbohydrates. Considering the effectiveness of this strong interaction, it is reasonable to hypothesize that the mutual influence between the heparin and solvent will also be strengthened. It remains however to be determined in which degree the water models that are so-far available for MD are able and appropriate to better describe this process. It should be noted that while the IdoA residues were fixed in the $^1C_4$ and $^2S_0$ forms, we have not used any kind of positional restraints in the rest of the heparin structure, neither in minimization nor in the MD simulation. This is an advantage compared to previous studies in which the use of restraints was necessary to avoid distortion of polysaccharide structure during equilibration.[23] Also the absence of consensus in the heparin structure and dynamics highlights the MD simulations presented in the present work. The variations between run 5 and the NMR structure of heparin containing the $^2S_0$ conformation of IdoA residue[15] are equivalent to the variations between the hexasaccharide structure obtained by NMR and MD simulations[23] (i.e., 35° in the IdoA $\rightarrow$ GlcN $\phi$ angle) and even a half value of the variations between tetrasaccharide crystal[30] and its solution structure[6] (i.e., $\sim$60° in the GlcN $\rightarrow$ IdoA $\psi$ angle). So we believe that the approach presented can be a useful tool in heparin study with a accuracy relative to experimental data and at least comparable with previously related proceedings with a lower computational cost. Moreover, it uses a distinct forcefield parameterization that can contribute with new answers to carbohydrate simulation and study.

## 4. Conclusions

The lack of structural information concerning the intrinsic conformational flexibility of heparin as well as about the interaction of heparin with its target proteins is partially due to experimental difficulties with this complex polysaccharide. These difficulties were also found with theoretical approaches like MD simulations. Many efforts have been made in order to overcome these obstacles, and the major MD simulation of heparin before the present study was made with a hexasaccharide during 2 ns.[35]

Here we present a conformational study of a heparin decasaccharide using MD simulation. This simulation was stable during the performed 3 ns with $\sim$6000 water molecules, characterizing a system with a physiological sodium concentration. Atomic charges suitable for MD calculations of sulfated sugars were parameterized using Löwdin population analysis and compared with previously related atomic charge schemes for sulfated and non-sulfated carbohydrates. The performed simulations were carried out with heparin presenting IdoA in either the $^2S_0$ or $^1C_4$ conformations. The conformational behavior of the glycosidic linkage is in agreement with the available NMR experimental data, as well as previous crystallographic and MD data. We believe that this protocol can be useful in future studies, overcoming the difficulties in determining the interaction sites of heparin and other glycosaminoglycans of different sizes and sugar compositions with their target proteins. The rationalization and molecular modeling of heparin interactions with its binding proteins can be an important tool for the development of new antithrombotic agents and also for monitoring the distinct activity profiles of heparins of different origins, including their low-molecular-weight components and other derivatives.

## Acknowledgements

We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), MCT, the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), MEC, Brasília, DF, Brazil, and the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), RS, Brazil for their financial support. We are indebted to Prof. Paulo A. Netz (ULBRA), Prof. Ernesto Caffarena (Fiocruz), Prof. Laurent Dardenne (LNCC), Carlos Maurício Sant'Anna (UFRRJ), and Prof. Ivarne Tersariol (UMC) for scientific discussion and contributions to this paper.

## References

1. Nader, H. B.; Pinhal, M. A. S.; Baú, E. C.; Castro, R. A. B.; Medeiros, G. F.; Chavante, S. F.; Leite, E. L.; Trindade, E. S.; Shinjo, S. K.; Rocha, H. A. O.; Tersariol, I. L. S.; Mendes, A.; Dietrich, C. P. *Braz. J. Med. Biol. Res.* **2001**, *34*, 699–709.
2. Silva, M. E.; Dietrich, C. P. *J. Biol. Chem.* **1975**, *250*, 6841–6846.
3. Danielsson, A.; Raub, E.; Lindahl, U.; Björk, I. *J. Biol. Chem.* **1986**, *261*, 15467–15473.
4. Mulloy, B.; Forster, M. J. *Glycobiology* **2000**, *10*, 1147–1156.
5. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Casu, B.; Torri, G.; Bossennec, V.; Perly, B.; Sinay, P.; Petitou, M.; Choay, J. *Carbohydr. Res.* **1990**, *195*, 157–167.
6. Mikhailov, D.; Mayo, K. H.; Vlahov, I. R.; Toida, T.; Pervin, A.; Linhardt, R. J. *Biochem. J.* **1996**, *318*, 93–102.
7. Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567–4588.
8. Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141–155.

9. IUPAC-IUB Commission on Biochemical Nomenclature. *J. Mol. Biol.* **1970**, *52*, 1–17.
10. (a) van Aalten, D. M. F.; Bywater, B.; Findlay, J. B. C.; Hendlich, M.; Hooft, R. W. W.; Vriend, G. *J. Comput. Aided Mol. Des.* **1996**, *10*, 255–262; (b) http://davapc1.bioch.dundee.ac.uk/programs/prodrg/prodrg.html.
11. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
12. MOLDEN. Schaftenaar, G. CAOS/CAMM Center, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, 1997.
13. Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
14. van der Spoel, D.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Hess, B.; Feenstra, K. A.; Lindahl, E.; van Drunen, R.; Berendsen, H. J. C. GROMACS User Manual Version 3.0, Nijenborgh 4, 9747 AG Groningen, The Netherlands, 2001.
15. Mulloy, B.; Forster, M. J.; Jones, C.; Davies, D. B. *Biochem. J.* **1993**, *293*, 849–858.
16. Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981; pp 331–342.
17. Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
18. de Groot, B. L.; Grubmüller, H. *Science* **2001**, *294*, 2353–2357.
19. Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
20. Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
21. Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
22. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092.
23. Mikhailov, D.; Linhardt, R. J.; Mayo, K. H. *Biochem. J.* **1997**, *328*, 51–61.
24. Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 55–57.
25. Ferro, D. R.; Pumilia, P.; Ragazzi, M. *J. Comput. Chem.* **1997**, *18*, 351–367.
26. Leach, A. R. *Molecular Modeling Principles and Applications*; Longman: Singapore, 1996. pp 115–116.
27. Clare, B. W.; Supuran, C. T. *J. Mol. Struct.* **1998**, *428*, 109–121.
28. Engelsen, S. B.; Monteiro, C.; de Penhoat, C. H.; Pérez, S. *Biophys. Chem.* **2001**, *93*, 103–127.
29. Cros, S.; Petitou, M.; Sizun, P.; Pérez, S.; Imberty, A. *Bioorg. Med. Chem. Lett.* **1997**, *5*, 1301–1309.
30. Jin, L.; Abrahams, J. P.; Skinner, R.; Petitou, M.; Pike, R. N. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 14683–14688.
31. Uedaira, H.; Okouchi, S.; Tsuda, S.; Uedaira, H. *Bull. Chem. Soc. Jpn.* **2001**, *74*, 1857–1861.
32. Fahem, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C. *Science* **1996**, *276*, 1116–1120.
33. Homans, S. W. *Biochemistry* **1990**, *29*, 9110–9118.
34. Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541–10545.
35. Angulo, J.; Nieto, P. M.; Martin-Lomas, M. *Chem. Commun. (Cambridge)* **2003**, *13*, 1512–1513.