Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 342 (2007) 779–783
Carbohydrate RESEARCH

Rapid Communication

Relationship between structure and three-bond proton–proton coupling constants in glycosaminoglycans

M. Hricovíni[a,*] and F. Bízik[b]

[a]*Institute of Chemistry, Slovak Academy of Sciences, 845 38 Bratislava, Slovakia*
[b]*Institute of Virology, Slovak Academy of Sciences, 842 45 Bratislava, Slovakia*

Received 2 November 2006; received in revised form 4 January 2007; accepted 5 January 2007
Available online 12 January 2007

Abstract—Theoretical calculations using the DFT theory at the B3LYP/6-311++G** level were used to determine the molecular geometry of various glycosaminoglycan (GAG) molecules. Three-bond proton–proton spin–spin coupling constants ($^3J_{H–C–C–H}$) were then computed and compared with the published experimental data of selected mono- and disaccharides. The computed $^3J_{H–C–C–H}$ values showed a strong dependence on the molecular geometry and varied up to 12 Hz. This dependence was expressed in a simple analytical form relating $^3J_{H–C–C–H}$ and torsion angles. The population of conformers in heparin and other biologically active GAGs has also been estimated using the computed coupling constants.

Keywords: Glycosaminoglycans; DFT calculations; Three-bond coupling constants; Conformation

Glycosaminoglycans (GAGs) are linear polysaccharides constituted by repeating units of hexosamine and uronic (glucuronic or iduronic) acid. GAGs are involved in a number of biological functions, such as cell growth and differentiation, blood coagulation, angiogenesis, and viral invasion.[1] The knowledge of their 3D structure and intermolecular complexes with proteins has been therefore the subject of numerous biochemical and biophysical studies.[2,3] The 3D structures of GAG monosaccharide units in aqueous solutions are primarily determined from high-resolution NMR data. The magnitudes of three-bond proton–proton coupling constants ($^3J_{H–C–C–H}$), together with molecular modeling data, indicate that various conformations ($^4C_1$, $^1C_4$, and $^2S_0$) of iduronic acid (IdoA) residues contribute to conformational equilibria in GAGs.[4] The populations of these forms depend upon substitution of the IdoA residues and the structure of neighboring units.[5] Such conformational flexibility of the IdoA residues in GAGs is a rather specific phenomenon in carbohydrate chemistry. It has its origin in the structural arrangements of these molecules, mainly the presence of bulky and negatively charged sulfate and carboxylate groups. Both types of groups have effects not only on the conformational properties of IdoA residues but also on NMR parameters.[6] As $^3J_{H–C–C–H}$ values are influenced by atomic charges (among other parameters), coupling constants should be interpreted with appropriately parametrized Karplus-type relationships rather than those used for neutral molecules. However, conformer abundances in GAGs have been determined to date exclusively through the Haasnoot–Altona parametrization of $^3J_{H–C–C–H}$.[7] The applications of the latter type of parametrization did not lead to satisfactory agreements between computed averaged (based on the geometry of conformers) and experimental $^3J_{H–C–C–H}$ values in several biologically active GAGs. In the present study, we examine the properties of $^3J_{H–C–C–H}$ in GAG molecules and compare their magnitudes with the experimental data. Based on the computed couplings, a simple analytical form is proposed relating the dependence of $^3J_{H–C–C–H}$ and torsion angles in the H–C–C–H array of atoms.

* Corresponding author. Tel.: +421 2 59410323; fax: +421 2 59410222; e-mail: hricovini@savba.sk

0008-6215/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2007.01.003

DFT-computed $^{3}J_{H–C–C–H}$ coupling constants in saccharides **1**–**4** (Fig. 1) are given in Table 1. $^{3}J_{H–C–C–H}$ values are presented for three different conformers, namely, $^{4}C_1$, $^{1}C_4$, and $^{2}S_0$ for monosaccharide methyl 2-*O*-sulfo-α-L-iduronate monosodium salt (IdoA2S-OMe). Two conformers ($^{1}C_4$ and $^{2}S_0$) of the IdoA2S residue and the $^{4}C_1$ form in the N,6-sulfated α-D-glucosamine residue (GlcN,6S) are analyzed in the disaccharide (GlcN,6S-IdoA2S-OMe) corresponding to the regular sequence of polysaccharide heparin. The E_0 form was analyzed in conformationally rigid 1,6-anhydro-2-sulfamino-3-*O*-sulfo-D-glucopyranose. The geometry of these compounds was obtained by energy minimization using the B3LYP functional and the 6-311++G** basis set (compounds **1** and **2**) or B3LYP/6-31+G* method (compounds **3** and **4**). The values of the torsion angles of the H–C–C–H array of atoms in these compounds covered a large interval and the corresponding $^{3}J_{H–C–C–H}$ ranged from 0.54 Hz up to 12.20 Hz. The smallest value (0.54 Hz) was obtained for 108.2° in compound **4**, the largest one (12.20 Hz) for 179.4° in compound **3**. Altogether, 35 values of $^{3}J_{H–C–C–H}$ were obtained for compounds **1**–**4** (Table 1). The computed data enabled the determination of the relationship between $^{3}J_{H–C–C–H}$ and torsion angles. The dependence, together with the computed coupling constants, is shown in Figure 2. This relationship can be expressed in the following form:

$$^{3}J_{H–C–C–H} = 9.6\cos^2\phi - 0.6\cos\phi + 0.2$$

where ϕ is the torsion angle of the H–C–C–H atom array. The overall shape of the curve is typical for three-bond coupling constants depending upon the torsion angles where the lowest $^{3}J_{H–C–C–H}$ values correspond to angles close to 90° and 270° and the largest one to *anti-periplanar* conformation (Fig. 2). However, influences of both sulfate and carboxylate groups result in such magnitudes of coupling constants that do not correspond to torsion angle dependence exclusively (Table 2, Fig. 2). For example, the magnitude of $^{3}J_{H3–C3–C4–H4}$ is larger (3.12 Hz) than the $^{3}J_{H4–C4–C5–H5}$ value (2.30 Hz) in **1** though the H3–C3–C4–H4 and H4–C4–C5–H5 dihedral angles would agree better with the reversed magnitudes. A strong effect of the N–SO_3^-, O–SO_3^-, and the COO^- groups upon coupling constants in **1**–**4** is also seen in the GlcN,6S residue in **2**. $^{3}J_{H2–C2–C3–H3}$ is larger (11.03 Hz) than the $^{3}J_{H3–C3–C4–H4}$ (9.92 Hz) in the GlcN,6S residue though the torsion angles would again suggest the opposite trend (166.0° and 171.9°, respectively) (the IdoA2S residue in the $^{2}S_0$ form). This is not the case when the IdoA2S residue adopts the $^{1}C_4$ form in **2**. Then $^{3}J_{H2–C2–C3–H3}$ is larger (11.34 Hz) than the $^{3}J_{H3–C3–C4–H4}$ (9.40 Hz) and is in accordance with the torsion angles (179.7° and 169.9°, respectively). The above evidence illustrates a relatively complex effect of the N–SO_3^- group conformation on the structure of the α-D-glucopyranose ring (different bond lengths, torsion angles within the GlcN,6S ring and electron densities depending upon the iduronate conformation) and consequently on the magnitude of coupling constants. Similar stereoelectronic effects of the COO^- group, including the influence of the O-5 oxygen lone-pairs, also affected the above mentioned values of both $^{3}J_{H3–C3–C4–H4}$ and $^{3}J_{H4–C4–C5–H5}$ in **1**.

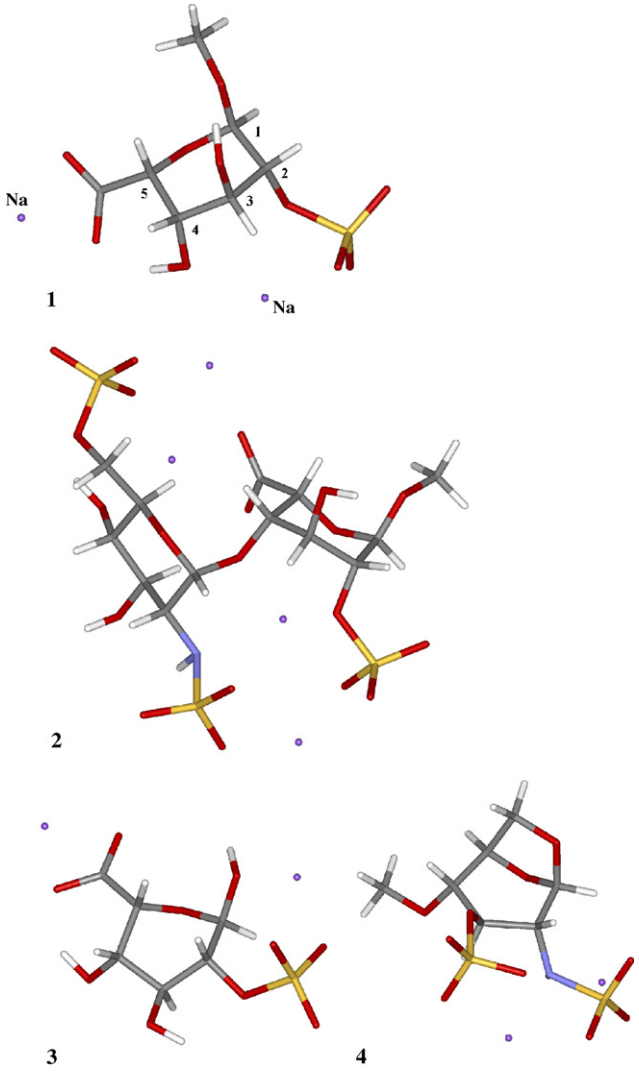

Figure 1. Structures of the compounds IdoA2S-OMe (**1**), GlcN,6S-IdoA2S-OMe (**2**), IdoA2S-OH (**3**), and 1,6-anhydro-GlcN,3S (**4**).

Computed geometry, energy, and coupling constants enabled a detailed analysis of the IdoA2S conformer populations in both compounds **1** and **2**. Recently, we have analyzed the conformer distribution in monosaccharide IdoA2S-OMe.[8] Computed energies showed that the most stable conformer is the $^{1}C_4$ form (99%) and the $^{4}C_1$ form contributes to the equilibrium with about 1%. Based on these populations, averaged coupling constants agreed well with experimental values.[4] In the present study, minute differences in both the computed averaged $^{3}J_{H–C–C–H}$ values and the conformer

Table 1. Computed three-bond proton–proton coupling constants (in Hz) in various sulfated carbohydrate derivatives

Compound	Pyranose conformation	Array of atoms	Torsion angle	$^3J_{H-C-C-H}$
1 IdoA2S-OMe	$^1C_4$	H1–C1–C2–H2	71.6	1.67
		H2–C2–C3–H3	288.8	3.00
		H3–C3–C4–H4	74.6	3.12
		H4–C4–C5–H5	48.8	2.30
	$^2S_0$	H1–C1–C2–H2	143.4	4.76
		H2–C2–C3–H3	183.2	10.53
		H3–C3–C4–H4	149.3	4.49
		H4–C4–C5–H5	33.9	4.11
	$^4C_1$	H1–C1–C2–H2	162.8	7.31
		H2–C2–C3–H3	195.0	9.01
		H3–C3–C4–H4	177.5	10.22
		H4–C4–C5–H5	303.5	6.53
2 GlcN,6S-IdoA2S-OMe	$^1C_4$ IdoA2S	H1–C1–C2–H2	71.9	1.61
		H2–C2–C3–H3	285.9	2.60
		H3–C3–C4–H4	78.5	2.54
		H4–C4–C5–H5	53.4	2.72
	$^4C_1$ GlcN,6S	H1–C1–C2–H2	55.5	3.70
		H2–C2–C3–H3	179.7	11.34
		H3–C3–C4–H4	169.9	9.40
		H4–C4–C5–H5	193.3	9.20
	$^2S_0$ IdoA2S	H1–C1–C2–H2	146.5	5.41
		H2–C2–C3–H3	180.8	11.15
		H3–C3–C4–H4	147.8	4.87
		H4–C4–C5–H5	36.2	5.09
	$^4C_1$ GlcN,6S	H1–C1–C2–H2	68.1	2.17
		H2–C2–C3–H3	166.0	11.03
		H3–C3–C4–H4	171.9	9.92
		H4–C4–C5–H5	205.8	7.66
3 IdoA2S-OH	$^2S_0$	H1–C1–C2–H2	140.8	4.59
		H2–C2–C3–H3	179.4	12.20
		H3–C3–C4–H4	155.0	5.69
		H4–C4–C5–H5	31.2	4.66
4 1,6-Anhydro-GlcN,3S	E_0	H1–C1–C2–H2	77.5	1.04
		H2–C2–C3–H3	244.6	1.14
		H3–C3–C4–H4	108.2	0.54

distribution were obtained when the $^3J_{H-C-C-H}$ values were fitted to the experimental values and populations

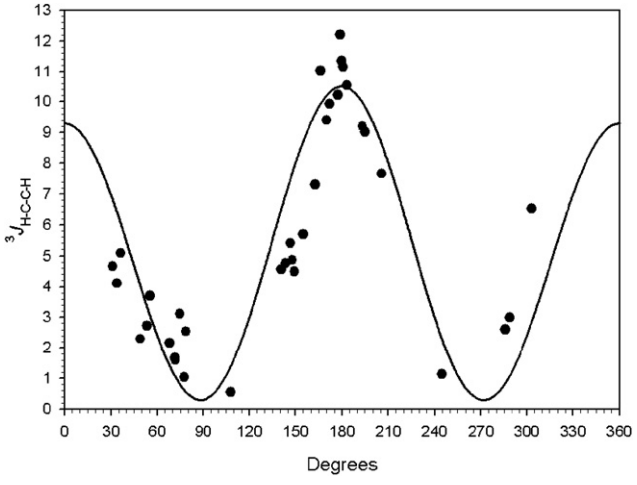

Figure 2. Relationship between $^3J_{H-C-C-H}$ (values in Hz) and torsion angles in the H–C–C–H atom array. Dots represent the computed $^3J_{H-C-C-H}$ values in compounds **1**–**4**.

were determined from the fitting procedure. In this case (Table 2), the ratio shifted very slightly to 97:3 ($^1C_4$:$^4C_1$; compared to 99:1 as determined from energies) and showed very good agreement of conformer distribution obtained by two independent methods. Very good agreement between experimental[9] and the computed averaged $^3J_{H-C-C-H}$ values were also obtained when evaluating the presence of two forms ($^1C_4$:$^2S_0$, 79:21) of the IdoA2S residue in heparin disaccharide GlcN,6S-IdoA2S-OMe. Unlike in the monosaccharide, the skewed $^2S_0$ form is the minor form in this disaccharide. The presence of the skewed form in conformational equilibrium is in agreement with the conformational properties of heparin where the two forms ($^1C_4$ and $^2S_0$) of the IdoA2S residue are in equilibrium. The population of these forms in heparin was determined using the geometry and $^3J_{H-C-C-H}$ values in the monomer IdoA2S-OMe. In this case, the ratio of the chair and the skew forms is 65:35 ($^1C_4$:$^2S_0$) indicating that the skewed form is more populated in the polymer than in the disaccharide and is also in agreement with

Table 2. Experimental (third column) and computed, best-fit (fourth column) three-bond proton–proton coupling constants (values in Hz) in various GAGs

		$^3J_{H–C–C–H}$ Exp.[a]	$^3J_{H–C–C–H}$ Comp.	$^1C_4$	$^2S_0$	$^4C_1$
IdoA2S-OMe	H1–H2	1.8	1.8	97	0	3
	H2–H3	3.3	3.2			
	H3–H4	3.4	3.3			
	H4–H5	2.2	2.4			
GlcN,6S-IdoA2S-OMe						
IdoA2S	H1–H2	2.3	2.4	79	21	0
	H2–H3	4.6	4.4			
	H3–H4	3.4	3.1			
	H4–H5	2.6	3.2			
GlcN,6S	H1–H2	3.6	3.4			
	H2–H3	11.4	11.3			
	H3–H4	9.5	9.5			
	H4–H5	10.4	8.9			
Pentasaccharide						
IdoA2S	H1–H2	4.0	3.5	42	58	0
	H2–H3	7.5	7.4			
	H3–H4	3.6	3.9			
	H4–H5	3.1	3.3			
Heparin						
IdoA2S	H1–H2	2.6	2.7	65	35	0
	H2–H3	5.9	5.6			
	H3–H4	3.4	3.6			
	H4–H5	3.1	2.9			

The computed conformer abundances in the IdoA2S residue are given in the last three columns.
[a] Taken from Refs. 4 and 9.

experiments.[4,9,10] An even higher population of the $^2S_0$ form ($^1C_4$:$^2S_0$, 42:58) was obtained in heparin-pentasaccharide.[11] The computed populations in pentasaccharide and heparin are comparable with those previously estimated,[4] however, the present data suggest higher populations of the $^1C_4$ chair form in both cases.

In summary, the present data show that the DFT calculations at the B3LYP/6-311++G** level can account for the effect of the charged sulfate and carboxylate groups and provide geometries and $^3J_{H–C–C–H}$ values in very good agreement with the experimental data. Calculations for compounds **1**–**4** generated a number of theoretical values that enabled us to propose the relationship between $^3J_{H–C–C–H}$ and torsion angles in GAGs. As the relationship was obtained using compounds bearing sulfate and carboxylate groups, it appears more suitable for its application to GAGs than relationships parametrized considering only neutral molecules. It should help determine conformer populations more accurately and thus to understand GAG properties in solution[4,5,11–13,10,14] and in complexes with proteins[2,3,15–18] in more detail.

1. Methods

The geometries of investigated compounds (Fig. 1) have been completely optimized with the JAGUAR program[19] using density functional theory (DFT) with Lee–Young–Parr (B3LYP)[20] correlation functional and the 6-311++G** basis set. Hybrid functionals with the Slater local functional/Becke88 nonlocal gradient correlation[21] and Vosko–Wilk–Nusair local functionals[22] were used. Geometry optimizations were obtained with the gradient optimization routine, the convergence criteria were set to 1×10^{-5}. Starting geometries were based on the previous data.[8,12,17] NMR proton–proton spin–spin coupling constants were computed with the Gaussian03 program[23] at the B3LYP level of theory.[24]

Acknowledgements

This research was supported by VEGA Grant No. 2/5075/25, APVT Grant No. 51-034504, and EU Grant No. QLK3-2002-02049.

References

1. *Heparin—Chemical and Biological Properties*; Lane, D. A., Lindahl, U., Eds.; CRC Press: Boca Raton, FL, 1989.
2. Conrad, H. E. *Heparin-Binding Proteins*; Academic Press: San Diego, 1998.
3. Casu, B.; Lindahl, U. *Adv. Carbohydr. Chem. Biochem.* **2001**, *57*, 159–208.
4. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Torri, G.; Casu, B.; Gatti, G.; Jacquinet, J. C.; Sinay, P.; Petitou, M.; Choay, J. *J. Am. Chem. Soc.* **1986**, *108*, 6773–6778.

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. Hricovíni, M.; Nieto, P. M.; Torri, G. In *NMR Spectroscopy of Glycoconjugates*; Jimenez-Barbero, J., Peters, T., Eds.; Wiley-VCH, 2002; pp 189–229.
7. Haasnoot, C. A. G.; DeLeeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783–2792.
8. Hricovíni, M. *Carbohydr. Res.* **2006**, *341*, 2575–2580.
9. LaFerla, B.; Lay, L.; Guerrini, M.; Poletti, L.; Panza, L.; Russo, G. *Tetrahedron* **1999**, *55*, 9867–9880.
10. Guerrini, M.; Agulles, T.; Bisio, A.; Hricovíni, M.; Lay, L.; Naggi, A.; Poletti, L.; Sturiale, L.; Torri, G.; Casu, B. *Biochem. Biophys. Res. Commun.* **2002**, *292*, 222–230.
11. Torri, G.; Casu, B.; Gatti, G.; Petitou, M.; Choay, J.; Jacquinet, J.-C.; Sinay, P. *Biochem. Biophys. Res. Commun.* **1985**, *128*, 134–140.
12. Hricovíni, M.; Guerrini, M.; Bisio, A. *Eur. J. Biochem.* **1999**, *261*, 789–801.
13. Ernst, S.; Venkataraman, G.; Sasisekharan, V.; Langer, R.; Cooney, C. L.; Sasisekharan, R. *J. Am. Chem. Soc.* **1998**, *120*, 2099–2107.
14. Angulo, J.; Nieto, P. M.; Martin-Lomas, A. *Chem. Commun.* **2003**, *13*, 1512–1513.
15. Capila, I.; Linhardt, R. J. *Angew. Chem., Int. Ed.* **2002**, *41*, 390–412.
16. Mulloy, B.; Forster, M. J. *Glycobiology* **2000**, *10*, 1147–1156.
17. Hricovíni, M.; Guerrini, M.; Bisio, A.; Torri, G.; Petitou, M.; Casu, B. *Biochem. J.* **2001**, *359*, 265–272.
18. Canales, A.; Angulo, J.; Ojeda, R.; Bruix, M.; Fayos, R.; Lozano, R.; Jiménez-Gallego, G.; Martin-Lomas, M.; Nieto, P. M. *J. Am. Chem. Soc.* **2005**, *127*, 5778–5779.
19. JAGUAR 3.5, Schrodinger, Portland, 1998.
20. Lee, C.; Yang, W.; Paar, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
21. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
22. Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1208.
23. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. GAUSSIAN *03, Revision C.02*; Gaussian: Wallingford CT, 2004.
24. Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560–571.