Pergamon

Bioorganic & Medicinal Chemistry, Vol. 5, No. 7, pp 1301–1309, 1997

PII: S0968-0896(97)00087-4
0968-0896/97 $17.00 + 0.00

Combined NMR and Molecular Modeling Study of an Iduronic Acid-Containing Trisaccharide Related to Antithrombotic Heparin Fragments

Soizic Cros,[a] Maurice Petitou,[b] Philippe Sizun,[c] Serge Pérez[d] and Anne Imberty[d,*]

[a]*Ingénierie Moleculaire, INRA, BP 1627, 44316 Nantes Cédex 03, France*
[b]*SANOFI Recherche, 195 route d'Espagne, 31000 Toulouse, France*
[c]*SANOFI Recherche, 371 rue du Professeur Blayac 34000 Montpellier, France*
[d]*CERMAV†-CNRS, BP 53, F-38041 Grenoble Cédex 09, France*

Abstract—An iduronic acid-containing trisaccharide, methyl-*O*-(4-*O*-methyl-2,3,6-tri-*O*-sulfo-α-D-glucopyranosyl-(1→4)-*O*-(2-*O*-sulfo-α-L-idopyranosyluronic acid)-(1→4)-*O*-2,6-di-*O*-sulfo-α-D-glucopyranoside, related to antithrombotic heparin fragments has been subjected to a combined NMR and molecular modeling investigation. The conformational behavior of the two constituting disaccharide segments was investigated using a systematic grid search approach with the MM3 force field along with the proper parameters for the sulfate ester group. The exploration of the potential energy surfaces of the trisaccharide was performed through the use of the CICADA methods interfaced with the MM3 force field. In all cases, the 2-*O*-sulfo-α-L-iduronate moiety was given the three favored ring conformations $^1C_4$, $^4C_1$, and $^2S_0$. Conformations were clustered into families, four of which are likely to exhibit significant occupancy in solution. The different low-energy conformational families display different orientations at the glycosidic linkages and/or different ring shapes for the iduronate ring. The $^2S_0$ conformation is the major one for the 2-*O*-sulfo-α-L-iduronate but is still in equilibrium with the $^1C_4$ ring shape. The occurrence of such a conformational equilibrium in solution was probed via high-resolution NMR spectroscopy through measurements of coupling constants and NOE build-up. These results are in keeping with the observation that 2-*O*-sulfated pentasaccharides display a similar affinity for antithrombin III as their 2-*N*-sulfated counterparts. © 1997 Elsevier Science Ltd.

Introduction

In the polysaccharide heparin, the unique pentasaccharide sequence shown in Figure 1 is involved in binding to the serine protease inhibitor antithrombin III (AT-III).[1-4] This pentasaccharide contains an L-iduronic acid unit, which may adopt different conformations ($^4C_1$, $^1C_4$, and $^2S_0$) according to the nature and the substituents of the neighboring residues, two sulfated glucosamine units.[5,6] It has thus been suggested, though direct experimental evidence has yet to be obtained, that the presence of a remarkable 3-*O*-sulfated glucosamine unit next to iduronic acid in the AT-III binding sequence pushes the conformation toward $^2S_0$ and thereby facilitates the binding of the protein.[7] In order to simplify the preparation of such complex sulfated oligosaccharides, we have shown that 2-*O*-sulfated glucose units can be introduced instead of *N*-sulfated glucosamine units without affecting the biological activity.[8] The aim of the present work is to assess to what extent the conformation of L-iduronic acid is influenced by this structural modification. Using molecular modeling and ^{1}H NMR we have studied the conformation of the iduronic acid containing trisaccharide (see Fig. 1) corresponding to the FGH sequence present in the genuine AT-III binding pentasaccharide. Comparison with previous studies carried out on the latter[5-9] indicates that the substitution of *O*-sulfates for *N*-sulfates does not alter the conformational equilibrium of the 2-*O*-sulfated L-iduronic acid unit.

*Tel: +33 (0)4 76 03 76 36; Fax: +33 (0)4 76 54 72 03; e-mail: imberty@cermav.cnrs.fr
†Affiliated with University Joseph Fourier, Grenoble, France.

Results and discussion

Relaxed maps of the constituent disaccharides

Two disaccharides F–G and G–H are needed to build the trisaccharide. Each of them was studied by a systematic grid search approach using the MM3 program.[10] The three favored conformations of the iduronate ring ($^2S_0$,$^1C_4$, and $^4C_1$) were considered, yielding therefore three different (Φ,Ψ) energy maps for these dimers. Consequently, six adiabatic maps were calculated; they are shown in Figure 2.

For the F–G disaccharide, the adiabatic maps relative to the $^4C_1$ and $^1C_4$ conformations of the iduronate ring exhibit only one main low-energy region (energy minimum at [80, 270] and [90, 210] for $^4C_1$ and $^1C_4$, respectively). The potential energy surface (PES)

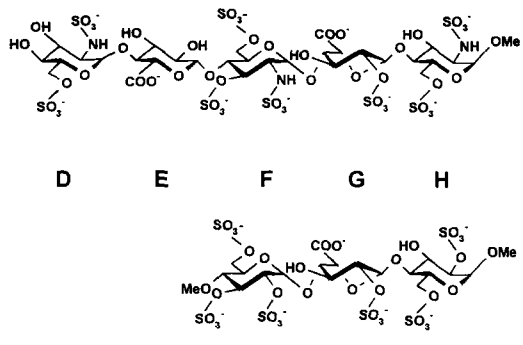

Figure 1. Structure of the pentasaccharide representing the antithrombin III binding site of heparine, and of the trisaccharide (4-*O*-methyl-2,3,6-tri-*O*-sulfate-α-D-glucose(1→4)2-*O*-sulfate-α-L-iduronate(1→4)2,6-di-*O*-sulfate-α-L-methyl-glucoside studied here. The L-iduronic ring is shown in the $^2S_0$ conformation.

computed with an $^2S_0$ conformation of the iduronate ring displays two low-energy minima, A (80, 200) and B (80, 60). The exocyclic group of the F residue displays a GT conformation. It has to be noted that a conformational transition between a $^1C_4$ (or $^4C_1$) and $^2S_0$ conformation of the iduronate ring occurs spontaneously in the calculation when the Φ and Ψ angles are in the region 80, 40. This creates the small low-energy region of the PES calculated for the F–G ($^1C_4$) and $^4C_1$ disaccharides.

When compared to the F–G linkage, the PES calculated for the G–H disaccharide is more extended, which indicates a greater flexibility for this linkage. Irrespective of the ring conformation of the G residue, the adiabatic maps are very similar and show two main energy regions. The (Φ,Ψ) values of the two minima of the PES calculated for the G($^1C_4$)-H, G($^4C_1$)-H and G($^2S_0$)-H disaccharides are, respectively, A (270, 200), (290, 240), (280, 250) and B (250, 60), (280, 80) and (270, 60). In this case, no conformational interconversion between the $^1C_4$ and $^2S_0$ conformation is observed.

These potential energy surfaces are very similar to those calculated by the MM2 program for the constituting disaccharides of the heparin pentasaccharide fragment[9] and of the disaccharide repeating sequence of heparin,[11] the only difference being that the B minimum did not appear in the work by Ragazzi et al.[9]

NMR experiments on the trisaccharide (F–G–H)

The proton chemical shifts and T_1 values of the trisaccharide F–G–H are collected in Table 1. The values of the chemical shifts are in agreement with previous values obtained for the *N*-sulfated trisaccharide[12] and for the F–G–H moiety of the pentasaccharide fragment of heparin.[13] Only the H–2F and H–2G

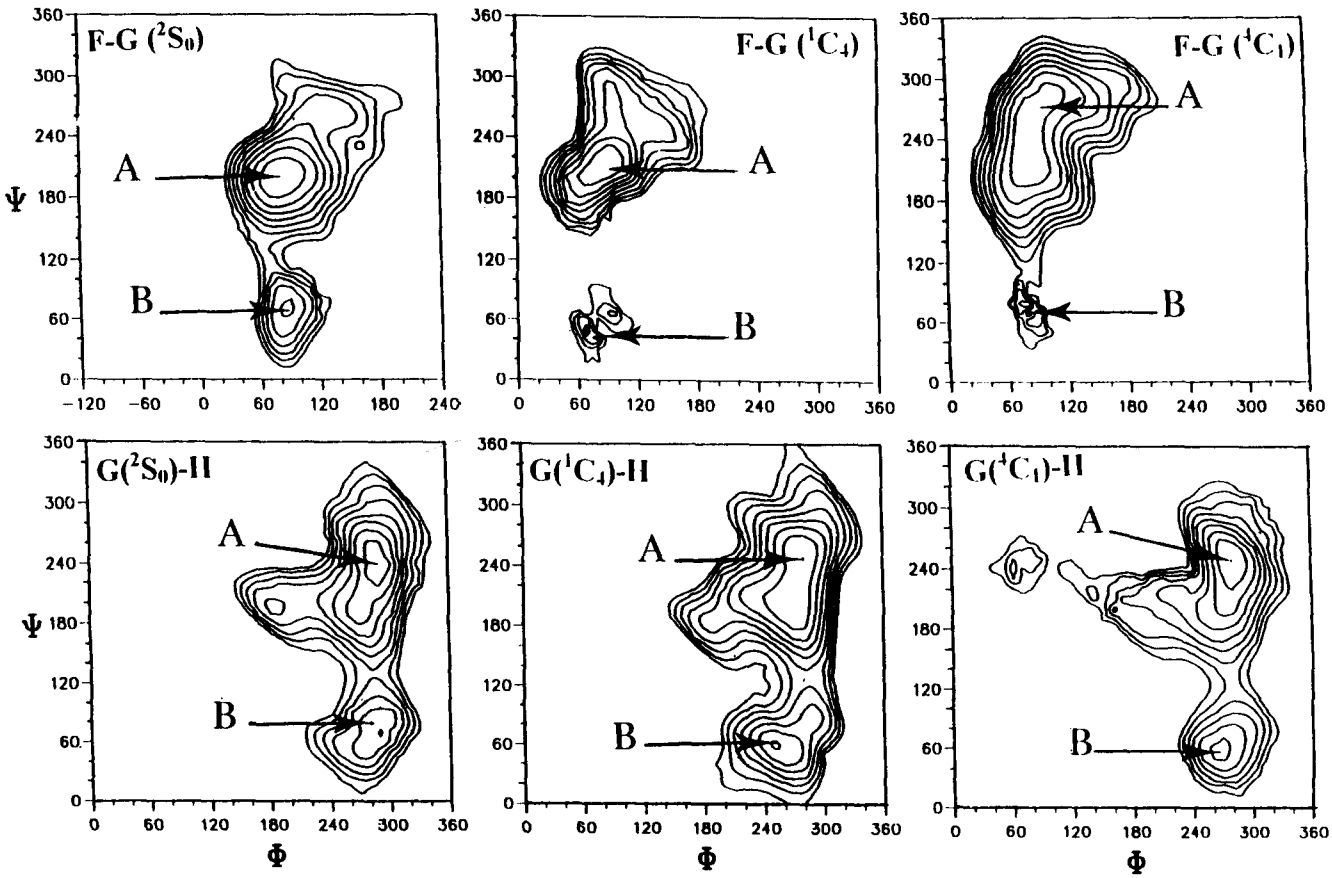

Figure 2. (Φ,Ψ) Energy maps of all the constituting-disaccharides of the trisaccharide. Energy contours have been drawn by step of 1 kcal mol^{-1} above the minimum of each map. Energy minima have been indicated by a label.

Table 1. 500 MHz ^{1}H chemical shift (ppm), ^{1}H T_1 values (s) and coupling constants (Hz) measured for the trisaccharide

Protons	δ_H (ppm)	T_1 (s)	Coupling constants	$^3J_{H-H}$ (Hz)
H–1F	5.64	1.05	$^3J_{H-1F\_H-2F}$	3.6
H–2F	4.29		$^3J_{H-2F\_H-3F}$	9.5
H–3F	4.56	2.05	$^3J_{H-3F\_H-4F}$	9.5
H–4F	3.51	1.10	$^3J_{H-4F\_H-5F}$	9.5
H–5F	4.08	0.95	$^3J_{H-5F\_H-6rF}$	2.4
H–6sF	4.31		$^3J_{H-5F\_H-6sF}$	2.1
H–6rF	4.20			
H–1G	5.11	0.95	$^3J_{H-1G\_H-2G}$	4.4
H–2G	4.28		$^3J_{H-2G\_H-3G}$	8.3
H–3G	4.10	1.35	$^3J_{H-3G\_H-4G}$	4.3
H–4G	4.17		$^3J_{H-4G\_H-5G}$	3.5
H–5G	4.71	1.30		
H–1H	5.09	1.40	$^3J_{H-1H\_H-2H}$	3.6
H–2H	4.22		$^3J_{H-2H\_H-3H}$	9.4
H–3H	3.84	1.45	$^3J_{H-3H\_H-4H}$	9.5
H–4H	3.81	1.35	$^3J_{H-4H\_H-5H}$	9.5
H–5H	3.96	0.90	$^3J_{H-5H\_H-6rH}$	4.0
H–6sH	4.43	0.40	$^3J_{H-5H\_H-6sH}$	2.1
H–6rH	4.33			

displays smaller values in the present study. This can be attributed to the presence of *O*-sulfate moiety instead of an *N*-sulfate one, at these two positions.

Experimental $^3J_{H-H}$ values are listed in Table 1. Values for the F and H residues are very similar and correspond to a $^4C_1$ conformation of the glucopyranosyl ring.[12] The coupling constants of the idose residue reflect the shape of the ring and therefore display different values depending on the neighboring residues, the substituents, the concentration, or the salt content.[12–15] The J_{H-H} measured for the G residue in the present work are in close agreement with the ones measured on the *N*-sulfated trisaccharide,[5,8,11,12] indicating therefore that the substitution of *N*-sulfate by *O*-sulfate in position 2 of the neighboring residues does not affect the conformational behavior of the idose ring.

The following NOEs were clearly observed in each ring: H–1F/H–2F, H–3F/H–4F, H–4F/H–2F, H–1G/H–2G, H–1G/H–3G, H–3G/H–5G, H–4G/H–5G, H–2G/H–5G, the latter one being characteristic of the $^2S_0$ conformer. Some other intraresidue NOEs of smaller amplitudes were also observed such as: H–2F/H–3F, H–4F/H–5F, H–6rF/H–6sF, H–2G/H–3G, H–4H/H–5H, H–2H/H–6sH, H–6sH/H–6rH. Six interresidue NOEs were large enough to be quantified. In addition to the two NOEs across the glycosidic linkage, that is H–1F/G–4H and H–1G/H–4H, four others are observed. The additional NOE through the F–G linkage (H–1F/H–3G) has been observed in the pentasaccharide fragment of heparin.[9,16] Three additional NOEs are observed through the G–H linkage: H–1G/H–5H, H–1G/H–6sH, and H–1G/H–6rH. The evolution of the interresidue NOEs as a function of mixing times is presented in Figure 3.

A correlation time of 3.25×10^{-10} s was calculated for the trisaccharide F–G–H from the measured average ^{13}C T_1 value of the methine carbon and using proton–carbon distance of 0.11 nm.[17,18]

Potential energy surface of the trisaccharide (F–G–H)

The conformational behavior of the trisaccharide displayed in Figure 1 has been fully investigated. Three CICADA runs were performed, each using a different conformation of the iduronate ring as a starting point. In all cases, the search was stopped after 1100 minima were found, and the total number of MM3 calculations was then 19,000.

The complete ensemble of conformations resulting from the CICADA analysis has been clustered into different conformational families, within an energy window of 5 kcal mol^{-1} above the global minimum. The conformational families having a population of more than 1% are listed in Table 2. For the trisaccharide with the G residue in a $^2S_0$ conformation (CIC1), one family overwhelms the others, representing more than 90% of the whole population. The conformers of this family adopt the global minimum (A) for both F–G and G–H glycosidic linkages. A second and a third family exist; they represent only 6% of the population. These two families exhibit large variations in their Ψ angle about the G–H linkage. Family 3 corresponds to the B minimum, whereas the Ψ angle for family 2 is only 30° different from the A minimum. No conformational transitions are observed for the conformation of the iduronate ring in this run.

The CICADA run starting with a trisaccharide having a G monomer in a $^1C_4$ conformation (CIC2) reflects a more flexible conformational behavior for the oligosaccharide. Conformational transitions are observed for the iduronate ring. Consequently, for the two most populated families (>78% of the conformations), the iduronate ring, initially in $^1C_4$ conformation, switches to an $^2S_0$ shape. These two most populated conformations correspond closely to the best families of the previous run. Only a low population of the conformers (10%) remains in a $^1C_4$ form. This population is distributed into two families, which differ by the orientation of the Φ torsion angle at the G–H glycosidic linkage.

For the CICADA run starting with the trisaccharide having the iduronate residue in $^4C_1$ conformation (CIC3), no interconversion of the ring is observed. Nevertheless, some flexibility appears at the F–G glycosidic linkage. Both families adopt the low energy conformation (A) for the F–G and G–H disaccharides, but they display a difference of about 30° for the Ψ torsion angle at the F–G glycosidic linkage. This is because of the occurrence of a large low-energy plateau in the F–G ($^4C_1$) potential energy map.

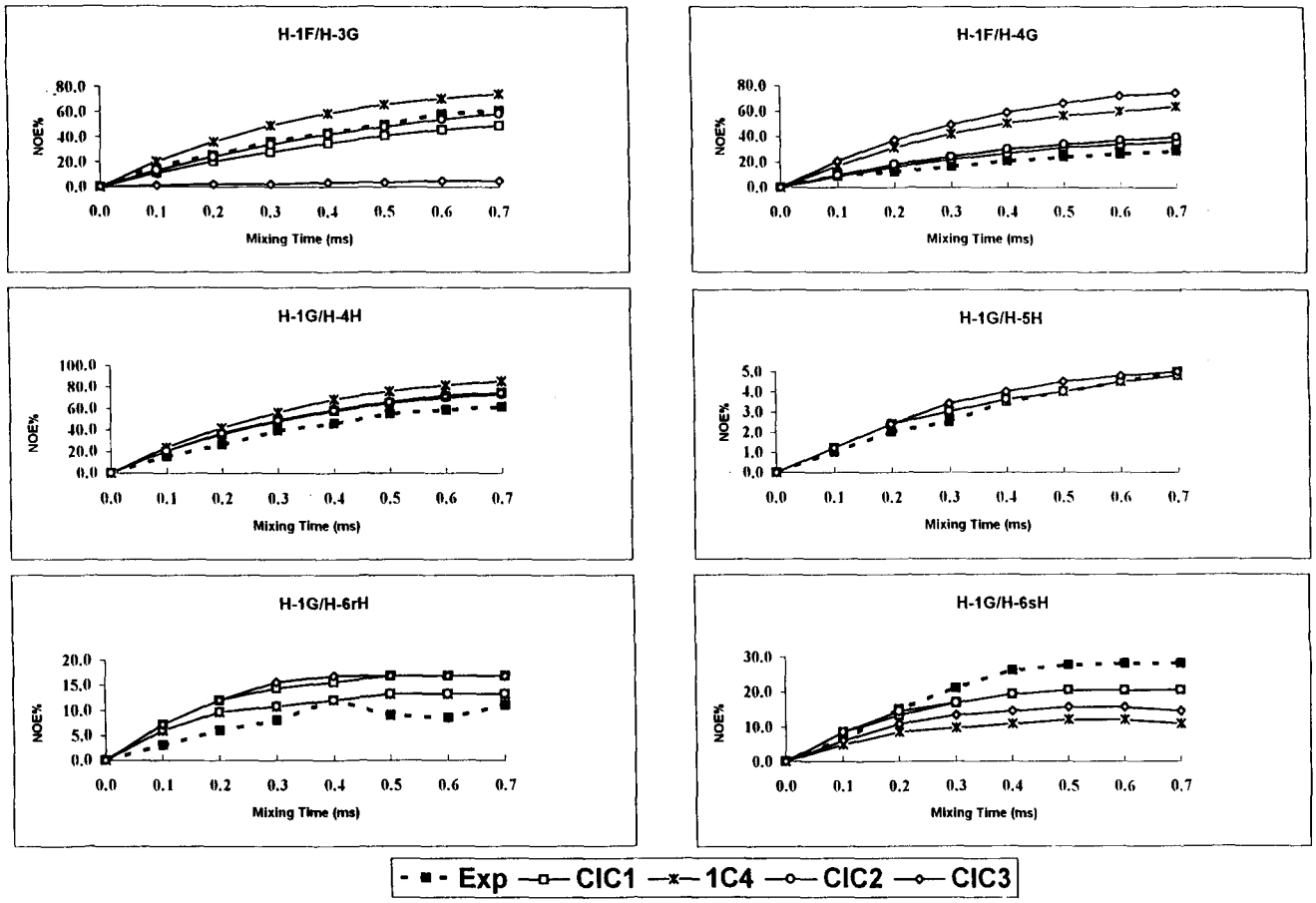

Figure 3. Comparison of experimental (Exp) and theoretical NOE build-up curves for the trisaccharide. CIC1, CIC2, and CIC3 represent the populations of conformers of the three CICADA calculations whereas $^1C_4$ represents the subset of CIC2 population with the idose ring in the $^1C_4$ shape.

Figure 4 displays the most important conformational families found by CICADA with the iduronate ring being in a $^1C_4$ conformation. For each family, the lowest energy conformation and also those with the largest variations for the linkage torsion angles within an energy window of 5 kcal mol^{-1} are drawn. Inside each family, only small variations are observed about the glycosidic linkage. The major variations occur for the orientations of the *O*-sulfate groups mainly for the 2-position of the G residue and at the 6-position of the H residue.

As it has been previously observed,[5,6] the conformation of the G residue can interconvert from $^1C_4$ to $^2S_0$ without affecting the global conformation of the trisaccharide, since the Φ and Ψ values remain unaltered. The major difference between these two low-energy conformations resides in the orientation of the *O*-sulfate group of the iduronate residue.

Comparison of calculated and experimental NMR parameters

Coupling constants. Theoretical $^3J_{H-H}$ coupling constants have been calculated for all the glycosidic rings of the trisaccharide F–G–H following the methodology described in the Methods section. The F and H residues remain in their $^4C_1$ chair conformations. As for the G residue, variations in the averaged coupling constants of each run can be correlated to the iduronate ring conformation (Table 3). The $^1C_4$ conformation yields small values for the four intraring $^3J_{H-H}$ coupling constants. The $^2S_0$ and $^4C_1$ conformations yield a maximum value for the $^3J_{H2-H3}$ and $^3J_{H3-H4}$ coupling constants respectively and medium values for the other coupling constants.

None of the three conformers alone can reproduce the experimental values. The CICADA run which started with the $^1C_4$ iduronate shape (CIC2), and which resulted in a population of 78% $^2S_0$ and 22% of $^1C_4$, gives theoretical coupling constants which are in excellent agreement with the experimental ones. (It can be noted that when calculating the percentage of the three shapes directly from the coupling constants, a distribution of P^2S_0= 76%, P^1C_4 = 24% and P^4C_1 = 0% is predicted.) This distribution is in agreement with those previously determined on the equivalent trisaccharide of heparin.[5,6]

NOE build-ups. Ensemble average NOE curves were calculated on different populations of the trisaccharide

Table 2. Characteristics of the trisaccharide conformational families

Families	Idose ring	Φ3	Ψ3	Φ4	Ψ4	Erel	%
Run CIC1							
Family 1							
conf-min	$^2S_0$	78.1	195.7	286.9	246.2	0.0	91.1
min		66.0	188.0	283.7	244.8		
max		84.0	201.8	297.2	255.8		
Family 2							
conf-min	$^2S_0$	79.9	193.9	270.7	192.5	1.6	3.45
min		74.1	193.7	265.7	183.3		
max		80.6	198.1	275.3	200.6		
Family 3							
conf-min	$^2S_0$	76.9	192.7	273.7	57.8	2.0	1.6
min		76.1	192.1	263.8	55.9		
max		80.3	196.9	279.8	65.19		
Run CIC2							
Family 1							
conf-min	$^2S_0$	78.1	194.6	287.2	243.9	0.0	70.8
min		71.4	191.7	283.0	241.5		
max		80.6	197.6	295.9	251.6		
Family 2							
conf-min	$^2S_0$	77.1	196.5	272.9	185.2	1.0	7.8
min		75.3	194.7	268.8	182.9		
max		82.4	198.1	277.1	192.4		
Family 3							
conf-min	$^1C_4$	82.7	206.1	270.3	197.6	1.2	6.6
min		81.0	205.8	268.3	195.3		
max		83.5	208.0	271.2	199.8		
Family 4							
conf-min	$^1C_4$	93.3	284.2	280.5	247.9	1.3	3.3
min		91.5	275.2	276.4	244.5		
max		99.8	293.9	285.1	256.1		
Run CIC3							
Family 1							
conf-min	$^4C_1$	78.8	221.5	281.0	249.4	0.0	51.7
min		68.5	204.5	278.4	241.0		
max		84.3	231.0	300.6	268.4		
Family 2							
conf-min	$^4C_1$	79.0	256.3	283.1	247.7	0.004	23.7
min		77.5	249.8	281.8	243.9		
max		81.6	266.6	300.9	264.9		
Family 3							
conf-min	$^4C_1$	105.0	285.0	282.0	249.7	0.6	11.4
min		90.3	265.7	276.0	239.7		
max		150.9	298.7	288.2	259.2		

Only the families having an energy-weighted population more than 1% at 298 K have been listed. For each starting conformation of the iduronate ring ($^2S_0$, $^1C_4$, and $^4C_1$) and each family, conf-min indicates the glycosidic linkages torsion angles and relative energy of the lowest energy conformation, whereas min and max are the limit values for each torsion angle within an energy window of 5 kcal mol^{-1}.

as a function of the mixing time. The theoretical interresidue NOEs are drawn and superimposed with the corresponding experimental data (Fig. 3). NOEs across the F–G glycosidic linkage are more sensitive to the idose conformation that those across the G–H linkage.

Populations with only $^1C_4$ or $^4C_1$ (CIC3) pseudorotamer do not reproduce the experimental values. The agreement is much better with the CIC1 conformation which is constituted of $^2S_0$ conformers. However, when quantifying the comparison by calculating an agreement factor (Table 4), the CIC2 run, (78%, ($^2S_0$)/22% ($^1C_4$)) represents the best agreement.

Conclusion

It has been suggested that the conformation of L-iduronic acid in the pentasaccharide binding site of heparin plays a major role in antithrombin III binding and activation.[5–7,9] Earlier studies indicated that this conformation is strongly dependent on the nature of the substituent borne by L-iduronic acid,[5,9] and it was thus of interest to investigate the conformational properties of the present trisaccharide in which L-iduronic acid is surrounded by two 2-*O*-sulfated glucose units in place of the *N*-sulfated glucosamine units usually present in heparin and the antithrombin binding sequence. The results indicate that this structural modification hardly affects the conformational properties of L-iduronic acid, and that the $^2S_0$ conformer is still the major one, in

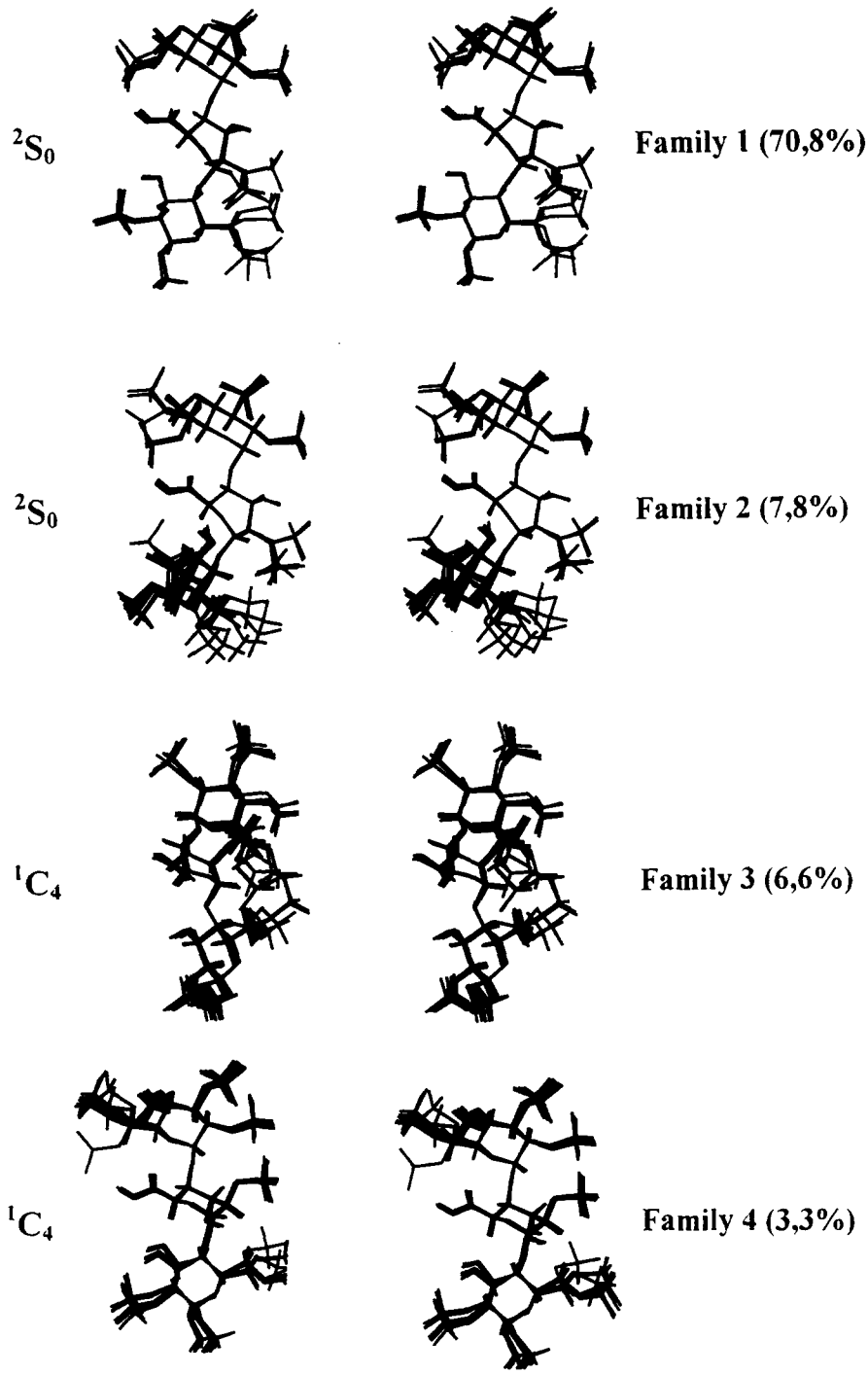

Figure 4. Conformational families determined for the trisaccharide during the CICADA run CIC2.

equilibrium with the $^1C_4$ one. The results are in keeping with the observation that pentasaccharides containing 2-*O*-sulfonates instead of 2-*N*-sulfonates display a similar affinity for antithrombin III.[8,19]

Methods

NMR experiments

Trisaccharide (9.6 mg) was dissolved in 0.5 mL of D_2O (99.96%, Commissariat à l'Energie Atomique, Saclay France). ^1H NMR spectra (500 MHz) were obtained with an AMX 500 spectrometer, at 300 K. Typically 32 K data were used, with a spectral width of 6000 Hz. Resolution enhancement of free induction decay (Fid), was performed by a Lorentzien–Gaussian method. Assignment of the ^1H signals was obtained from the COSY phase[20] sensitive spectrum, acquired as 4 K × 512 points (sine bell weighting was applied in each dimension), zero-filled to 4 K × 2 K prior to double Fourrier Transform. $^3J_{H,H}$ coupling constants were obtained from spectra 1D and completed with COSY spectrum. The NOESY phase sensitive[21] was acquired

Table 3. Comparison of theoretical and experimental coupling constants for the 2-*O*-sulphate iduronic ring (G)

$^3J_{H-xG\_-H-x+1G}$	CIC1	CIC2	$^1C_4$[a]	CIC3	Exp
$^3J_{H-1G\_-H-2G}$	5.86	5.19	1.86	5.09	4.4
$^3J_{H-2G\_-H-3G}$	10.15	8.88	2.37	5.49	8.3
$^3J_{H-3G\_-H-4G}$	5.03	4.68	2.76	9.29	4.3
$^3J_{H-4G\_-H-5G}$	3.26	2.95	1.18	5.25	3.5

[a]Only the $^1C_4$ conformations extracted from the population obtained with the starting trisaccharide with G in $^1C_4$ form.

at 290 K, as 2 K × 400 points, with mixing periods of 100, 200, 300, 400, 500, 600, and 700 ms. NOESY crosspeak intensities were evaluated with the UXNMR program.

Molecular modeling

Nomenclature. A schematic representation of the pentasaccharide fragment of the heparin is shown in Figure 1 along with the labelling of the residues and the atoms of interest. The orientation of the *O*-sulfate group in position *x* is described by two torsion angles defined as:

$$\mu_x = C_{(x-1)} - C_x - O_x - S_x$$

$$\nu_x = C_x - -O_x - S_x - O_1S_x$$

Whereas the *O*-sulfate in position 6 needs a third torsion angle for defining its orientation, which can be referred as GG, GT, and TG.[22]

$$\omega_x = O_5 - C_5 - C_6 - O_6/C_4 - C_5 - C_6 - O_6$$

The three favored conformations of the α-L-iduronate ring ($^2S_0$, $^1C_4$, and $^4C_1$) are described by the values of three intraring torsion angles defined as χ^1, χ^2, and χ^3. For each ring shape, the minimum and maximum values of these angles are reported in Table 5.

$$\chi^1 = \Theta(O_5G - C_1G - C_2G - C_3G)$$

$$\chi^2 = \Theta(C_2G - C_3G - C_4G - C_5G)$$

$$\chi^3 = \Theta(C_3G - C_4G - C_5G - O_5G)$$

The torsional angles of the glycosidic linkages of a disaccharide $i(1 \rightarrow x)j$ are defined as:

Table 4. R factor calculated for all inter-glycosidic build-up NOEs

NOEs	CIC1	$^1C_4$	CIC2	CIC3
H-1F/H-3G	0.207	0.300	0.053	0.921
H-1F/H-4G	0.284	1.365	0.417	1.790
H-1G/H-4H	0.234	0.437	0.226	0.254
H-1G/H-5H	0.062	0.062	0.062	0.227
H-1G/H-6rH	0.732	0.357	0.357	0.774
H-1G/H-6sH	0.243	0.536	0.259	0.408
All	0.260	0.529	0.210	0.725

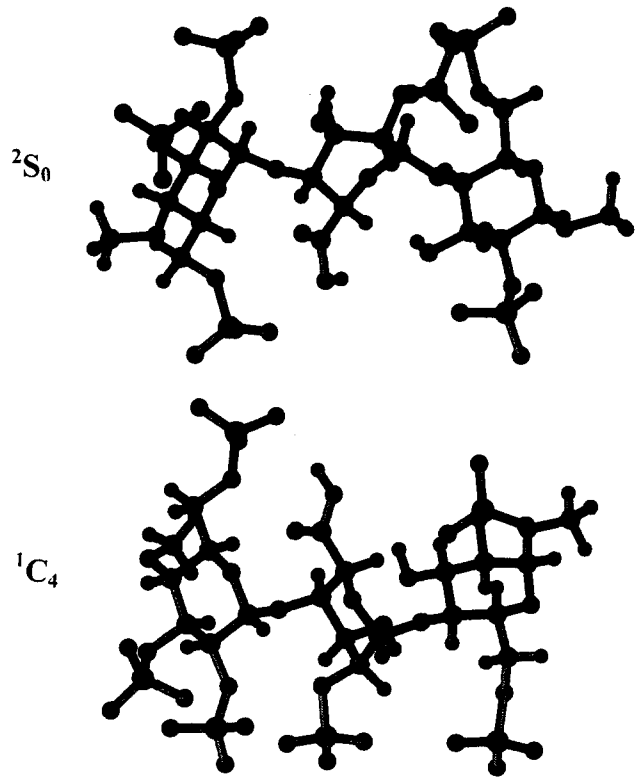

Figure 5. Lowest energy conformation for the trisaccharide with idose ring having either the $^1C_4$ or the $^2S_0$ conformation.

$$\Phi = O_5i - C_1i - O_xi - C_xi$$

$$\Psi = C_1i - O_xi - C_xi - C_xj - C_{x-1}j$$

The signs of the torsion angles are in agreement with the IUPAC-IUB conventions.[23]

Energy calculations. Geometry optimizations were performed using the molecular mechanics program MM3(92).[24,25] The MM3 force field has been extensively used for the study of carbohydrate structures.[26,27] The block-diagonal minimization method, with the default energy-convergence criterion ($0.00008 \times n$ kcal mol^{-1} per five iterations, *n* being the number of atoms) was used for geometry optimization. The dielectric constant was set to 78 in order to lower the influence of hydrogen bonding on the potential energy surface.

Modeling of monosaccharides. All the starting blocks were taken from MONOBANK, a database of three-dimensional structures for monosaccharides.[28] The three favored conformations $^1C_4$, $^4C_1$, and $^2S_0$ of the 2-*O*-sulfate-α-L-iduronate were constructed using geometrical characteristics previously determined by molecular mechanics.[29]

All monosaccharides that constitute the modified heparin fragment contain at least one *O*-sulfate group. The parameters relative to this group have been determined by the MM2 program and then have been added to the MM3 force-field.[30] For all the monomers constituting the trisaccharide, the preferred orientation

of all the *O*-sulfate groups was assessed by the DRIVER option of the MM3. The conformation was then fully optimized.

Relaxed maps of the disaccharides. The two disaccharides constituting the trisaccharide displayed in Figure 1 have been studied by a systematic grid search method. The calculations were always performed on the methyl glycoside. Starting from a refined geometry, the procedure drives Φ and Ψ torsion angles in steps of 10 over the whole angular range while the MM3 program provides full geometry relaxation. Several maps are calculated for each disaccharide in order to take into account the three staggered orientations of the hydroxymethyl groups. The three favored conformations ($^2S_0$, $^1C_4$, and $^4C_1$) of the iduronic residue are also taken into account for the disaccharides F–G and G–H. For each disaccharide, the results of these calculations are synthesized in a so-called adiabatic map where only the conformer with the lowest energy for each (Φ, Ψ) value is considered. Iso-energy contours are then plotted by the interpolation of 1 kcal mol^{-1} within a limit of 8 kcal mol^{-1}.

CICADA calculations. Exploration of the potential energy surface (PES) of the oligosaccharide was performed with a recent method, CICADA (channels in conformational space analyzed by driver approach).[31] The CICADA program, which is an interface to the MM3 force-field, has been proven to be very efficient in conformational studies of oligosaccharides.[32,33] Input for the CICADA program consists mainly of one or a few conformers in MM3 format and a file containing the list of torsion angles to be driven/monitored. During the CICADA calculations, each individual torsion angle is driven in each direction from the initial conformation at a given increment. At each step, the structure is optimized except for the driven torsion angle. When CICADA detects a minimum, the conformation is fully optimized, including the driven torsion angle. The resulting structure is compared with the previously stored ones and stored as a new geometry if it is one that has not yet been discovered. Structure corresponding to energy maxima, the transition states are also saved. Calculations stop when no new conformers (local minima) are found within a desired energy window.

For the F–G–H trisaccharide, several conformations were built and fully refined to serve as starting points for the cicada runs. The driven torsion angles were Φ and Ψ at each linkage, the first torsion angle of all *O*-sulfated groups, and the ω angle in the case of a sulfated group in position 6. The step of increment of the driven torsion angles was set at 20°. The number of driven torsion angles, and therefore the dimensionality of the potential energy surface to explore, was 13. Three torsion angles which depend on the conformation of the cycle of the iduronate residue were not driven but were monitored. Two conformations were considered different if one of the driven or monitored angles differed by at least 30°. A relative energy cut-off of 50 kcal mol^{-1} was applied for exploring the PES, whereas a relative energy cut-off of 10 kcal mol^{-1} was applied for considering a new geometry as a new conformer. The search was stopped after 9000 minima on the potential energy surface were found.

Analysis of the potential energy surface. The conformation and transition states found by CICADA were analyzed by the PANIC program,[34] which explores the paths along the PES. Conformations were clustered into families,[35] and their relative populations were calculated applying a Boltzmann distribution at a temperature of 25 °C. The alignment of the different conformers of each family of the trisaccharide was performed with a FIT procedure of the SYBYL molecular modeling software.[36] Only the ring atoms and the glycosidic oxygens of the iduronate residue (G) were fitted.

Calculation of theoretical NMR data

Coupling constants. For each conformation of the trisaccharide, coupling constants $^3J_{H-H}$, for vicinal hydrogen atoms were calculated using a Karplus-type equation,[37] which accounts for J dependence on the dihedral angle (ϕ) of the H–C–C–H fragment and for the electronegativities and orientation of the substituents.

NOE build-up. The conformational analysis yielded the population of conformers of the trisaccharide from which the ensemble average $\langle r^{-6} \rangle$ values were obtained using the full relaxation approach.[38] Ensemble average NOEs were calculated from the relaxation matrix $\langle R \rangle$ as described previously.[39] A value of 0.325 ns was used for the τ_c value. Theoretical build-up NOE curves were obtained as a function of the mixing times used in the

Table 5. Maximum and minimum values of the torsion angles χ^1, χ^2, and χ^3, which determined the puckering of the 2-*O*-sulphate-iduronic-acid. The values of the torsion angles of the idose residue after minimization by MM3 program (Conf) are also indicated

	χ^1			χ^2			χ^3		
	Min	**Max**	**Conf**	**Min**	**Max**	**Conf**	**Min**	**Max**	**Conf**
$^2S_0$	0	80	23.5	0	30	27.3	30	80	33.1
$^1C_4$	−90	−20	−54.6	−90	−20	−56.5	20	90	61.3
$^4C_1$	30	90	54.2	30	90	53.6	−90	−20	−52.0

experiments. The assessment of the agreement between observed (NOE_0) and calculated data (NOE_I), was performed using a relationship recently proposed to describe the quality of the fit in NMR/modeling studies approach.[40]

$$R_{NOE} = \frac{\sum | NOE_0 - NOE_C |}{\sum | NOE_0 |}$$

References

1. Casu, B.; Oreste, P.; Torri, G.; Zoppetti, G.; Choay, J.; Lorneau, J.-C.; Petitou, M.; Sinaÿ, P. *Biochem. J.* **1981**, *197*, 599.

2. Choay, J.; Lormeau, J.-C.; Petitou, M.; Sinaÿ, P.; Fareed, J. *Annals N.Y. Acad. Sci.* **1981**, *370*, 644.

3. Lindahl, U.; Bäckström, G.; Thunberg, L.; Leder, I. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6551.

4. Thunberg, L.; Bäckström, G.; Lindahl, U. *Carbohydr. Res.* **1982**, *100*, 393.

5. Ferro, D. R.; Provasoli, A.; Ragazzi, M.; Torri, G.; Casu, B.; Gatti, G.; Jacquinet, J.-C.; Sinaÿ, P.; Petitou, M.; Choay, J. *J. Am. Chem. Soc.* **1986**, *108*, 6773.

6. Forster, M. J.; Mulloy, B. *Biopolymers* **1993**, *33*, 575.

7. Casu, B.; Petitou, M.; Provasoli, A.; Sinaÿ, P. *Trends Biochem. Sci.* **1988**, *13*, 221.

8. Petitou, M.; Jaurand, G.; Derrien, M.; Duchaussoy, P.; Choay, J. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 95.

9. Ragazzi, M.; Ferro, D.R.; Perly, B.; Torri, G.; Casu, B.; Sinaÿ, P.; Petitou, M.; Choay, J. *Carbohydr. Res.* **1990**, *195*, 169.

10. MM3, **1992** QCPE, Creative Arts Building 181, Indiana University, Bloomington, IN 47405, U.S.A.

11. Mulloy, B.; Forster, M. J.; Jones, C.; Davies, D. B. *Biochem. J.* **1993**, *293*, 849.

12. van Boeckel, C. A. A.; Aelst, S. F.; Wagenaar, G. N.; Mellema, J.-R. *Recueil des travaux des Pays-Bas* **1987**, *106*, 19.

13. Torri, G.; Casu, B.; Petitou, M.; Choay, J.; Jacquinet, J.-C.; Sinaÿ, P. *Biochem. Biophys. Res. Commun.* **1985**, *128*, 134.

14. Ferro, D. R.; Provasoli, A.; Ragazzi, M. *Carbohydr. Res.* **1990**, *195*, 157.

15. Petitou, M; Duchaussoy, P; Lederman, I; Choay, J; Jacquinet, J.-C.; Sinaÿ P; Torri G. *Carbohydr. Res.* **1987**, *167*, 67.

16. Hricovini, M.; Torri, G. *Carbohydr. Res.* **1995**, *268*, 159.

17. Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546.

18. Hervé du Penhoat, C.; Imberty, A.; Roques, N.; Michon, V.; Mentech, J.; Descotes, G.; Pérez, S. *J. Am. Chem. Soc.* **1992**, *113*, 3720.

19. van Boeckel, C. A. A.; Petitou, M. *Angew. Chem. Int. Ed. Engl.* **1993**, *32*, 1671.

20. Marion, D.; Wuthrich, K. *Biochem. Biophys. Res. Com.* **1983**, *113*, 967.

21. Bodenhausen, G.; Kogler, H.; Ernst, R. R. *J. Mag. Res.* **1984**, *58*, 370.

22. Marchessault, R. H.; Pérez, S. *Biopolymers* **1979**, *18*, 2369.

23. IUPAC–IUB Commission on Biochemical Nommenclature. *Arch. Biochem. Biophys.* **1971**, *145*, 255.

24. Allinger, N. L.; Zhu, Z.-Q. S.; Chen, K. *J. Am. Chem. Soc.* **1992**, *114*, 6120.

25. Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1990**, *112*, 8293.

26. Pérez, S.; Imberty, A.; Carver, J. P. *Adv. Comput. Biol.* **1994**, *1*, 147.

27. French, A. D.; Dowd, M. K. *J. Mol. Struct. (Theochem)* **1993**, *286*, 183.

28. Pérez, S.; Delage, M.-M. *Carbohydr. Res.* **1991**, *212*, 253.

29. Ragazzi, M.; Ferro, D. R.; Provasoli, A. *J. Comp. Chem.* **1986**, *7*, 105.

30. Lamba, D.; Glover, S.; Mackie, W.; Rashid, A.; Sheldrick B.; Pérez, S. *Glycobiology* **1994**, *4*, 151.

31. Koca, J. *J. Mol. Struct. (Theochem)* **1994**, *308*, 13.

32. Koca, J.; Pérez, S.; Imberty, A. *J. Comp. Chem.* **1995**, *16*, 796.

33. Imberty, A.; Mikros, E.; Koca, J.; Mollicone, R.; Oriol, R.; Pérez, S. *Glycoconj. J.* **1995**, *12*, 331.

34. Koca, J. *J. Mol. Struct.* **1993**, *291*, 255.

35. Imberty, A.; Pérez, S. *Glycobiology* **1994**, *4*, 351.

36. SYBYL 6.04 Tripos Associates, 1699 S. Hanley Road, Suite 303, St Louis, MO 63144, U.S.A.

37. Haasnoot, C. A. G. ; de Leeuw, F. A. A. M.; Altona, C. *Tetrahedron* **1980**, *36*, 2783.

38. Brisson, J.-R.; Carver, J. P. *Biochemistry* **1983**, *32*, 1362.

39. Peters, T.; Weimar, T. *J. Biomol. NMR* **1994**, *4*, 97.

40. Shriver, J.; Edmonson, S. *Biochemistry* **1993**, *32* 1610.

(*Received in U.S.A.* 24 *December* 1996; *accepted* 10 *February* 1997)