Glycoconjugate Journal (1997) **14:** 973–981

EGC1
Conformational analysis of methyl 6-*O*-[(*R*)- and (*S*)-1-carboxyethyl]-*α*-D-galactopyranoside by MM and Langevin dynamics simulations

Roland Stenutz and Göran Widmalm*

Department of Organic Chemistry, Arrhenius Laboratory Stockholm University, S-106 91 Stockholm, Sweden

The conformational space of methyl 6-*O*-[(*R*)- and (*S*)- 1-carboxyethyl]-*α*-D-galactopyranoside has been investigated. A grid search employing energy minimization at each grid point over the three major degrees of freedom, namely ϕ, ψ and ω, identified low energy regions. The *R*-isomer shows five low energy conformers within ca. 1 kcal mol^{-1} of the global energy minimum. The *S*-isomer has two conformers within a few tenths of a kcal mol^{-1} of the global energy minimum. Langevin dynamics simulations have been have been performed at 300 K for 30 ns of each isomer. The ϕ dihedral angle has as its major conformer (g−) for the *R*-isomer whereas it is the (g+) conformer for the *S*-isomer. For the ψ dihedral angle the (t) conformer has the highest population for both isomers. The dihedral angle ω has the (g+) conformer most highly populated, both for the *R*- and *S*-isomer. The above five and two conformational states for the *R*- and *S*-isomers, respectively, make up 90% in each case of the populated states during the Langevin dynamics (LD) simulations. Rate constants for the ω dihedral angle have been calculated based on a number correlation function. Three bond homo- and heteronuclear, i.e. proton and carbon-13, coupling constants have been calculated from the dynamics trajectories for comparison to experimental values. The heteronuclear coupling constant H2′,C6 has been measured for the *S*-isomer and found to be 3.3 Hz. The *J* value calculated from the LD simulations, namely 2.6 Hz, is in fair agreement with experiment. A comparison to the X-ray structure of the *R*-isomer shows that the conformation of the crystalline compound occupies the low energy region most highly populated as a single *R*-conformer (30%) during the LD simulations.

***Keywords*: Carboxyethyl, Langevin dynamics, long-range carbon-proton coupling constants, molecular mechanics, number correlation function, NMR spectroscopy**

Introduction

In molecular recognition processes the three dimensional shape of a molecule is of utmost importance. The configurational change at only one carbon atom may cause a dramatic change in binding of a ligand to its receptor. Substitution, even with small groups such as *O*-acetyl, phosphate or 1-carboxyethyl, can cause great variation in the biological response.

Substitution of carbohydrates by 1-carboxyethyl groups occur in some bacterial polysaccharides. We have in the present study investigated the three dimensional structure of methyl 6-*O*-[(*R*)- and (*S*)-1-carboxyethyl]-α-D-galactopyranoside. The *R*-isomer is a constituent of the capsular polysaccharide from a sheep pathogen, namely *Butyrivibrio fibrisolvens* strain X6C61 [1]. We have recently synthesized 1-carboxyethyl substituted sugars and characterized these with NMR and CD spectroscopy [2].

Low energy conformations were investigated by molecular mechanics and energy minimizations using a grid search, i.e. the potential energy of the system was evaluated at regular intervals in space. The dynamics of the system was also studied using molecular dynamics simulations (MD) which are based on Newton's equations of motion. The variant of the simulation technique called Langevin dynamics was employed for inclusion of solvent effects.

From the grid search five and two low energy conformers could be identified for the *R*- and *S*-isomer, respectively. The energy difference for these conformers to the global energy minimum was small, which leads to a significant population of all of these conformers at room temperature.

Simulations were performed using Langevin dynamics. The simulation times were well into the nanosecond time regime and between most conformers a large number of transitions took place. This means that for this part of the

*To whom correspondence should be addressed.

0282–0080 © 1997 Chapman & Hall

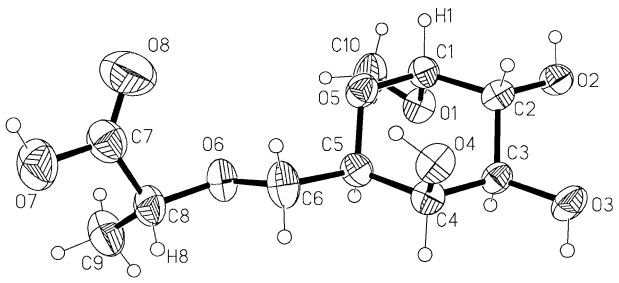


Figure 1. X-ray crystal structure of methyl 6-*O*-[(*R*)-1-carboxyethyl]-*α*-D-galactopyranoside.

conformational space, sampling is good. In Langevin dynamics coupling to a heat bath occurs and the molecule experiences random collisions with a virtual solvent. The above mentioned five and two conformers for the *R*- and *S*-isomers, respectively, make up approximately 90% of the populated structures during the dynamics simulation time.

An experimental way to investigate the conformation of a molecule is to measure three-bond homo- and heteronuclear coupling constants, usually proton and carbon-13. These are then evaluated using a Karplus type relationship, i.e. the dependence of the coupling constant on the dihedral angle of the atoms involved. From the dynamics trajectories we have calculated the averaged coupling constants of interest for the three dihedral angles that make up the conformational space. One of these heteronuclear coupling constants was measured, namely $^3J_{C,H}$ for H2′,C6, and found to be 3.3 Hz to be compared to 2.6 Hz calculated from the simulations. The other calculated values will be compared to experimental values to be measured. Recently, the crystal structure of the *R*-isomer (Figure 1) was determined [3]. It was found to have an extended conformation. The grid search also identified this conformer as a low energy one which was the most populated single conformer in the Langevin dynamics simulation of the *R*-isomer.

Methods

Simulation

Simulations were performed using Quanta/CHARMm version 4.0 (Molecular Simulations Inc., Burlington, MA, USA) with PARM22 running on Silicon Graphics Inc. workstations. The methyl 6-*O*-[(*R*)- and (*S*)-1-carboxyethyl]-α-D-galactopyranosides were built by substituting position 6 of methyl α-D-galactopyranosides by a 1-carboxyethyl group. The total charge of the molecule was attenuated to -0.5 and spread by the Gasteiger method [4]. A constant dielectric of unity was used in all simulations. The atoms in the 1-carboxyethyl group are primed. The three torsional angles of interest are defined as follows:

ϕ = C6-O6-C2′-H2′, ψ = C5-C6-O6-C2′ and

ω = O5-C5-C6-O6.

The conformational space was sampled in vacuo at 30° intervals in ϕ, ψ and ω. At each grid point the three dihedral angles were restrained before energy minimization was performed using 50 steps of steepest descent followed by 100 steps of Newton-Raphson or until the rms gradient was below 0.01 kcal (mol-Å)$^{-1}$.

Low energy regions were identified and for these parts an additional grid search was performed with 10° increments.

The Langevin dynamics (LD) simulation technique was used to simulate the effects of solvent water [5]. This includes a collision frequency $\gamma = \zeta/m$, where ζ is the friction constant and m the atomic mass. A value of $\gamma = 50$/ps was applied to carbons and oxygens. The lowest energy conformers, as obtained from the grid search, were used as starting points for the LD simulations. Each dynamics run was initiated using a random seed for obtaining initial velocities. Three 10 ns runs at 300 K were performed for each isomer using a time step of 1.0 fs.

NMR spectroscopy

The *S*-isomer was dissolved in D_2O (pD 8, Na salt) to a concentration of 0.06 M. NMR spectra were measured at 303 K on a Varian Unity 500 MHz NMR spectrometer. Measurement of long-range $^{13}C,^1H$ coupling constants were performed using a pulsed field gradient version [6] of ^{13}C site selective excitation with band-selective proton decoupling during the aquisition time as devised by Nishida *et al.* [7]. The ^{13}C site selective excitation used a Gaussian shaped pulse of 200 ms duration. A spectral width of 4000 Hz was sampled with 16 384 data points using 20 000 transients. The FID was processed using Varian VNMR software and Felix 2.3 (Biosym Technologies, San Diego, CA, USA). Zero-filling eight times and multiplication of the FID with an exponential weighting function with a line broadening of 0.3 Hz was applied prior to Fourier transformation and *J*-doubling. Eight delta functions in the frequency domain was used for the *J*-doubling procedure [8, 9].

Coupling constants

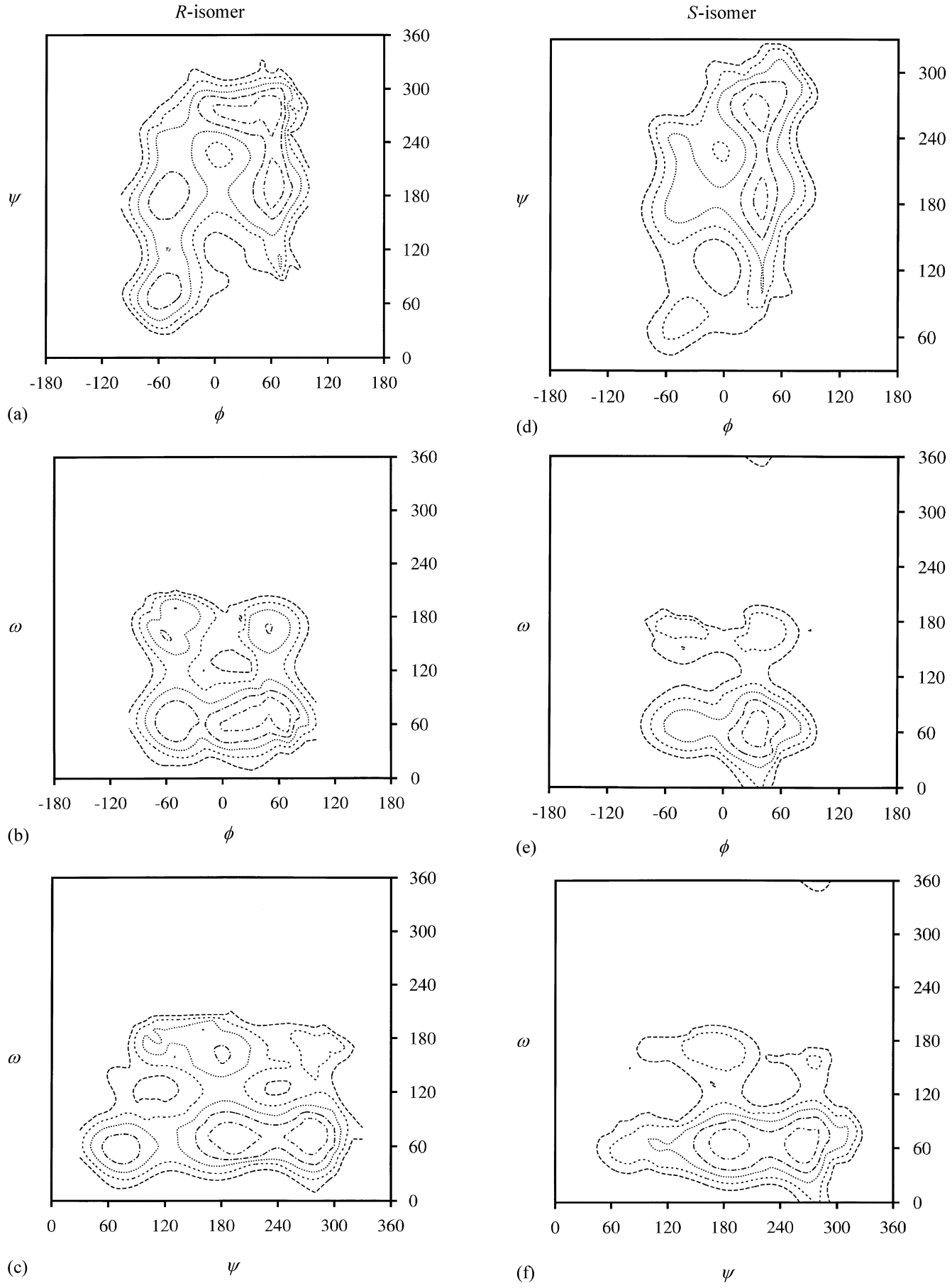
The 3JH,H coupling constants for H5,H6R and H5,H6S were calculated using equation D of Haasnoot *et al.* [10] and the equation of Tvaroska *et al.* [11] was used for the $^3J_{CH}$ couplings of C2′,H6R; C2′,H6S and C6,H2′. The coupling constants for every step of the dynamics trajectories were calculated and averaged.

Results and Discussion

Grid search

A three dimensional grid search was performed on both isomers. From preliminary studies a carboxylate group was

Figure 2. Ramachandran plots. The isoenergy levels are shown at 1 kcal mol^{-1} increments above the global energy minimum.


R-isomer
S-isomer
ψ
ϕ
ω
(a)
(b)
(c)
(d)
(e)
(f)
-180 -120 -60 0 60 120 180
0 60 120 180 240 300 360

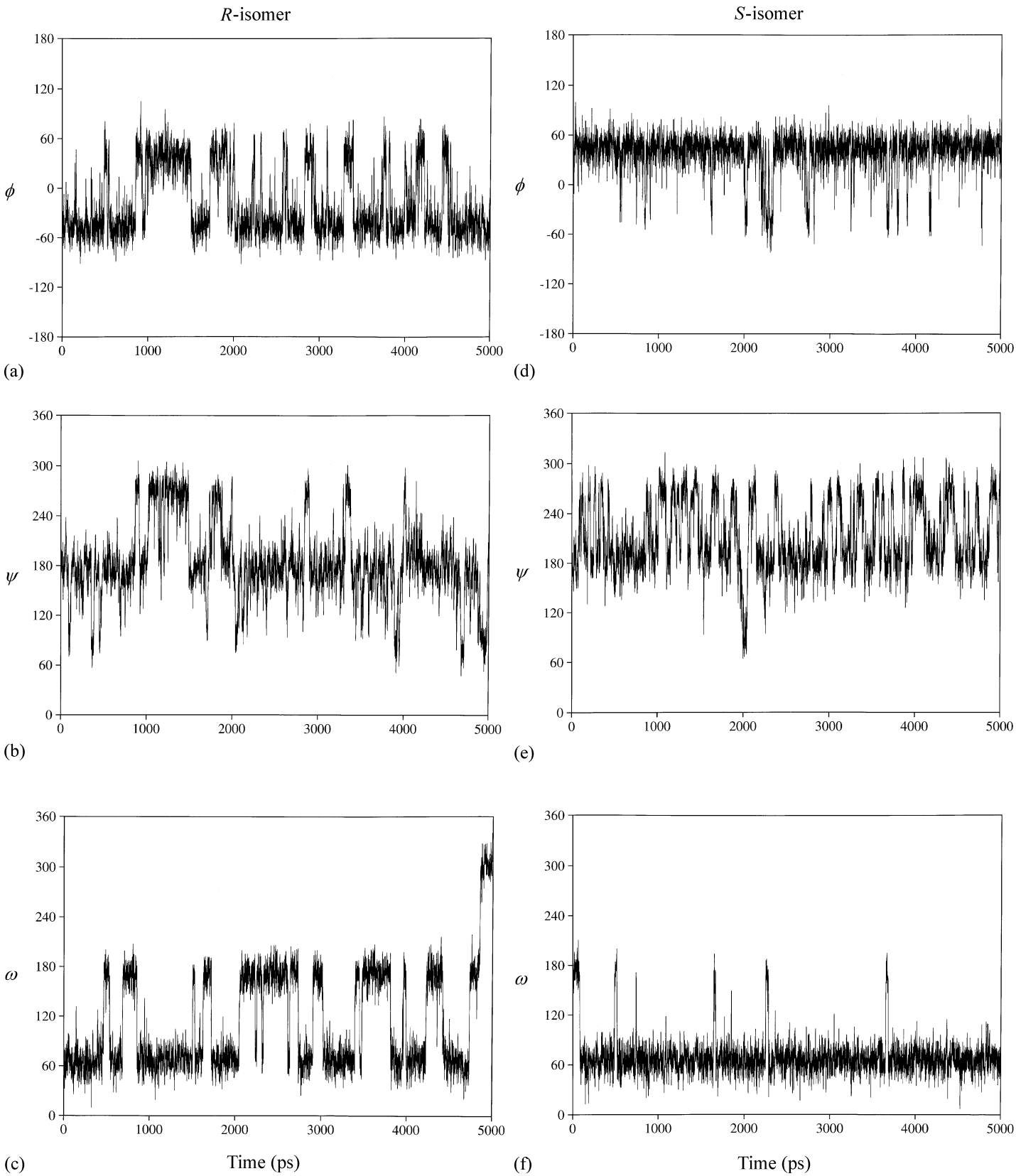


Figure 3. Dihedral angle trajectories from LD simulation.

chosen since previous NMR studies had been performed with a sodium counter ion but more importantly hydrogen bonding was observed between the carboxylic OH group and the ring oxygen O5 when molecular dynamics was performed in vacuo. In addition, prior to the grid search, the energy of the system having hydroxyl hydrogen bonding around the pyranose ring as clockwise or counter clockwise was investigated. The counter clockwise conformers, i.e. H-bonds towards O1, were found to have lower energy and this directionality was subsequently used in the grid search.

For both isomers a 3D grid search at 30° increments was performed. Regions of low energy were identified and for these additional grid searches at 10° increments were obtained. Projections of energy minimized regions from 3D space were then performed as ϕ vs. ω, ϕ vs. ψ and ψ vs. ω for the *R*-isomer and the *S*-isomer (Figure 2). From the low energy regions further energy minimization was done. Five low energy conformers were identified for the *R*-isomer and two for the *S*-isomer (Table 1). The *R*-isomer shows its global minimum (ϕ, ψ, ω) as the conformer (g+/g−/g+). The five low energy conformers all reside within ca. 1 kcal mol^{-1} and should thus all be populated to a significant extent. The *S*-isomer has its global energy minimum conformer also at (g+/g−/g+) and the next local energy minimum conformer resides only a few tens of a kcal mol^{-1} above the global energy minimum. Thus, both conformers should be populated about equally.

Dynamics simulation

Dynamics simulations were performed at 300 K. Langevin dynamics (LD) was chosen as this technique also makes it possible to include solvent effects. It is based on molecular dynamics to which a friction constant and a random force have been added. The collision frequency γ was set to 50/ps which should be appropriate for this kind of system based on previous simulations using Langevin dynamics [12, 13]. The molecule experiences random collisions which model solvent molecules. For each isomer three 10 ns simulations were performed.

These were analysed as trajectories for the *R*-isomer and the *S*-isomer (Figure 3). Scatter plots of the *R*-isomer and the *S*-isomer (Figure 4) show the sampling of conformational states. Results from the LD runs have also been tabulated (Table 2).

Table 1. Major conformers: Angles and populations were taken from LD, energies and structures from energy minimizations

Conformer		*Angle*	*Energy (Kcal mol^{-1})*	*Population*
(*R*)-isomer				
ϕ	g$^-$	−50°		
ψ	t	178°	30.4	32%
ω	g$^+$	66°		
ϕ	g$^-$	−52°		
ψ	g$^+$	75°	30.4	8%
ω	g$^+$	57°		
ϕ	g$^+$	31°		
ψ	g$^-$	−90°	29.9	16%
ω	g$^+$	67°		
ϕ	g$^+$	50°		
ψ	t	−169°	31.0	11%
ω	g$^+$	66°		
ϕ	g$^+$	51°		
ψ	t	175°	30.2	20%
ω	t	173°		
(*S*)-isomer				
ϕ	g$^+$	46°		
ψ	g$^-$	−92°	29.1	35%
ω	g$^+$	68°		
ϕ	g$^+$	52°		
ψ	t	−172°	29.3	52%
ω	g$^+$	64°		

Table 2. Dihedral angles and populations

		ϕ		ψ		ω	
		Angle	*Pop.*	*Angle*	*Pop.*	*Angle*	*Pop.*
(*R*)-isomer							
	g$^+$	41°	28%	94°	10%	65°	65%
1st Run	t	–	–	177°	75%	172°	34%
	g$^-$	−45°	72%	269°	15%	302°	1%
	g$^+$	40°	34%	85°	16%	66°	72%
2nd Run	t	180°	3%	179°	62%	171°	19%
	g$^-$	−46°	63%	269°	22%	305°	9%
	g$^+$	39°	28%	87°	15%	66°	66%
3rd Run	t	179°	2%	177°	68%	171°	22%
	g$^-$	−49°	70%	268°	17%	305°	12%
	g$^+$	40°	30%	89°	14%	66°	68%
Average	t	180°	2%	178°	68%	171°	25%
	g$^-$	−47°	68%	269°	18%	304°	7%
(*S*)-isomer							
	g$^+$	45°	92%	93°	1%	65°	96%
1st Run	t	–	–	187°	65%	168°	3%
	g$^-$	−37°	8%	265°	34%	311°	1%
	g$^+$	45°	89%	89°	3%	65°	91%
2nd Run	t	–	–	187°	60%	169°	3%
	g$^-$	−36°	11%	264°	37%	305°	6%
	g$^+$	46°	94%	–	–	65°	97%
3rd Run	t	–	–	188°	56%	169°	3%
	g$^-$	−35°	6%	265°	44%	–	–
	g$^+$	45°	92%	91°	2%	65°	95%
Average	t	–	–	187°	60%	169°	3%
	g$^-$	−36°	8%	265°	38%	308°	2%

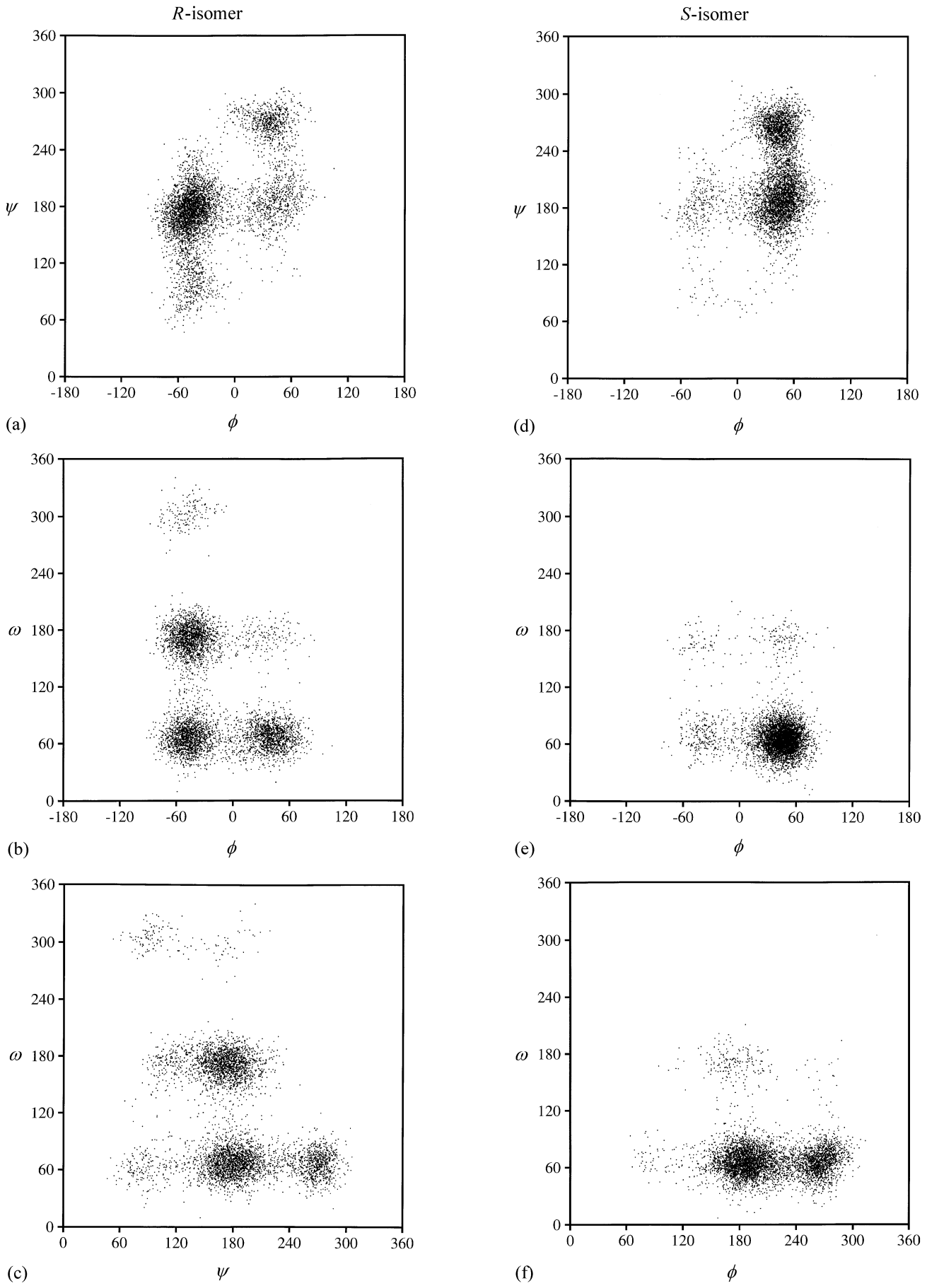


Figure 4. Scatter plots from LD simulation.

During LD the (g+) and (g−) states of ϕ, all three states of ψ and the (g+) and (t) states of ω are sampled >10% for the *R*-isomer. ϕ and ω occupy essentially only (g+), whereas ψ populates the states (t) and (g−) > 10% for the *S*-isomer. The dihedral angle ω does not occupy the (g−) state to any significant extent, i.e. where a 1,3-dipolar interaction takes place between O4 and O6. For both isomers more than half of the population of ψ resides in the (t) state, which gives an extended structure.

A comparison for the ϕ dihedral angle shows that the (t) state is negligible for both isomers; the difference observed is that the *R*-isomer has as its major state (g−) whereas the *S*-isomer has the state (g+). The five states identified from the grid search account for a population of 87% for the *R*-isomer and the two states of the *S*-isomer acount also for a population of 87% during the LD simulation.

Transitions between wells

Transitions between regions in conformational space (Table 3) have been calculated. When a barrier on a potential energy surface is lower than kT there is sufficient energy for interconversion between different states without obstruction by a barrier. As can be observed from the Ramachandran maps of the *R*-isomer and the *S*-isomer (Figure 2) the barriers between some energy minima are low, approximately 2–3 kT. Interconvertion between different states are numerous for (g+/g−) of ϕ and (t/g−) of ψ for both isomers and (g+/t) of ψ of the *R*-isomer. A substantial number of transitions also occur in ψ for (g+/t) of the *S*-isomer and in ω for (g+/t) of both isomers. We have only calculated the rate constants for one set of transitions in one run since, inter alia, low barrier heights make the rate description invalid.

We also conclude that some regions of conformational space are well sampled. It is obvious from the compilation of transitions that the greater the number, the smaller the deviation in each of the 10 ns simulation from the average.

Transition rates

For a bistable potential, transitions between the two wells A and B can be identified [14]. N_A and N_B are defined as the respective populations in the wells and k_{AB} and k_{BA} are the forward and the backward rate constants for the transition process. To investigate transition rates a number correlation function C(t) can be defined [14, 15] where each state along a trajectory is identified as unity if the molecule has the conformation of the reactant well and zero if it is in the product well. The relaxation time, τ, of C(t) is related as follows to the rate constant given in equation (1):

$$\tau = 1/(k_{AB} + k_{BA}) \quad (1)$$

Furthermore, the rate constant k_{AB} can be written as equation (2):

$$k_{AB} = (1.0 - N_A)/\tau \quad (2)$$

Simple counting [14] can also be used to obtain the rates of barrier crossing. The rate is then equal to the number of events divided by the time of occupancy in the reactant well.

For the rate description to be valid a sufficiently high barrier is needed, *ie* greater that several kT [16]. Transitions between different wells in the system under study may be influenced by the conformational states of the other dihedral angles, but the effect is judged to be small (see *eg* the grid search, Figure 2). A number correlation function C(t) has been calculated for the dihedral angle ω of the *R*-isomer for the first 9 ns of simulation where transitions occur only between wells of the conformational states (g+) and (t). The exponential decay of C(t) for ω is observed from the graph (Figure 5) and a value obtained for the isomerization rate

Table 3. Transitions in dihedral angles.

	Number of transitions								
	ϕ			ψ			ω		
	$g^- \rightarrow g^+$	$g^- \rightarrow t$	$g^+ \rightarrow t$	$g^- \rightarrow g^+$	$g^- \rightarrow t$	$g^- \rightarrow t$	$g^- \rightarrow g^+$	$g^- \rightarrow t$	$g^+ \rightarrow t$
(*R*)-isomer									
1st run	414	0	0	0	188	274	2	1	81
2nd run	392	4	2	1	247	209	7	3	60
3nd run	402	3	3	0	198	270	12	5	77
Average	403	2	2	0	211	251	7	3	73
(*S*)-isomer									
1st run	264	0	0	0	545	40	4	0	33
2nd run	318	0	0	0	570	42	12	0	24
3rd run	224	0	0	0	575	30	6	0	24
Average	269	0	0	0	563	37	7	0	27

constant was obtained from ln C(t) and found to be:

$\tau(R, \omega) = 82$ ps

The following rate constants can then be deduced from equation (2):

$k(R, \omega, g+/t) = 4.2/ns$

$k(R, \omega, t/g+) = 8.2/ns$

By the simple counting procedure one obtains a higher rate than above as movement occurs over the dividing line of the surface but a real transition never takes place and the dihedral angle returns and remains in the well form where it started. By only counting transitions where the dihedral angle makes it into the adjacent well one obtains good agreement with the isomerization rate constant obtained from the number correlation function. During the 9 ns of simulation, 46 of these transitions took place between the states (g+) and (t). The following rate constants can then be obtained:

$k(R, \omega, g+/t) = 3.9/ns$

$k(R, \omega, t/g+) = 7.4/ns$

The isomerization rate constant from this counting is then obtained from equation (1).

$\tau(R, \omega) = 88$ ps

The statistical error in a relaxation time τ, can be calculated from a trajectory by equation (3) [17];

$$\text{err} = \tfrac{1}{2}(2\tau/T) \quad (3)$$

where T is the length of the simulation. The calculated error is then 13%. The isomerization rate constant obtained by the counting procedure differ by 7% to the one observed from the number correlation function. Thus, the obtained rate constant using the number correlation function or counting agree with each other, within statistical error.

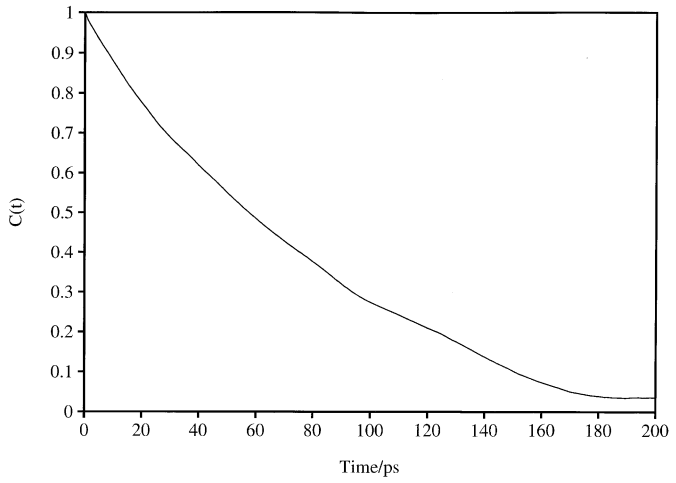


Figure 5. Number correlation function C(t) for the dihedral angle ω of the *R*-isomer.

Coupling constants

Coupling constants are a valuable tool in conformational analysis. The J value of two nuclei separated by three bonds can by a Karplus relationship give information on the dihedral angle between the atoms. Homonuclear proton–proton coupling constants have been used extensively but also in more recent time heteronuclear relationships.

We have measured the three-bond heteronuclear coupling constant between H2′ and C6. This has been performed by selective excitation of C6, 69.6 ppm (Figure 6a), followed by evolution of the long-range coupling, and polarization transfer to proton for detection. During aquisition of the

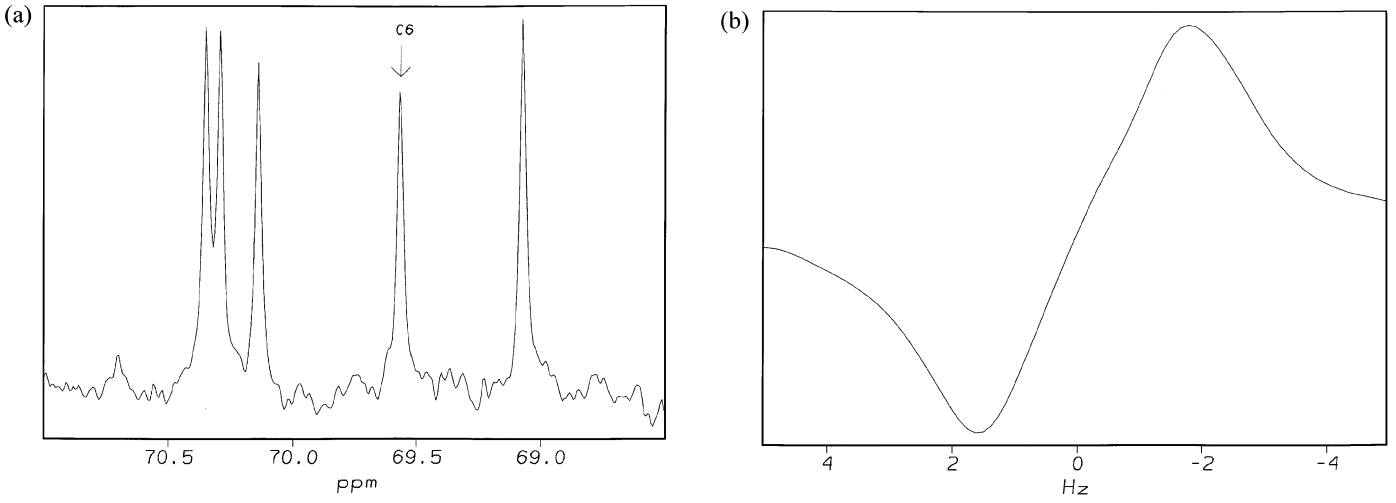


Figure 6. (a) Part of the ^{13}C NMR spectrum of the *S*-isomer. (b) ^{1}H detected NMR spectrum of the *S*-isomer showing the long range coupling between C6 and H2′.

FID, selective decoupling was performed on the methyl group, 1.33 ppm, to facilitate measurement of the J value for the H2′, 3.92 ppm, spin–spin coupled to C6. The long-range H2′,C6 coupling constant can be observed in anti-phase in the NMR spectrum (Figure 6b) and measured directly or extracted by the J-doubling procedure. The measured value using the J doubling technique is 3.3 Hz.

Homo- and heteronuclear coupling constants (Table 4) have been calculated from the LD simulations for the ϕ, ψ and ω dihedral angles. Differences in the populations of the ω angle are also reflected in the H5,H6R and H5,H6S J values being ca. 7.6/3.8 for the R-isomer and ca. 9.7/1.5 for the S-isomer. The heteronuclear coupling constants for the ψ dihedral angle are all around 3 Hz. The same is observed for the H2′,C6 J value of both isomers from the LD simulations.

A comparision of the simulated J value being 2.6 Hz for H2′,C6 to the experimental one of 3.3 Hz shows fair agreement although this parameter can only differentiate gauche conformers from the trans conformer which has a calculated J value of 6.8 Hz.

X-ray structure

Recently, the X-ray structure (Figure 1) of the R-isomer, in its protonated form, was determined [3]. The dihedral angles ϕ, ψ and ω had values of $-35.0°$, $170.7°$ and $76.5°$, respectively. This conformation can be defined as a (g−/t/g+) for the states of staggered conformers. On energy minimization in vacuo of the R-isomer having the ϕ, ψ and ω angles restrained at the dihedral angles obtained from the X-ray structure the energy difference is only 1.2 kcal mol^{-1} to the global energy minimum. Thus, the conformation of the X-ray structure in vacuo has an energy similar to those five low energy conformers identified from the grid search. This conformational state is also the one which has the highest single population (32%) during the LD simulations of the R-isomer. Although the crystal structure is in the protonated form it is of great interest that the molecule crystallizes into the most highly populated conformation, as determined from Langevin dynamics simulations.

Table 4. Calculated coupling constants.

Simulation	J_{H5H6R}	J_{H5H6S}	$J_{H6R\text{-}C2'}$	$H_{H6S\text{-}C2'}$	$J_{C6H2'}$
***(R)*-isomer**					
1st run	7.39	4.34	2.47	2.81	3.05
2nd run	7.99	3.33	2.92	2.97	3.21
3nd run	7.59	3.71	2.74	2.85	3.11
Average	7.66	3.79	2.71	2.88	3.12
(*S*)-isomer					
1st run	9.83	1.44	3.11	2.68	2.98
2nd run	9.47	1.67	3.21	2.78	2.01
3rd run	9.93	1.34	3.23	2.92	2.94
Average	9.74	1.48	3.18	2.79	2.64

Conclusion

The conformational space of methyl 6-*O*-[(*R*)- and (*S*)-1-carboxyethyl]-α-D-galactopyranoside has been investigated. Five low energy conformers for the R-isomer and two low energy conformers for the S-isomer have been identified from molecular mechanics energy minimizations and Langevin dynamics simulations. The conformational states differ for the ϕ dihedral angle where the major one is (g−) for the R-isomer and (g+) for the S-isomer. For both isomers the major population for the ψ dihedral angle is in the (t) state and for the ω dihedral angle the major conformation is in the (g+) state. Comparison of calculated $^3J_{C,H}$ values from LD simulations for H2′,C6 to that measured by NMR spectroscopy show fair agreement. The conformational state of the crystal structure can be identified as the same one as that of the single most highly populated one in the Langevin dynamics simulations.

Acknowledgements

This work has been supported by a grant from the Swedish Natural Science Research Council. We thank the Swedish NMR Centre for putting NMR facilities at our disposal.

References

1 Andersson M, Ratnayake S, Kenne L, Eriksson L, Stack RJ (1993) *Carbohydrate Res* **246**: 291.
2 Andersson M, Kenne L, Stenutz R, Widmalm G (1994) *Carbohydrate Res* **254**: 35.
3 Eriksson L, Pilotti Å, Stenutz R, Widmalm G (1996) *Acta Cryst Sect* **C52**: 2285.
4 Gasteiger J, Marsili M (1980) *Tetrahedron* **36**: 3219.
5 Venable RM, Pastor RW (1988) *Biopolymers* **27**: 1001
6 Nishida T, Widmalm G, Sandor P, (1996) *Magn Reson Chem* **34**: 377.
7 Nishida T, Widmalm G, Sandor P (1995) *Magn Reson Chem* **33**: 596.
8 McIntyre L, Freeman R (1992) *J Magn Reson* **96**: 425.
9 del Rio-Portilla F, Blechta V, Freeman R (1994) *J Magn Reson Ser A* **111**: 132.
10 Haasnoot CAG, de Leeuw FAAM, Altona C (1980) *Tetrahedron* **36**: 2783.
11 Tvaroska I, Hricovíni M, Petráková E (1989) *Carbohydrate Res* **189**: 359.
12 Widmalm G, Pastor RW (1992) *J Chem Soc Faraday Trans* **88**: 1747.
13 Hardy BJ, Egan W, Widmalm G (1995) *Int J Biol Macromol* **17**: 149.
14 Loncharich RJ, Brooks BR, Pastor RW (1992) *Biopolymers* **32**: 523.
15 Pastor RW, Karplus M (1989) *J Chem Phys* **91**: 211.
16 Zhou H-X (1989) *Chem Phys Lett* **164**: 285.
17 Zwanzig R, Ailawadi A (1969) *Phys Rev* **182**: 280.

Received 19 March 1996, revised 4 June 1996