Journal of Biomolecular NMR, **16:** 59–77, 2000.
KLUWER/ESCOM

Conformational analysis of the Sd^a determinant-containing tetrasaccharide and two mimics in aqueous solution by using ^{1}H NMR ROESY spectroscopy in combination with MD simulations

José L. Jiménez Blanco, Johannes J.M. van Rooijen, Paul J.A. Erbel, Bas R. Leeflang, Johannis P. Kamerling & Johannes F.G. Vliegenthart*
Department of Bio-Organic Chemistry, Bijvoet Center, Utrecht University, Padualaan 8, NL-3584 CH Utrecht, The Netherlands

Received 15 October 1999; Accepted 3 December 1999

Key words: full-relaxation matrix, molecular dynamics, oligosaccharide conformation, ROESY, Sd^a determinant, Tamm–Horsfall glycoprotein

Abstract

The conformational behaviour of the spacer-linked synthetic Sd^a tetrasaccharide β-D-Gal*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→O)$(CH_2)_5NH_2$ (**1**) and the two mimics β-D-Gal*p*-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→O)$(CH_2)_5NH_2$ (**2**) and β-D-Glc*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→O)$(CH_2)_5NH_2$ (**3**) were investigated by ^{1}H NMR spectroscopy in combination with molecular dynamics (MD) simulations in water. Experimental 2D ^{1}H ROESY cross-peak intensities (ROEs) of the tetrasaccharides were compared with calculated ROEs derived from MD trajectories using the CROSREL program. Analysis of these data indicated that the oligosaccharidic skeletons of the compounds **1**–**3** are rather rigid, especially the β-D-Hex(NAc)-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p* fragments. The α-Neu5-Ac-(2→3)-β-D-Gal*p* linkage occurred in two different energy minima in the three-dimensional structure of the compounds **1**–**3** in aqueous solution. Experimental data and dynamics simulations supported the finding that the higher energy rotamer (CHEAT forcefield) was abundant in compounds **1** and **3** due to the existence of a hydrogen bond between the carboxyl group of the sialic acid and the acetamido group of the terminal monosaccharide (GalNAc or GlcNAc) unit. The conformational similarity between **1** and **3** leads to the suggestion that also their activities will be alike.

Introduction

The Sd^a determinant β-D-Gal*p*NAc-(1→4)-[α-Neu5-Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→R) (Donald et al., 1983; Williams et al., 1984) was firstly discovered on human erythrocytes (Macvie et al., 1967) and consequently termed a blood antigen. However, Sd^a activity is also abundantly present in stomach, kidney, and colon tissues (Watkins, 1995 and references cited therein). A structurally related pentasaccharide, the Cad antigen β-D-Gal*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→3)-[α-Neu5Ac-(2→6)]-D-Gal*p*NAc, was isolated as alditol from glycophorin-A (Blanchard et al., 1983). Moreover, an Sd^a-active ganglioside, carrying the typical Sd^a determinant, was isolated from Cad erythrocytes (Gillard et al., 1988). The Sd^a determinant is typically found as a terminal sequence in N-glycans of human Tamm–Horsfall glycoprotein (TH-gp) (Williams et al., 1984; Hård et al., 1992; van Rooijen et al., 1998, 1999). With regard to the physiological function of TH-gp, its binding to neutrophils has been reported and a role as a specific ligand of neutrophil integrins has been suggested (Toma et al., 1994). This binding should facilitate neutrophil migration across renal epithelium, for example in the case of

*To whom correspondence should be addressed. E-mail: vlieg@pobox.uu.nl

tubulointerstitial nephritis. TH-gp was also reported to display immunosuppressive activity via its N-glycans (Muchmore and Decker, 1985; Dall' Olio et al., 1991; Sathyamoorthy et al., 1991). Furthermore, TH-gp has been suggested to be involved in the prevention of urinary tract and urinary bladder infection by inhibitory action of its glycans against the fimbriae-mediated adherence of *Escherichia coli* to uroepithelial cells (Duncan, 1988; Parkkinen et al., 1988).

In order to gain insight into the biological functioning of the Sd^a determinant at the molecular level and to predict biological activities of its mimics, the study of the conformational space explored by these glycans is essential. Several conformational analyses have already been performed on free oligosaccharides of the gangliosides GD1a, α-Neu5Ac-(2→3)-β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→O)-ceramide, GM1a, β-D-Gal*p*-(1→3)-β-D-Gal*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→O)-ceramide, and GM2, α-Neu5Ac-(2→3)-[β-D-Gal*p*NAc-(1→4)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*-(1→O)-ceramide, which have structures related to the Sd^a determinant. Sabesan et al. (1984, 1991) studied the conformational properties of these gangliosides using 1D and 2D NOESY experiments and molecular mechanics calculations (MM) and concluded that the glycosidic linkage α-Neu5Ac-(2→3)-β-D-Gal*p* for the branched sialic acid units was found to be extremely rigid. Other conformational studies of gangliosides and related oligosaccharides in non-aqueous solvents (Acquotti et al., 1990; Scarsdale et al., 1990; Levery, 1991) have shown that the presence of an α-Neu5Ac substituent at the linear oligosaccharide sequence can have a significant influence on their conformational structure. The structural perturbations in GM1a, GM2 and GD1a appeared to be stabilised by the formation of a hydrogen bond between the *N*-acetyl group of GalNAc and the carboxyl group of Neu5Ac (Acquotti et al., 1990; Scarsdale et al., 1990; Levery, 1991). Recently, some conformational studies of the GM1a oligosaccharide have been achieved in aqueous solution (Bernardi and Raimondi, 1995; Brocca et al., 1998). However, the search of conformations was performed using a homogenous dielectric environment that does not take into account the solvation effects of water molecules. So far, no conformational analyses on the Sd^a determinant or structurally related oligosaccharides (i.e. those of the gangliosides) using MD simulations with a non-continuum dielectric solvation model have been reported.

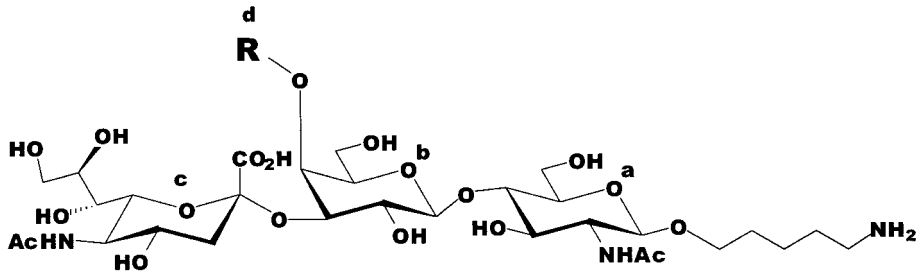
In this study we have analysed the three-dimensional structure of a synthetic Sd^a tetrasaccharide β-D-Gal*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→O)$(CH_2)_5NH_2$ (**1**), and the two mimics β-D-Gal*p*-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→O)$(CH_2)_5NH_2$ (**2**) and β-D-Glc*p*NAc-(1→4)-[α-Neu5Ac-(2→3)]-β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc-(1→O)$(CH_2)_5NH_2$ (**3**), using 2D ROESY ^1H NMR spectroscopy in combination with non-restrained MD simulations in a non-continuum water solvation model. The conformational analysis of these oligosaccharides can help to predict recognition phenomena with lectins or antibodies since, in many cases, the most stable conformation in free oligosaccharides is generally the one recognised by protein binding sites (Poveda and Jiménez-Barbero, 1998).

Materials and methods

The compounds **1**–**3** were synthesized (van Seeventer et al., 1997) (Figure 1) and provided by Dr. P.B. van Seeventer (Bijvoet Center, Utrecht University).

^1H and ^{13}C NMR spectroscopy

Compounds **1**–**3** were exchanged in ^2H_2O (99.9% ^2H, MSD Isotopes), with intermediate lyophilization, and finally dissolved in 500 μL ^2H_2O (99.96% ^2H, MSD Isotopes). 1D ^1H NMR and 2D ROESY experiments in ^1H_2O:^2H_2O (9:1, v/v) were performed at pH 6.0 using phosphate buffer (50 mM Na_2HPO_4/50 mM NaH_2PO_4 containing 0.1 mM NaN_3). 1D and 2D ^1H NMR measurements at 500 or 600 MHz were carried out on Bruker AMX-500 and AMX-600 spectrometers, respectively (Bijvoet Center, Department of NMR Spectroscopy, Utrecht University). ^1H chemical shifts (δ) were expressed in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured by reference to internal acetate (δ 1.908). Suppression of the $^1HO^2H$ signal for measurements in ^2H_2O was performed by presaturation during the relaxation delay for 1 s. For 1D ^1H NMR experiments in ^1H_2O:^2H_2O (9:1, v/v) of the tetrasaccharides, a 1D NOESY sequence was applied with a homo-spoil pulse of 10 ms and a consecutive recovery delay of 15 ms during the NOESY mixing time, and relaxation delays of 1, 5 and 10 s and variable temperatures were used in order to obtain the



Compound	R
1	β-D-GalNAc
2	β-D-Gal
3	β-D-GlcNAc

Figure 1. Synthetic analogue of the Sd^a tetrasaccharide determinant (**1**) and mimic compounds (**2** and **3**). Abbreviations used: **a**, β-D-GlcNAc; **b**, β-D-Gal; **c**, α-Neu5Ac; **d**, R.

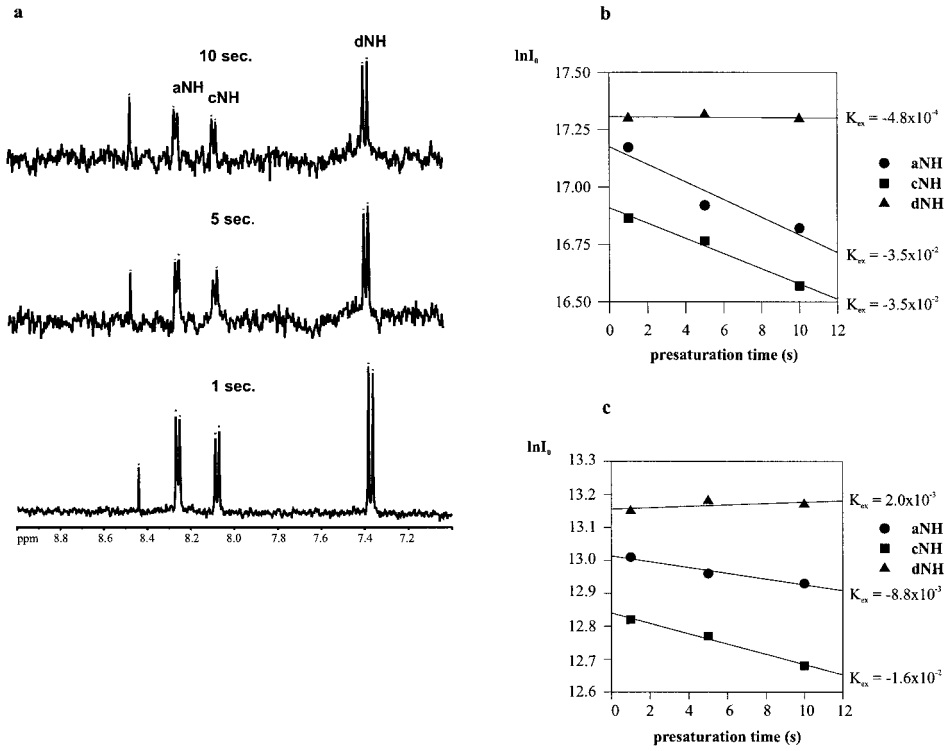


Figure 2. (a) 1D ^1H NMR spectrum of compound **1**, with presaturation times 1, 5 and 10 s, recorded in ^1H_2O-^2H_2O (9:1, v/v), pH 6.0 (NaH_2PO_4/Na_2HPO_4) at a probe temperature of 298 K. (b) Graphic representation of the logarithm of NH-proton intensities versus presaturation time for compound **1**. (c) Graphic representation of the logarithm of NH-proton intensities versus presaturation time for compound **3**. Abbreviations used are the same as for Figure 1.

exchange rates and temperature coefficients, respectively. All 2D spectra were processed on a Bruker Aspect Station using UXNMR, or on a Silicon Graphics (Indigo^2) Workstation using the local TRITON software (R. Kaptein, R. Boelens and J.A. van Kuik, Bijvoet Center, Utrecht University).

The 2D TOCSY experiments were recorded at 278 K using MLEV-17 mixing sequences (Bax and Davis, 1985a) at a field strength corresponding to 9.6–8.9 kHz (90° pulse width of 26–28 μs), with the offset frequency placed on the ^{1}HO^{2}H frequency position. The mixing times were 12 or 80 ms, and the spectra were recorded at a spectral width of 2000 Hz in each dimension.

The 2D ROESY (Bax and Davis, 1985b) spectra of **1**–**3** were recorded with a modified pulse sequence at 278 K. Spin-lock times of 75, 150, 225, or 300 ms were applied for **1**, and 25, 50, 80, 115, 145, or 200 ms for **2** and **3**, at a field strength of 9.4 kHz. The modified pulse sequence had three different carrier frequencies. Initially, the frequency was set to the water resonance for presaturation ($\delta_{presat.} \cong 4.76$ ppm). The presaturation time was 0.8 s. Next, the frequency was shifted to the centre of the spectral region ($\delta_{spectrum} \cong 3.1$ ppm). Prior to spin-locking, the frequency was moved off resonance (δ_{SL} = 12.6, 10.0 and 13.5 ppm for **1**, **2** and **3**, respectively). Since setting the frequency on this NMR spectrum region took a few milliseconds, a Δ-180°-Δ sequence was used to set the frequency and to refocus the magnetisation. Following the spin-lock pulse, a similar Δ-180°-Δ sequence was introduced in which the carrier frequency was set to $\delta_{spectrum} \sim 3.1$ ppm. The spectral width was ∼2200 Hz in each dimension. Peak intensities were determined by integration in a rectangular area around the peak maximum, or in the case of overlap, around the centre of the peak cluster.

^{1}H-^{1}H COSY (Bax and Freeman, 1981a, b) spectra of **1**–**3** were recorded at 278 K using a spectral width of 4032 Hz.

Natural abundance 1D ^{1}H-decoupled ^{13}C measurements were carried out on a Bruker AC-300 spectrometer operating at 75 MHz at 300 K. Chemical shifts (δ) are expressed in ppm downfield from external tetramethylsilane, but were actually measured with reference to internal acetone (δ 31.55).

A ^{1}H-^{13}C HMQC (Bodenhausen and Ruben, 1980) experiment was carried out at 500 MHz for the ^{1}H and at 126 MHz for the ^{13}C nuclei at 278 K, using spectral widths of 4032 and 15092 Hz, respectively. Phase-sensitive handling of data in the ω_1 direction was achieved by application of States-TPPI (Archer et al., 1992).

Molecular mechanics calculations

The molecular mechanics calculations were performed using the CHEAT forcefield (Grootenhuis and Haasnoot, 1988) in the INSIGHTII program (BIOSYM/Molecular Simulations, 1995, *InsightII 95.0 Molecular Modelling System*, San Diego, CA) on a Silicon Graphics Indigo^2 or O_2 computer. The CHEAT forcefield is based on the treatment of hydroxyl groups as extended atoms, which contain adjusted parameters for the aqueous state. This avoids the appearance of unrealistic intramolecular hydrogen bonds in vacuum and reduces the number of local energy minima, because the large number of possible hydroxyl group dihedral angles has disappeared. The forcefield was modified to have explicitly the hydrogen atom of the HO8 group of Neu5Ac, in view of the presence of an intramolecular hydrogen bond between HO8 and the carboxyl group of Neu5Ac (Acquotti et al., 1990; Poppe and van Halbeek, 1991). The atoms of the methyl, *N*-acetyl and carboxyl substituents, and the HO8 atoms of the Neu5Ac residues were implemented on the basis of the existing general valence forcefield (CHARMm).

For monosaccharide residues the co-ordinates from X-ray analyses were used. Only the dihedral angles of glycosidic bonds (ϕ and ψ) were variable and defined according to the IUPAC-IUB recommendations (IUPAC-IUB, 1983): $\phi = \theta(O_{5A}\text{-}C_{1A}\text{-}O_{XB}\text{-}C_{XB})$ and $\psi = \theta(C_{1A}\text{-}O_{XB}\text{-}C_{XB}\text{-}C_{(X-1)B})$. The glycosidic bond angle τ, (C_1-O_X-C_X), has fixed values of 117°. Solvation was not taken into account.

Iso-energy contour plots were calculated for glycosidic linkages by making steps of 5° in ϕ and ψ. The calculated iso-energy maps (French, 1989; Tran et al., 1989) are plotted at intervals of 1 kcal/mol with respect to the calculated global energy minimum.

Molecular dynamics

Molecular dynamics (MD) simulations were carried out on a Silicon Graphics Indigo^2 or O_2 computer using the GROMOS-87 program package (van Gunsteren, 1987). Positions of hydrogen atoms of the united CH, CH_2 and CH_3 atoms were calculated after simulation using 'ideal' tetrahedral geometry and a C-H bond length of 0.11 nm. Neutral charge groups were applied with a cutoff radius of 1.2 nm. The negatively charged carboxyl group was neutralised by assembling

a sodium cation. The SPC model (Berendsen et al., 1981) was used to describe the water molecules.

Simulations were carried out for ethyl glycoside analogues of **1**–**3**, placed in a truncated octahedron with a minimum distance of 1.2 nm to the edge of the box. Subsequently, the box was filled with water molecules (between 1100–1300 molecules). Periodic boundary conditions were applied together with a cutoff radius of 1.5 nm. All bond lengths were constrained and the water molecules were kept rigid using the SHAKE method (van Gunsteren and Berendsen, 1977).

Before the simulations were started, an energy minimisation was performed on the whole system. Initial velocities were obtained from Maxwellian distribution at 300 K and the simulation continued with steps of 0.002 ps. No constraints were applied in order to give the glycoside the opportunity to make conformational transitions. The simulations were performed at a constant pressure of 1 atm with a relaxation time (τ_p) of 0.5 ps and at a constant temperature of 300 K for compound **1** and 278 K of compounds **2** and **3**, by loose coupling of solvent and solute separately to temperature baths with relaxation times (τ_t) of 0.1 and 0.1 ps, respectively. An equilibrium of the total energy was reached after about 10 ps in each run. The first 10 ps were considered to be the equilibration time of the system and were not used for analysis.

For hydrogen bonding analysis, two criteria were applied to establish hydrogen bond formation: (1) a distance of less than 0.24 nm between the donor hydrogen atom and the acceptor, and (2) a bond angle of more than 120° (van Eijck and Kroon, 1989).

MD simulations (500 ps) were performed by starting with the lowest energy conformations at the β-D-Hex*p*(NAc)-(1→4)-β-D-Gal*p* and β-D-Gal*p*-(1→4)-β-D-Glc*p*NAc linkages as obtained from MM simulations minima (Table 1). MD simulations B,D,F correspond to a relative energy minimum in the α-Neu5Ac-(2→3)-β-D-Gal*p* linkage ($\phi,\psi = -32°,-142°$), where C1 of Neu5Ac adopts an anti arrangement with respect to Gal C3. In the MD simulations A,C,E, the α-Neu5Ac-(2→3)-β-D-Gal*p* linkage was situated on the global energy minimum ($\phi,\psi = 59°,-129°$) obtained from MM calculations.

Theoretical models of the ethyl glycoside conformations consisted of ^{1}H-^{1}H distance matrices, calculated by $r = <r^{-3}>^{-1/3}$ for each complete MD trajectory. The distance matrices of compounds **1** and **3** were formed by 44 aliphatic protons and the distance matrices of compound **2** were formed by 41 aliphatic protons.

ROESY spectra analysis

The theoretical basis of the ROESY spectra analysis applied in this report has been described by Lommerse et al. (1995a), including a treatment for TOCSY effects and for intramolecular flexibility by various calculation methods. Here, for the ROESY analysis Methods I and III are applied. Method I is based on the assumption that the molecule is in a single, rigid, average conformation, experiencing only one rotation correlation time (τ_0) describing the overall motion. Method III assumes different mobilities for each monosaccharide residue. However, it does not take into account fast internal motion within a monosaccharide ring. These fast internal motions are included in the effective local rotation correlation times ($\tau_{eff,k}$) of each monosaccharide residue and the interresidual vectors are described by $\tau_{eff,k}$ of the most mobile residue.

The experimental ROE data were transferred into matrices, in which each column ω_1 (ω_2 is constant) is scaled according to M_0-scaling (Leeflang and Kroon-Batenburg, 1992). The CROSREL program reconstructs ROE intensities from ethyl glycoside models obtained from MD simulations. These theoretical ROE intensities were compared with experimental data. Rotation correlation times (τ_0, $\tau_{eff,k}$) were determined for each theoretical model by fitting calculated intraresidual ^{1}H-^{1}H ROEs to the experiment. In general, those intraresidual ^{1}H-^{1}H pairs were taken, in which one of the protons is an anomeric proton. Furthermore, the intraresidual ^{1}H-^{1}H pairs were preferably chosen in a non-overlapping part of the spectrum, without the presence of strong TOCSY relay effects, and in particular without overlapping interresidual ^{1}H-^{1}H pairs. Overall rotation correlation times (τ_0, applied in Method I) were determined by a simultaneous fit of four intraresidual ROEs, one corresponding to each monosaccharide residue. The effective local rotation correlation times ($\tau_{eff,k}$, applied in Method III) were obtained for each linkage as described above.

The R_w value was used to evaluate the fit of a set of calculated ROEs to the experimental data. A ROE matrix was calculated of each obtained geometric model applying Methods I and III of the full relaxation rate matrix analysis and using effective correlation times of the models (Lommerse et al., 1995b).

Table 1. Initial conditions of the MD simulations A–F of the compounds **1**–**3**

Compound	**1**		**2**		**3**	
	A(ϕ,ψ)	B(ϕ,ψ)	C(ϕ,ψ)	D(ϕ,ψ)	E(ϕ,ψ)	F(ϕ,ψ)
GlcNAc-**a**-1-*O*-Ethyl	−64°, 180°	−64°, 180°	43°, 175°	−71°, 175°	−71°, 175°	−71°, 175°
Gal-**b**-(1→4)-GlcNAc-**a**	−63°, 123°	−63°, 123°	−63°, 123°	−110°, 110°	−63°, 123°	−63°, 123°
Neu5Ac-**c**-(2→3)-Gal-**b**	59°, −129°	−32°, −142°	59°, −129°	−32°, −142°	59°, −129°	−32°, −142°
R-**d**-(1→4)-Gal-**b**	−68°, −110°	−68°, −110°	−64°, −111°	−110°, −140°	−65°, −112°	−65°, −112°
Temp. (K)	300	300	278	278	278	278
Number of H_2O molecules	1237	1272	1126	1218	1271	1238

Results and discussion

The three-dimensional structures in water of the Sd[a] determinant **1** and two structurally related mimics **2** and **3** are investigated. Complete assignments of the proton and carbon chemical shifts were obtained by using different standard NMR techniques (1D ^{1}H and ^{13}C, and 2D TOCSY, ROESY and HMQC); the calculation of temperature coefficients and exchange rates of the NH protons is described in the *Proton and carbon chemical shift assignments* section. In the *Molecular dynamics trajectories* section, the MM calculations are described to obtain local energy minima at the glycosidic linkages of the compounds **1**–**3**. Starting from conformations at these different local minima, MD simulations of the compounds **1**–**3** have been carried out to obtain different trajectories. From these trajectories, information about intramolecular hydrogen bonding was extracted.

ROESY spectra of the three compounds at different mixing times have been recorded to construct ROE build-up curves that reflect the experimental distances between interresidual proton pairs. In the *ROE analysis by CROSREL* section, theoretical interresidual proton ROEs were calculated from these trajectories and compared with the experimental ones by applying the full-relaxation matrix theory with the CROSREL program.

Proton and carbon chemical shift assignments

The ^{1}H NMR chemical shifts and *J* couplings of **1**–**3** are listed in Table 2, and the 2D ROESY (150 ms) spectra of **1**–**3** are given in the Supplementary material. The results for **1** fully agree with the partial assignments of oligosaccharides structurally related to the Sd[a] determinant, such as gangliosides (Koerner et al., 1983; Sabesan et al., 1984; Acquotti et al., 1990; Sabesan et al., 1991) or an alditol derivative of the Sd[a] determinant (Donald and Feeney, 1986). Stereospecific assignments of H6pro*S* and H6pro*R* were performed by analogy with ^{1}H-NMR studies of stereospecifically deuterated D-hexoses (Nishida et al., 1987). The ^{1}H chemical shifts of the spacer were assigned by using homonuclear NMR techniques (COSY and TOCSY, Supplementary material). Of the methylene protons of the spacer, only those closest to the attachment to GlcNAc-**a** (–OCH_2–) are chemically non-equivalent. This is a consequence of the hindered rotation around the glycosidic bond due to the exoanomeric effect (Tvaroškova and Bleha, 1989).

To assign the H5, H6pro*S* and H6pro*R* protons of R-**d** and Gal-**b**, and the H8, H9pro*S* and H9pro*R* signals of Neu5Ac-**c**, 1D ^{13}C NMR measurements combined with HMQC experiments were carried out. The carbon chemical shifts (Supplementary material) agree with previous data for a ganglioside tetrasaccharide structurally related with compound **1** (Bock et al., 1984). The signals corresponding to the Neu5Ac-**c** H9a,9b protons are assigned as pro*S* and pro*R*, respectively, in accordance to a previous assignment of the Neu5Ac protons of GM1a (Acquotti et al., 1990). For the assignment of the two non-equivalent methylene protons (**a**-CHa,b) of the spacer, the same criteria were used as for the H6pro*R* and H6pro*S* atoms of β(1→6)-disaccharides described by Ohrui et al. (1985).

Comparison of the ^{1}H NMR data of **1**–**3** shows that the replacement of GalNAc-**d** (**1**) by Gal-**d** (**2**) has a strong influence on the resonance positions of the H1,2,3,4 of Gal-**b**. Furthermore, a clearly deviating set for Neu5Ac-**c** H3eq,3ax was found. Such remarkable shifts are not observed when replacing GalNAc-**d** (**1**) by GlcNAc-**d** (**3**) (van Seeventer et al., 1997).

The *J* couplings of Neu5Ac-**c** of compound **1**, $J_{6,7}$ = 2.1 Hz, $J_{7,8}$ = 10.1 Hz, $J_{8,9a}$ = 6.3 Hz and $J_{8,9b}$<1 Hz, indicate a lower flexibility of the glycerol

Table 2. ^{1}H chemical shifts and *J*-coupling constants of compounds **1**–**3**

Residue	GlcNAc-**a**	Gal-**b**	Neu5Ac-**c**	R-**d**
Compound 1				
H1	4.51 ($J_{1,2}$ 7.8 Hz)	4.55 ($J_{1,2}$ 9.8 Hz)	–	4.72 ($J_{1,2}$ 9.8 Hz)
H2	↑	3.35 ($J_{2,3}$ 7.8 Hz)	–	3.90 ($J_{2,3}$ 9.8 Hz)
H3ax	3.72	4.14 ($J_{3,4}$ 2.3 Hz)	1.92 ($J_{3ax,3eq}$ 13, $J_{3ax,4}$ 13 Hz)	3.67 ($J_{3,4}$ <1 Hz)
	↓			
H3eq	–	–	2.66 ($J_{3eq,4}$4.5 Hz)	–
H4	3.69 ($J_{4,5}$ 8.2 Hz)*	4.11 ($J_{4,5}$ <1 Hz)	3.78 ($J_{4,5}$ 10.4 Hz)*	3.91
H5	3.58 ($J_{5,6a}$ 4.2, $J_{5,6b}$ 6.1 Hz)*	3.77	3.83 ($J_{5,6}$ 11.8 Hz)	3.72
H6a (pro *S*)	3.99 ($J_{6a,6b}$12.6 Hz)	↑	3.49 ($J_{6,7}$ 2.1 Hz)	↑
H6b (pro *R*)	3.83	3.82	–	3.82
		↓		↓
H7	–	–	3.59 ($J_{7,8}$ 10.1 Hz)*	–
H8	–	–	3.75 ($J_{8,9b}$ <1 Hz)*	–
H9a (pro *S*)	–	–	3.89 ($J_{8,9a}$ 6.3 Hz)*	–
H9b (pro *R*)	–	–	3.60 ($J_{9a,9b}$ 12.1 Hz)*	–
Compound 2				
H1	4.51 ($J_{1,2}$ 7.9 Hz)	4.60 ($J_{1,2}$7.9 Hz)	–	4.73 ($J_{1,2}$ 7.9 Hz)
H2	↑	3.68 ($J_{2,3}$ 11.5 Hz)	–	3.48 ($J_{2,3}$ 8.8 Hz)
H3ax	3.70	4.22 ($J_{3,4}$ 2.3 Hz)	1.86 ($J_{3ax,3eq}$ 12, $J_{3ax,4}$ 12 Hz)	3.66
	↓			
H3eq	–	–	2.70 ($J_{3eq,4}$ 4.4 Hz)	–
H4	3.68 ($J_{4,5}$ 8.8 Hz)*	4.18 ($J_{4,5}$ <1 Hz)	3.68 ($J_{4,5}$ 11.5 Hz)*	3.89
H5	3.58 ($J_{5,6a}$ 3.7 Hz)*	3.75	3.84	3.73
H6a (pro *S*)	4.00 ($J_{6a,6b}$13.2 Hz)	↑	3.57	↑
H6b (pro *R*)	3.84	3.77	–	3.74
		↓		↓
H7	–	–	3.57	–
H8	–	–	3.83	–
H9a (pro *S*)	–	–	3.88	–
H9b (pro *R*)	–	–	3.60	–
Compound 3				
H1	4.50 ($J_{1,2}$ 7.8 Hz)	4.55 ($J_{1,2}$ 9.0 Hz)	–	4.78 ($J_{1,2}$ 9.4 Hz)
H2	↑	3.30 ($J_{2,3}$ 9.0 Hz)	–	3.71 ($J_{2,3}$ 7.7 Hz)*
H3ax	3.70	4.16	1.93 ($J_{3ax,3eq}$ 12.1, $J_{3ax,4}$12.1 Hz)	3.49
	↓			
H3eq	–	–	2.64 ($J_{3eq,4}$3.9 Hz)	–
H4	3.70	4.10 ($J_{4,5}$ <1 Hz)	3.75 ($J_{4,5}$ 7.8 Hz)*	3.46
H5	3.57 ($J_{4,5}$ 7.7 Hz)*	3.78	3.84	3.49
H6a (pro *S*)	3.99 ($J_{6a,6b}$11.2 Hz)	↑	3.46	3.89
H6b (pro *R*)	3.82 ($J_{5,6b}$ 7.7 Hz)	3.80	–	3.76
		↓		
H7	–	–	3.59	–
H8	–	–	3.73	–
H9a (pro *S*)	–	–	3.87	–
H9b (pro *R*)	–	–	3.59	–

**J*-couplings obtained from the COSY spectra.

Table 3. Temperature coefficients (ppm.K^{-1}) of NH-protons of compounds **1** and **3**

Compound	GlcNAc-**a**	Neu5Ac-**c**	R-**d**
1	0.0093	0.0074	0.0049
3	0.0085	0.0066	0.0039

side chain of Neu5Ac-**c**. This has been reported for various sialyl oligosaccharides as a consequence of the existence of a strong intramolecular hydrogen bond between the hydroxyl function at C8 and the carboxyl oxygens of sialic acid (Acquotti et al., 1990; Poppe et al., 1997). In compounds **2** and **3**, it was not possible to obtain the *J* couplings of the glycerol side chain in Neu5Ac using COSY experiments due to overlap of the signals.

To construct the experimental ROE build-up curves, the tracks of anomeric protons and the Gal-**b** H3,4 and Neu5Ac-**c** H3eq of **1**–**3** were chosen. In compounds **1** and **3**, the Gal-**b** H2 track was also included for the CROSREL analysis, whereas in compound **2**, the Gal-**d** H2 track was chosen. All these ROE peaks were scaled according to M_0-scaling. For all the compounds, the ROESY analysis is centered on the interresidual cross peaks R-**d** H1–Gal-**b** H3,4, Gal-**b** H1–GlcNAc-**a** H2,3,4 and Gal-**b** H3,4–Neu5Ac-**c** H3ax (Supplementary material). This last cross peak is more intense in compounds **1** and **3** than in compound **2**, indicating that for these compounds Neu5Ac-**c** and Gal-**b** must mainly adopt a conformation with anti-disposition between C1 of Neu5Ac-**c** and C3 of Gal-**b** (Poppe et al., 1989). The existence of a hydrogen bond between the NH of R-**d** residues and the carboxyl oxygens of Neu5Ac-**c** in the preferred solution conformations of **1** and **3** provides an adequate explanation of this phenomenon.

In order to obtain additional proof supporting the existence of intramolecular bonding between the NH of R-**d** residues and the carboxyl oxygens of Neu5Ac-**c** of compounds **1** and **3**, 1D ^{1}H NMR experiments in water were recorded at different temperatures (278, 288 and 298 K), to obtain the temperature coefficients, and also at different presaturation times (1, 5 and 10 s). First, a 2D ROESY experiment at 278 K (Hård et al., 1991) in water for compound **1** was carried out in order to assign the three NH signals observed in the 1D spectra. The peak at lowest field (δ ~8.40 ppm) was assigned to the NH-track of GlcNAc-**a** with H1, H3,4. The NH signals situated at high field (δ ~7.40 ppm) showed ROE cross peaks with H1, H2,4, H3,5 and NAc of GalNAc-**d**. Finally, the remaining NH signal at δ ~8.20 ppm was assigned to the NH of Neu5Ac-**c**, which only showed ROE peaks with H3ax. Assignments of NH protons of compound **3** were achieved by comparison with chemical shifts of the previously assigned NH peaks of compound **1**. The calculated temperature coefficients of the NH peaks of compounds **1** and **3** (Table 3) gave lower values for the NH of R-**d** than for the other NHs, indicating reduced solvent accessibility for the NH of the R-**d** residue. Similarly, a smaller decrease in the intensity of the NH signal of GalNAc-**d** with respect to the NH peaks of Neu5Ac-**c** and GlcNAc-**a** was observed at longer presaturation times (Figure 2). The plots of the ln I_0 for NH protons versus presaturation times (Figure 2b,c) show that the slope (exchange rate, K_{ex}) for the NH of R-**d** is closer to zero ($K_{ex} = -4.8 \times 10^{-4}$ and $K_{ex} = 2.0 \times 10^{-3}$ for compounds **1** and **3**, respectively) than for the NH-protons of the Neu5Ac-**c** and GlcNAc-**a** residues. These data support again the lower accessibility to solvent of the NH of the R-**d** unit. These results, together with the chemical shift and ROE data, confirm that this NH group is likely implicated in an intramolecular hydrogen bond with the carboxylic group of Neu5Ac-**c**.

Molecular dynamics trajectories

MD simulations starting from energy minima conformations obtained by MM calculations in the CHEAT forcefield (Table 1, A–F) have been performed for the compounds **1**–**3** in order to investigate time-averaged conformations and the flexibility within the oligosaccharides in terms of transitions at the glycosidic linkages. These starting conformations are chosen according to the ROESY data obtained. The ϕ,ψ dihedral angles of each glycosidic linkage of the MD trajectories are plotted in Figures 3–5. These plots show that the conformations are mainly around the starting energy minima from the MM calculations. Average dihedral angles of the complete trajectories are listed in Table 4. Hydrogen bonds which appeared for more than 10% of the total simulated MD trajectories are listed in Table 5.

ROE analysis by CROSREL

Theoretical ROE intensities were calculated from the MD simulation models applying Method I (assuming isotropic motion) and Method III (assuming anisotropic motion) and compared with experimental

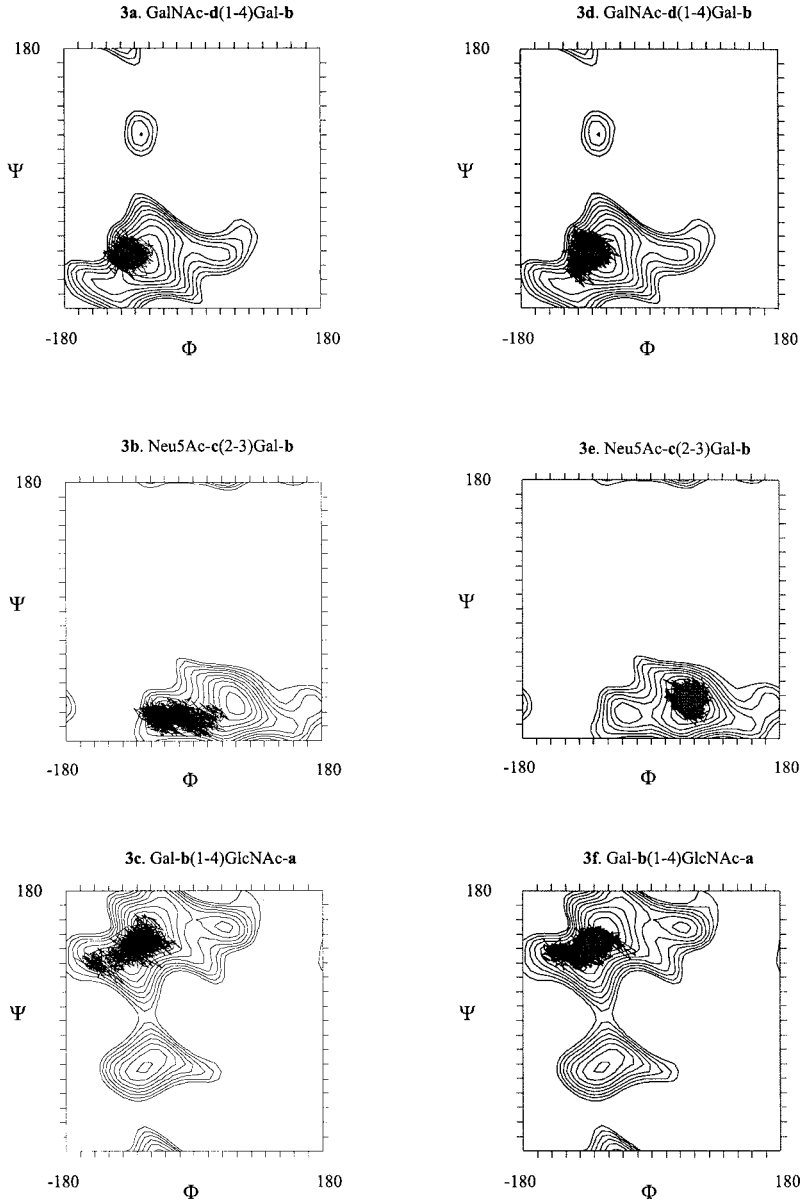


Figure 3. Plots in Φ,Ψ surfaces of the various glycosidic linkages for compound **1**. For each glycosidic linkage has been plotted: (1) Iso-energy contour levels as obtained by CHEAT calculations at intervals of 1 kcal/mol from the global minimum; (2) Φ,Ψ dihedral angle pairs of the A (**3a**–**c**) and B (**3d**–**e**) simulations.

ROEs from ROESY experiments, using the CROSREL program (Leeflang and Kroon-Batenburg, 1992). The intraresidual ^{1}H-^{1}H pairs, which were used to fit rotation correlation times, are given in Table 6. The resulting overall (τ_0) and local effective correlation times ($\tau_{eff,k}$), as obtained by Methods I and III, respectively (Lommerse et al., 1995a), are presented in Table 6. In general, the $\tau_{eff,k}$ values of Gal-**b** situated on the mass center of the molecules **1**–**3** are longer than the $\tau_{eff,k}$ values of the rest of the residues, in agreement with the previous results of Lommerse et al. (1995a, b); see Table 6.

The R_w values of the intraresidual (Table 7) and interresidual (Table 8) ROE ^{1}H-^{1}H pairs indicate the accuracy of the calculated ROE intensities by the used methods. The distances between intraresidual protons depend on monosaccharide ring conformations. For this reason, the intraresidual ^{1}H-^{1}H ROE inten-

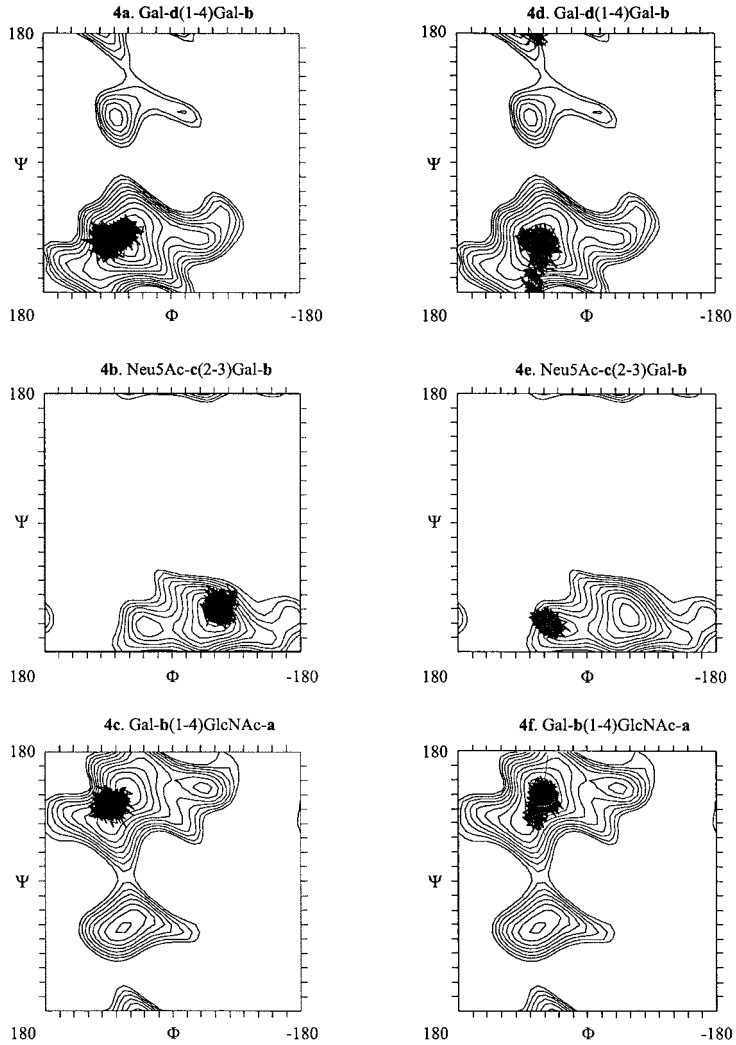


Figure 4. Plots in Φ,Ψ surfaces of the various glycosidic linkages for compound **2**. For each glycosidic linkage has been plotted: (1) Iso-energy contour levels as obtained by CHEAT calculations at intervals of 1 kcal/mol from the global minimum; (2) Φ,Ψ dihedral angle pairs of the C (**4a**–**c**) and D (**4d**–**e**) simulations.

sities will be approximately the same for the various theoretical models, resulting in similar intraresidual R_w values. Only those interresidual ^{1}H-^{1}H pairs are analysed that could be measured reliably (Table 8). Comparison of the theoretical and experimental interresidual ROE peaks is achieved by using build-up curves (Figures 6–8) and the R_w values (Table 8).

Theoretical ROE intensities from simulation B have a better fit than those from A to the experimental ones for compound **1**. In Table 8 it is shown that the R_w values of B are shorter than those of A. The build-up curves 6d–f fit better to the experimental ROEs than 6a–c (Figure 6). However, the application of Method III (anisotropic assumption) does not translate into an improvement of the fitting between experimental and theoretical interresidual ROE intensities as for bromelain (Lommerse et al., 1995b). This suggests that the conformational behaviour of compound **1** is rather isotropic. Also, these build-up curves and R_w values indicate that the β-D-Gal*p*NAc-**d**-(1→4)-β-D-Gal*p*-**b** and the β-D-Gal*p*-**b**-(1→4)-β-D-Glc*p*NAc-**a** glycosidic linkages adopt conformations around the corresponding global energy minimum (ϕ: $-93\pm\Delta15°$, ψ: $-103\pm\Delta15°$; ϕ: $-80\pm\Delta20°$, ψ: $108\pm\Delta20°$, respectively; Figure 3d,f), whereas those of α-Neu5Ac-**c**-(2→3)-β-D-

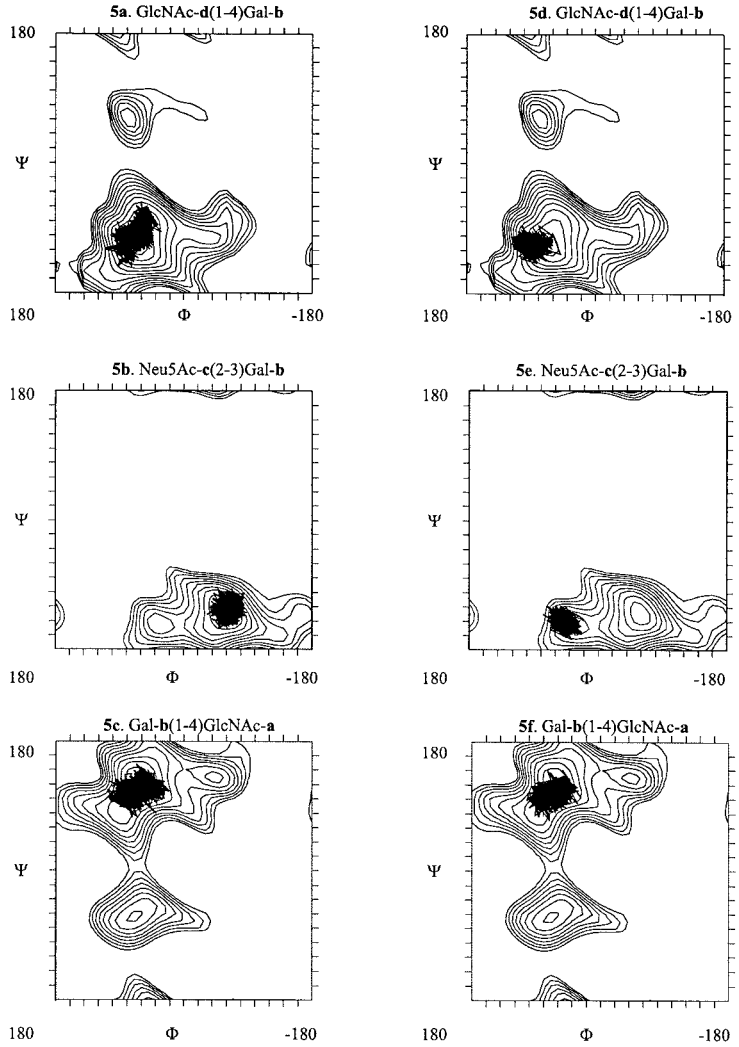


Figure 5. Plots in Φ,Ψ surfaces of the various glycosidic linkages for compound **3**. For each glycosidic linkage has been plotted: (1) Iso-energy contour levels as obtained by CHEAT calculations at intervals of 1 kcal/mol from the global minimum; (2) Φ,Ψ dihedral angle pairs of the E (**5a**–**c**) and F (**5d**–**e**) simulations.

Gal*p*-**b** are situated on a relative energy minimum (ϕ: $-29\pm\Delta15°$, ψ: $-150\pm\Delta15°$; Figure 3e). The dihedral angles predicted for β-D-Gal*p*NAc-**d**-(1→4)-β-D-Gal*p*-**b** and α-Neu5Ac-**c**-(2→3)-β-D-Gal*p*-**b** are in agreement with the previous ones obtained for GM1 gangliosides (Sabesan et al., 1984, 1991; Acquotti et al., 1990; Levery, 1991). For the β-D-Gal*p*-**b**-(1→4)-β-D-Glc*p*NAc-**a** bond, the average dihedral angles are closer to the values predicted in the SLe^x determinant by Rutherford et al. (1994). In Figure 9a it is shown that the averaged model B adopts an anti conformation at the α-Neu5Ac-**c**-(2→3)-β-D-Gal*p*-**b** bond similar to that proposed by Sabesan et al. (1991) and Poppe et al. (1989). This conformation approaches the NH of GalNAc-**d** and the carboxyl oxygens of Neu5Ac-**c** (∼2.5 Å) for hydrogen bonding (Table 5, ∼50–75% of the simulated time).

For compound **2**, the interresidual R_w values (Table 8) and the build-up curves 7a–f (Figure 7) show that each of the C and D simulations alone cannot describe satisfactorily its conformational equilibrium in water. Nevertheless, the CROSREL program allows the combination of the trajectories C and D in a convenient percentage in order to obtain a new model C/D that better fits the theoretical interresidual peaks to the experimental ones (Table 8, Figure 7g–i). Similarly to compound **1**, the application of the $\tau_{eff,k}$ values does not improve significantly the fitting between experi-

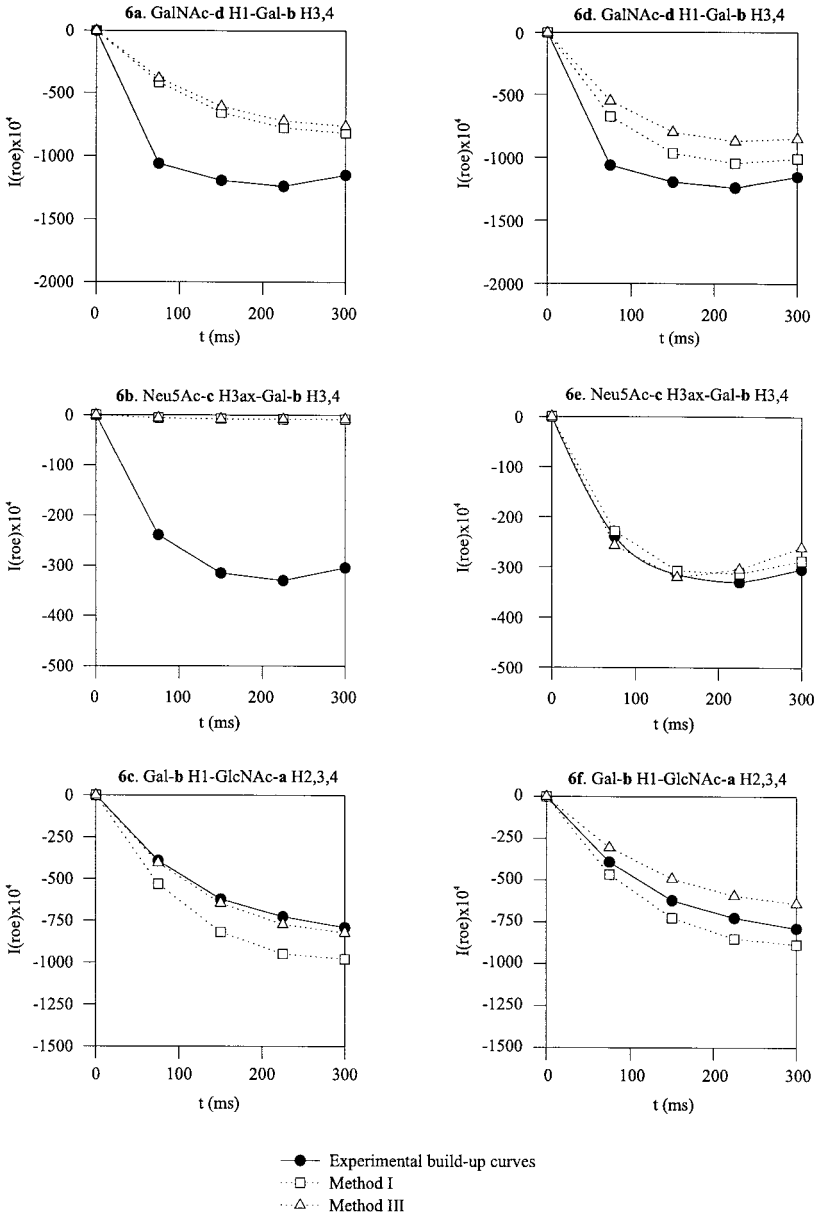


Figure 6. ROE build-up curves of interresidual ^{1}H-^{1}H pairs within compound **1**, both experimental and calculated for the A (**6a**–**c**) and B (**6d**–**f**) simulations using the CROSREL program.

mental and theoretical interresidual peaks (Figure 7), which again suggests an isotropic behaviour of the conformation in **2**. This C/D model is a combination of 85% of conformer C and 15% of conformer D and all the trajectories of the β-D-Gal*p*-**d**-(1→4)-β-D-Gal*p*-**b** and β-D-Gal*p*-**b**-(1→4)-β-D-Glc*p*NAc-**a** glycosidic linkages are mainly situated on the global energy minimum regions (Figure 4a,c,d,f). In the α-Neu5Ac-**c**-(2→3)-β-D-Gal*p*-**b** bond, the trajectories in the global energy minimum (ϕ: 60±Δ15°, ψ: −120±Δ15°; Figure 4b) prevail over those situated on the relative energy minimum (ϕ: −56±Δ15°, ψ: −120±Δ15°; Figure 4e). This energy minimum conformation has been observed in conformational studies of GM1a (Bernardi and Raimondi, 1995) and GM4 gangliosides (Poppe et al., 1989). Recently, a conformational analysis of Neu5Ac-(2→3)-β-D-Gal*p*-(1→4)-β-D-Glc*p* carried out by Milton et al. (1998)

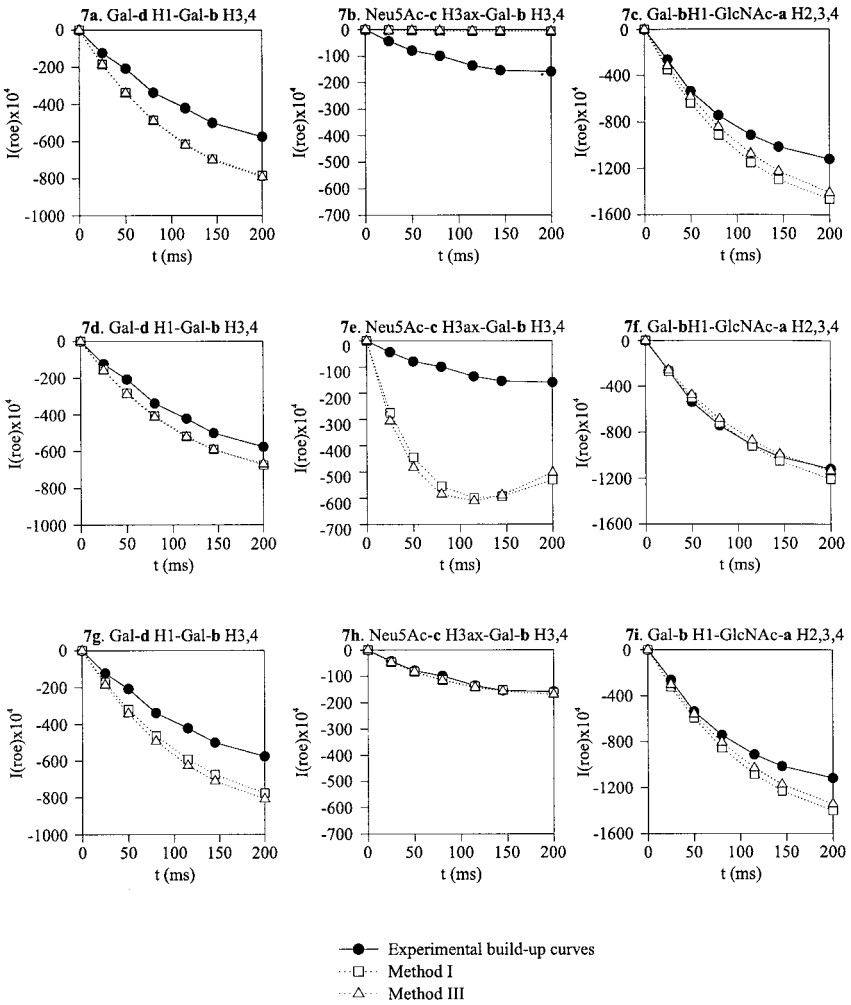


Figure 7. ROE build-up curves of interresidual ^{1}H-^{1}H pairs within compound **2**, both experimental and calculated for the C (**7a**–**c**), D (**7d**–**f**) simulations and a combination of two trajectories C/D (**7g**–**i**) using the CROSREL program.

Table 4. Average dihedral angles of the MD simulations for compounds **1**–**3**

Linkage	Gal-**b**(1→4)GlcNAc-**a** ϕ,ψ	Neu5Ac-**c**(2→3)Gal-**b** ϕ,ψ	R-**d**(1→4)Gal-**b** ϕ,ψ
Simulation	Compound **1**		
A	−88°,105°	55°,−130°	−80°,−103°
B	−80°,108°	−29°,−150°	−93°,−103°
	Compound **2**		
C	−83°,111°	60°,−120°	−85°,−101°
D	−62°,112°	−56°,−140°	−66°, −117°
	Compound **3**		
E	−63°,114°	64°,−121°	−67°,103°
F	−48°,−106°	−46°,−130°	−77°,105°

Table 5. Presence of hydrogen bonds* (% of the simulated trajectory) during the different long MD trajectories of compounds **1**–**3**

Hydrogen bond		MD simulations					
Donor	···Acceptor	A	B	C	D	E	F
R-**d** (NH/OH2)[a]	···Neu5Ac-**c** OC=O	–	50.4	–	18.0	–	83.8
R-**d** (NH/OH2)[b]	···Neu5Ac-**c** OC=O	–	25.0	–	–	–	–
R-**d** OH6	···Neu5Ac-**c** OC=O	–	–	–	–	11.8	–
Neu5Ac-**c** OH8	···Neu5Ac-**c** OC=O	22.6	15.2	–	53.9	–	94.0
Neu5Ac-**c** OH8	···Neu5Ac-**c** O6	–	–	–	25.3	–	27.6
Neu5Ac-**c** OH7	···Neu5Ac-**c** O9	–	17.8	–	–	12.0	–
Gal-**b** OH2	···GlcNAc-**a** O6	–	19.1	–	–	–	–
Gal-**b** OH2	···Neu5Ac-**c** O6	–	–	38.8	–	21.9	–
GlcNAc-**a** OH6	···Gal-**b** O2	–	–	27.7	–	–	–
GlcNAc-**a** OH3	···Gal-**b** O5	13.4	15.9	41.6	–	–	10.0

[a,b]Compounds **1** and **3** have NH in position 2 of R-**d**, compound **2** has OH in position 2 of R-**d**.
*Only hydrogen bonds are given which were formed for more than 10% of the complete MD trajectory.

Table 6. Rotation correlation times (ns) for the monosaccharide residues and the overall rotation correlation time of compounds **1**–**3** using different calculation methods

Simulation	Method I (τ_0)		Method III ($\tau_{eff,k}$)		
	Overall	GlcNAc-**a** H1-H3*	Gal-**b** H1-H5*	Neu5Ac-**c** H3ax-H3eq	GalNAc-**d** H1-H3*
A	0.43	0.23	0.64	0.53	0.35
B	0.56	0.23	0.96	0.79	0.34
	Overall	GlcNAc-**a** H1-H3*	Gal-**b** H1-H3*	Neu5Ac-**c** H3ax-H3eq	Gal-**d** H1-H3*
C	0.79	0.62	0.80	0.57	0.92
D	0.72	0.45	0.89	0.87	0.76
C/D	0.73	0.59	0.82	0.62	0.90
	Overall	GlcNAc-**a** H1-H3*	Gal-**b** H1-H3*	Neu5Ac-**c** H3ax-H3eq	GlcNAc-**d** H1-H3*
E	0.27	0.18	1.33	0.25	1.12
F	0.24	0.12	1.95	0.22	2.08
E/F	0.25	0.15	1.59	0.24	1.52

*These intraresidual ^{1}H-^{1}H pairs overlap with other cross peaks in the ROESY spectrum.

Table 7. Total R_w values for the intraresidual ROEs using various calculation methods for compounds **1**–**3**

Method	Simulation							
	A	B	C	D	C/D	E	F	E/F
I	0.162	0.187	0.142	0.163	0.106	0.254	0.354	0.388
III	0.148	0.120	0.110	0.125	0.091	0.229	0.269	0.243

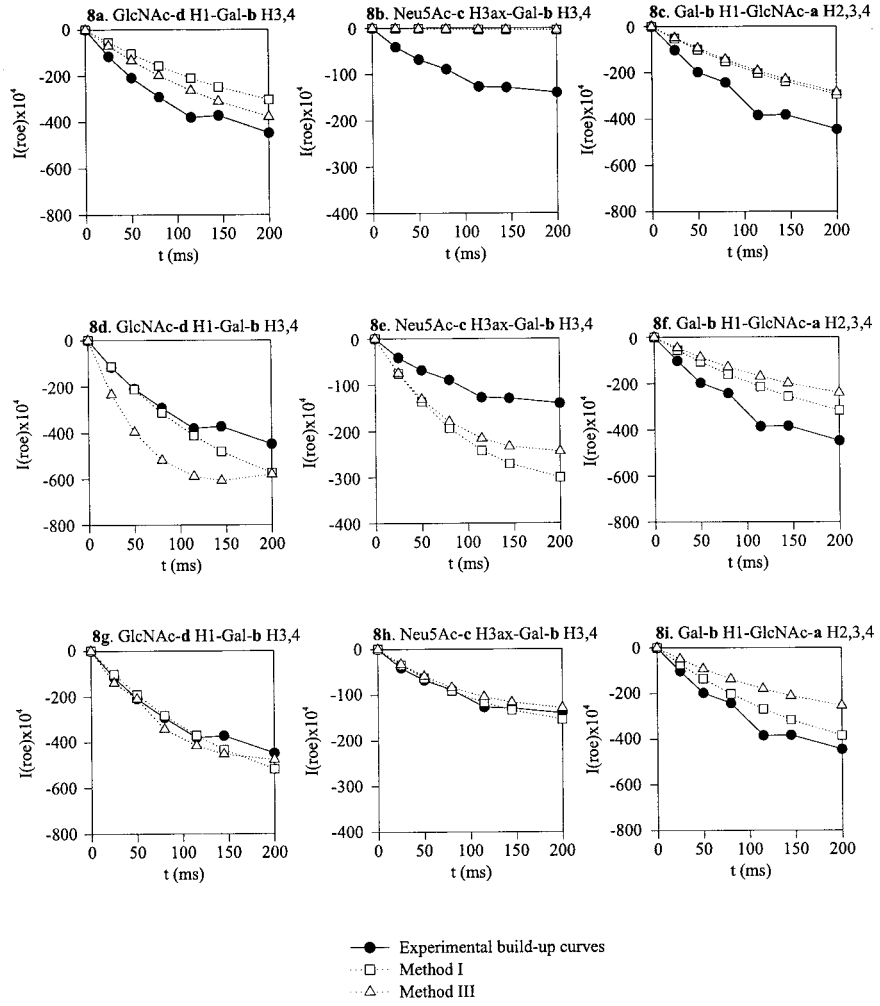


Figure 8. ROE build-up curves of interresidual ^{1}H-^{1}H pairs within compound **3**, both experimental and calculated for the E (**8a–c**), F (**8d–f**) simulations and a combination of two trajectories E/F (**8g–i**) using the CROSREL program.

also revealed the existence of a family of conformations around $\phi,\psi = 57°,-112°$. In Figure 9b it is shown that in the more predominant conformation C, the OH2 of Gal-**d** and the carboxyl oxygens of Neu5Ac-**c** are distant in space. The predominance of conformer C suggests that the hydrogen bond between the OH2 of Gal-**d** and the carboxyl oxygens of Neu5Ac-**c** observed during simulation D (Table 5) is not effective enough to shift the conformational equilibrium towards anti-conformer D.

Finally, the interresidual R_w values (Table 8) and build-up curves (Figure 8a–f) of compound **3** obtained from its E and F simulations indicate that an individual MD model is not suited to describe the conformational equilibrium of this molecule in aqueous solution. Again, the combination of the trajectories E/F fits quite well the experimental interresidual peaks (Figure 8g–i). However, in this case, the E and F models take part equally in the three-dimensional structure of **3** (E:F = 0.55:0.45). As for the previous compounds **1** and **2**, the application of Method III to calculate theoretical ROE peaks from the trajectories does not improve their fitting to the experimental ones, implying an isotropic tumbling of compound **3**. From the build-up curves (Figures 8a,d,g and 8c,f,i), it is deduced that this molecule adopts conformations around the global energy minima in the β-D-Glc*p*NAc-**d**-(1→4)-β-D-Gal*p*-**b** and β-D-Gal*p*-**b**-(1→4)-β-D-Glc*p*NAc-**a** linkages (See Figures 5a,d and 5c,f, respectively). In the α-Neu5Ac-**c**-(2→3)-β-D-Gal*p*-**b** bond, the inter-

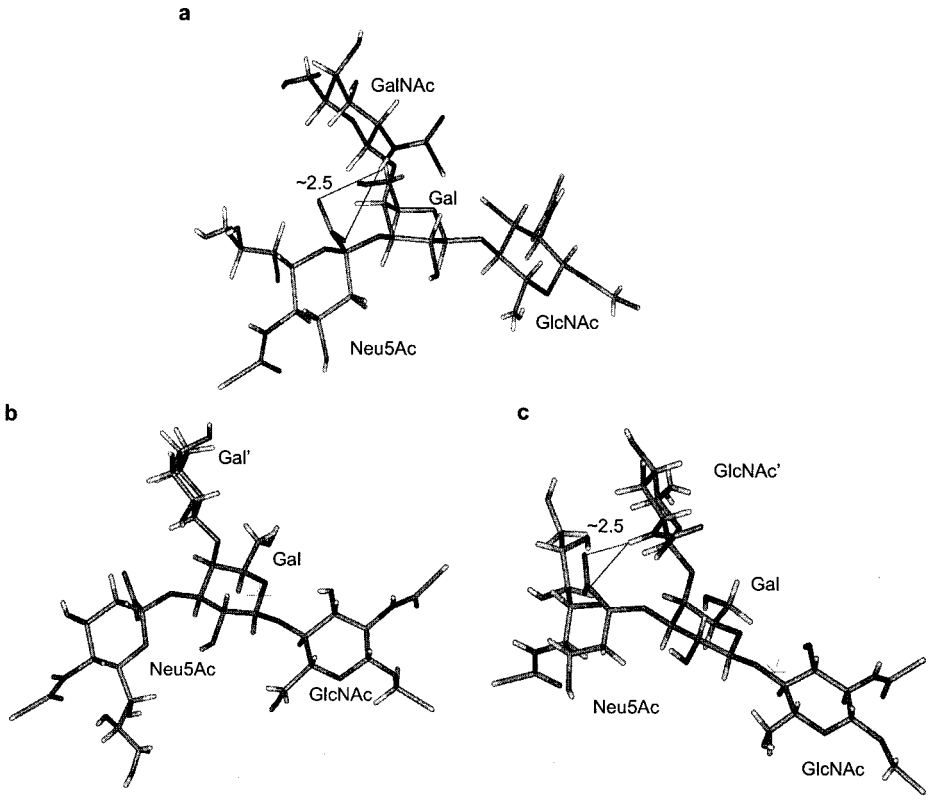


Figure 9. (a) Average simulated model from the B MD simulation of compound **1**. (b) Average simulated model from the C MD simulation of compound **2**. (c) Average simulated model from the F MD simulation of compound **3**.

residual theoretical ^{1}H-^{1}H pairs only show a correct fitting for the model E/F (Figure 8h). This means that the two local energy minima (ϕ: $64\pm\Delta15°$, ψ: $-121\pm\Delta15°$; Figure 5b and ϕ: $-46\pm\Delta15°$, ψ: $-130\pm\Delta15°$; Figure 5e) have an important presence in the conformational equilibrium of compound **3**. The trajectories situated on the relative energy minimum (Figure 5e) approach the NHAc group of GlcNAc-**d** and the carboxylic group of Neu5Ac-**c** (Figure 9c), favouring an intramolecular hydrogen bond (Table 5).

The hydrogen bond analysis of the simulations also predicts the existence of a hydrogen bond between the OH8 and the carboxyl groups of Neu5Ac-**c** in the A, B, D and F simulations (Table 5), in accordance with Acquotti et al. (1990), and Poppe and van Halbeek (1991).

Conclusions

The present report shows that the combination of MD simulations in water and NMR spectroscopy is a powerful tool to describe the conformational behaviour of oligosaccharide compounds. The application of this method on the Sd^{a} determinant, especially in combination with two mimics, provides a detailed model of the structure of the Sd^{a} determinant and the essential structural elements of this tetrasaccharide for biological functioning. The presence of the NHAc group in the terminal monosaccharide R-**d** of compounds **1** and **3** plays a significant role in the three-dimensional conformation adopted by these oligosaccharides in aqueous solution. Based on this MD study and the NMR data an intramolecular hydrogen bond R-**d**-NHAc$\cdots$Neu5Ac-**c**-OCO (presence 50–75%) favours the anti conformer in the α-Neu5Ac-**c**-(2$\rightarrow$3)-β-D-Gal*p*-**b** bond. A similar hydrogen bonding was suggested in some studies on the conformational properties of gangliosides (Acquotti et al., 1990; Scarsdale et al., 1990; Levery, 1991). In mimic **2**, where this NHAc group is absent, only $\sim$15% of the time the conformation resembles the Sd^{a} determinant, stabilised by a weak hydrogen bond between Gal-**d**-OH2$\cdots$Neu5Ac-**c**-OCO (presence 18%). And $\sim$85%

Table 8. Total R_w values* for the interresidual ROEs using various calculation methods for compounds **1**–**3**

Simulation	Overall	Gal-**b** H1-GlcNAc-**a** H2,3,4	Neu5Ac-**c** H3ax-Gal-**b** H3,4	GalNAc-**d** H1-Gal-**b** H3,4
	Method I			
A	0.543	+0.298	−0.975	−0.440
B	0.254	+0.160	−0.045	−0.225
	Method III			
A	0.570	+/−0.048	−0.975	−0.479
B	0.373	−0.197	+/−0.090	−0.349
	Overall	Gal-**b** H1-GlcNAc-**a** H2,3,4	Neu5Ac-**c** H3ax-Gal-**b** H3,4	Gal-**d** H1-Gal-**b** H3,4
	Method I			
C	0.421	+0.272	−0.959	+0.436
D	1.151	+/−0.051	+3.532	+0.216
C/D	0.237	+ 0.204	+/−0.069	+0.384
	Method III			
C	0.424	+0.161	−0.960	+0.446
D	1.203	+/−0.055	+3.701	+0.225
C/D	0.266	+0.148	+/−0.085	+0.461
	Overall	Gal-**b** H1-GlcNAc-**a** H2,3,4	Neu5Ac-**c** H3ax-Gal-**b** H3,4	GlcNAc-**d** H1-Gal-**b** H3,4
	Method I			
E	0.566	−0.401	−0.972	−0.403
F	0.580	−0.365	+1.115	+/−0.209
E/F	0.388	−0.385	+/−0.096	−0.171
	Method III			
E	0.566	−0.465	−0.975	−0.271
F	0.627	−0.503	+0.820	+0.638
E/F	0.425	−0.475	+/−0.122	+0.154

*Minus, plus or plus/minus shows if the calculated ROE build-up curve is smaller, larger or goes through the experimental curve, respectively.

of the time, the α-Neu5Ac-**c**-(2→3)-β-D-Gal*p*-**b** linkage is in the syn conformation, further stabilised by the hydrogen bond between Gal-**b**-OH2···Neu5Ac-**c**-OCO (presence 38.8%). Mimic **3** is found in an intermediate situation. In 45% of the simulation time, the higher energy anti conformer, similar to the native Sd^a determinant, is stabilised by the hydrogen bond GlcNAc-**d**-NH···Neu5Ac-**c**-OCO (presence 84%). However, there is a strong competition with the syn conformation, similar to mimic **2**, further stabilised by a hydrogen bond between GlcNAc-**d**-OH6···Neu5Ac-**c**-OCO (presence 12%). The difference of the population distributions of compounds **1** and **3** can be explained taking into account the competition between the hydrogen bonds GlcNAc-**d**-NH···Neu5Ac-**c**-OCO of conformer F and GlcNAc-**d**-OH6···Neu5Ac-**c**-OCO of conformer E (Table 5) in compound **3**. For β-D-GlcNAc, it is known that the rotamer tg at the C5–C6 bond is very unfavoured (Nishida et al., 1987) because of the synperiplanar repulsion between the OH4 and OH6 bonds (Perkin et al., 1977). Thus, the hydroxymethylene group of the GlcNAc-**d** residue should mainly adopt gg and gt dispositions. The latter would favour the presence of the GlcNAc-**d**-OH6···Neu5Ac-**c**-OCO hydrogen bond in the lower energy conformer E. Therefore, the competition between the two intramolecular hydrogen bonds could explain the almost equal presence of the two conformers E and F in the conformational equilibrium of compound **3**. In the case of the GalNAc-**d** residue of compound **1**, the C5–C6 bond is more flexible, since there is no synperiplanar repulsion. This explains the absence of the GalNAc-**d**-OH6···Neu5Ac-**c** in the A and B simulations. Thus, the presence of a unique hydrogen bond between the GalNAc-**d** and

Neu5Ac-**c** residues (Table 5) shifts the conformational equilibrium of **1** to the B conformation.

From the ROE build-up curves (Figures 6-8) and interresidual R_w values (Table 8) of compounds **1**–**3** it can be deduced that the conformational behaviour of these molecules is rather isotropic. This is in agreement with the observation that a generally better fitting of the calculated interresidual ROE peaks to the experimental ones was obtained assuming isotropic tumbling of the molecule (Method I). This means that the mobility of all the glycosidic linkages is similar. However, the α-Neu5Ac-**c**-(2→3)-β-D-Gal*p*-**b** linkage of compound **3** shows very short $\tau_{eff,k}$ values (0.22–0.25 ns). This is in agreement with fast transitions between the two minimum energy regions (Figures 5b,e) predicted for this compound.

The conformational resemblance of **1** and **3** suggests that mimic **3** could possibly display similar biological activity as Sd^a tetrasaccharide **1**. Further binding studies with lectins and antibodies of these structure-related tetrasaccharides to the Sd^a determinant are in progress in order to prove these predictions. It is evident that knowledge of the relation between three-dimensional structure and binding properties of these biomolecules could assist in the design of new carbohydrate-based therapeutic agents.

Acknowledgements

This investigation was supported by the European Commission DG XII-E. We thank Dr. P.B. van Seeventer for providing compounds **1**–**3**. We thank the authors of GROMOS for the use of their program package, and Dr. J.A. van Kuik, Dr. L.M.J. Kroon-Batenburg and Dr. P.H. Kruiskamp for their assistance in using this program. J.L.J.B. thanks the European Commission DG-XII-E (Biotechnology programme) for a postdoctoral grant (BIO4CT975071).

Supplementary material

The following material can be obtained from the authors upon request: Tables of ^{13}C chemical shifts of compounds **1**–**3** and ^{1}H chemical shifts of the aminopentyl spacer of compounds **1**–**3**, and 2D ^{1}H-^{1}H ROESY spectra, with a mixing time of 150 ms, of compounds **1**–**3** recorded in $^{2}H_2O$ measured at a probe temperature of 278 K with the assignments of the tracks of protons used for CROSREL analysis.

References

Acquotti, D., Poppe, L., Dabrowski, J., von der Lieth, C.-W., Sonnino, S. and Tettamanti, G. (1990) *J. Am. Chem. Soc.*, **112**, 7772–7778.

Archer, S.J., Baldisseri, D.M. and Torchia, D.A. (1992) *J. Magn. Reson.*, **97**, 602–606.

Bax, A. and Davis, D.G. (1985a) *J. Magn. Reson.*, **65**, 355–360.

Bax, A. and Davis, D.G. (1985b) *J. Magn. Reson.*, **63**, 207–213.

Bax, A. and Freeman, R. (1981a) *J. Magn. Reson.*, **42**, 164–168.

Bax, A. and Freeman, R. (1981b) *J. Magn. Reson.*, **44**, 542–561.

Berendsen, H.J.C., Postma, J.P.M., Van Gunsteren, W.F. and Hermans, J. (1981) In *Intermolecular Forces* (Ed., B. Pullman), Reidel, Dordrecht, pp. 331–342.

Bernardi, A. and Raimondi, L. (1995) *J. Org. Chem.*, **60**, 3370–3377.

Blanchard, D., Cartron, J.-P., Fournet, B., Montreuil, J., Van Halbeek, H. and Vliegenthart, J.F.G. (1983) *J. Biol. Chem.*, **258**, 7691–7695.

Bock, K., Pedersen, C. and Pedersen, H. (1984) *Adv. Carbohydr. Chem. Biochem.*, **42**, 193–225.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–189.

Brocca, P., Berthault, P. and Sonnino, S. (1998) *Biophys. J.*, **74**, 309–318.

Dall' Olio, F., Chiricolo, M., Malagolini, N., Franceschi, C. and Sefarini Cessi, F. (1991) *Cell. Immunol.*, **137**, 303–315.

Donald, A.S.R., Yates, A.D., Soh, C.P.C., Morgan, W.T.J. and Watkins, W.M. (1983) *Biochem. Biophys. Res. Commun.*, **115**, 625–631.

Donald, A.S.R. and Feeney, J. (1986) *Biochem. J.*, **236**, 821–828.

Duncan, J.L. (1988) *J. Infect. Dis.*, **158**, 1379–1382.

French, A.D. (1989) *Carbohydr. Res.*, **188**, 206–211.

Gillard, B.K., Blanchard, D., Bouhours, J.F., Cartron, J.-P., Van Kuik, J.A., Kamerling, J.P., Vliegenthart, J.F.G. and Marcus, D.M. (1988) *Biochemistry*, **27**, 4601–4606.

Grootenhuis, P.D.J. and Haasnoot, C.A.G. (1988) *Mol. Simul.*, **10**, 75–95.

Hård, K., Spronk, B.A., Hokke, C.H., Kamerling, J.P. and Vliegenthart, J.F.G. (1991) *FEBS Lett.*, **287**, 108–112.

Hård, K., Van Zadelhoff, G., Moonen, P., Kamerling, J.P. and Vliegenthart, J.F.G. (1992) *Eur. J. Biochem.*, **209**, 895–915.

Koerner Jr., T.A.W., Prestegard, J.H., Demou, P.C. and Yu, R.K. (1983) *Biochemistry*, **22**, 2676–2687.

Leeflang, B.R. and Kroon-Batenburg, L.M.J. (1992) *J. Biomol. NMR*, **2**, 495–518.

Levery, S.B. (1991) *Glycoconjugate J.*, **8**, 484–492.

Lommerse, J.P.M., Kroon-Batenburg, L.M.J., Kroon, J., Kamerling, J.P. and Vliegenthart, J.F.G. (1995a) *J. Biomol. NMR*, **6**, 79–94.

Lommerse, J.P.M., Kroon-Batenburg, L.M.J., Kroon, J., Kamerling, J.P. and Vliegenthart, J.F.G. (1995b) *Biochemistry*, **34**, 8196–8206.

Macvie, S.I., Morton, J.A. and Pickles, M.M. (1967) *Vox Sang.*, **13**, 485–492.

Milton, M.J., Harris, R., Probert, M.A., Field, R.A. and Homans, S.W. (1998) *Glycobiology*, **8**, 147–153.

Muchmore, A. and Decker, J.M. (1985) *Science*, **229**, 479–481.

Nishida, Y., Hori, H., Ohrui, H. and Meguro, H. (1987) *Carbohydr. Res.*, **170**, 106–111.

Ohrui, H., Nishida, Y., Watanabe, M., Hori, H. and Meguro, H. (1985) *Tetrahedron Lett.*, **26**, 3251–3254.

Parkkinen, J., Virkola, R. and Korhonen, T.K. (1988) *Infect. Immun.*, **56**, 2623–2630.

Perkin, S.J., Johnson, L.N., Phillips, D.C. and Dwek, R.A. (1977) *Carbohydr. Res.*, **59**, 19–34.
Poppe, L., Dabrowski, J., von der Lieth, C.-W., Numata, M. and Ogawa, T. (1989) *Eur. J. Biochem.*, **180**, 337–342.
Poppe, L. and Van Halbeek, H. (1991) *J. Am. Chem. Soc.*, **113**, 363–365.
Poppe, L., Brown, G.S., Philo, J.S., Nikrad, P.V. and Shah, B.H. (1997) *J. Am. Chem. Soc.*, **119**, 1727–1736.
Poveda, A. and Jiménez-Barbero, J. (1998) *Chem. Soc. Rev.*, **27**, 133–143.
Rutherford, T.J., Spackman, D.G., Simpson, P.J. and Homans, S.W. (1994) *Glycobiology*, **4**, 59–68.
Sabesan, S., Bock, K. and Lemieux, R.U. (1984) *Can. J. Chem.*, **62**, 1034–1045.
Sabesan, S., Duus, J.Ø., Fukunaga, T., Bock, K. and Ludvigsen, S. (1991) *J. Am. Chem. Soc.*, **113**, 3236–3246.
Sathyamoorthy, N., Decker, J.M., Sherblom, A.P. and Muchmore, A.V. (1991) *Mol. Cell. Biochem.*, **102**, 139–147, and references cited therein.
Scarsdale, J.N., Prestegard, J.H. and Yu, R.K. (1990) *Biochemistry*, **29**, 9843–9855.
Toma, G., Bates, J.M. and Kumar, S. (1994) *Biochem. Biophys. Res. Commun.*, **200**, 1974–1978.
Tran, V., Buleon, A., Imberty, A. and Pérez, S. (1989) *Biopolymers*, **28**, 679–690.
Tvaroškova, I. and Bleha, T. (1989) *Adv. Carbohydr. Chem. Biochem.*, **47**, 45–123.
Van Eijck, B.P. and Kroon, J. (1989) *J. Mol. Struct.*, **195**, 133–146.
Van Gunsteren, W.F. and Berendsen, H.J.C. (1977) *Mol. Phys.*, **34**, 1311–1327.
Van Gunsteren, W.F. (1987) *GROMOS, Groningen Molecular Simulation Package*, University of Groningen.
Van Rooijen, J.J.M., Kamerling, J.P. and Vliegenthart, J.F.G. (1998) *Glycobiology*, **8**, 1065–1075.
Van Rooijen, J.J.M., Voskamp, A., Kamerling, J.P. and Vliegenthart, J.F.G. (1999) *Glycobiology*, **9**, 21–30.
Van Seeventer, P.B., Kamerling, J.P. and Vliegenthart, J.F.G. (1997) *Carbohydr. Res.*, **299**, 181–195.
Watkins, W.M. (1995) In *Blood Cell Biochemistry: Molecular Basis of Major Human Blood Group Antigens*, Vol. 6 (Eds., Cartron, J.-P. and Rouger, P.), Plenum Press, New York, NY, pp. 351–375, and references cited therein.
Williams, J., Marshall, R.D., Van Halbeek, H. and Vliegenthart, J.F.G. (1984) *Carbohydr. Res.*, **134**, 141–155.