Glycoconjugate Journal (1993) **10**:365–380

Preferred conformations and dynamics of five core structures of mucin type *O*-glycans determined by NMR spectroscopy and force field calculations

ANNETTE POLLEX-KRÜGER[1], BERND MEYER[2], RAINER STUIKE-PRILL[2], VOLKER SINNWELL[3], KHUSHI L. MATTA[4] and INKA BROCKHAUSEN[1,5]*

[1] *Research Institute, The Hospital for Sick Children, Toronto, M5G 1X8 Ontario, Canada*
[2] *Complex Carbohydrate Research Center and Department of Biochemistry, University of Georgia, Athens, GA, USA*
[3] *Institut für Organische Chemie, Universität Hamburg, Hamburg, Germany*
[4] *Department of Gynecologic Oncology, Roswell Park Memorial Institute, New York State Department of Health, Buffalo, NY, USA*
[5] *Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada*

Received 23 June 1992

Glycosyltransferases acting on *O*-glycans have been shown to exhibit distinct specificity for the carbohydrate and the peptide moiety of their substrates. As an approach to study the 3-dimensional interactions between enzymes and *O*-glycan substrates, we determined the preferred conformations of five oligosaccharide-core structures of mucin type glycoproteins by NMR spectroscopy and by static and dynamic force field calculations. Seven oligosaccharides, representing five basic core structures, were investigated: Galβ(1-3)GalNAcαBzl (**1**, core 1), GlcNAcβ(1-6)[Galβ(1-3)]GalNAcαBzl (**2**, core 2), GlcNAcβ(1-3)GalNAcαBzl (**3**, core 3), GlcNAcβ(1-6)-[GlcNAcβ(1-3)]GalNAcαBzl (**4**, core 4), GlcNAcβ(1-6)GalNAcαBzl (**5**, core 6), the elongated core 2, Galβ(1-4)GlcNAcβ(1-6)[Galβ(1-3)]GalNAcαpNp (**6**) and GalNAcα-Bzl (**7**). The dynamic behaviour of the molecules was studied by Metropolis Monte Carlo (MMC) simulations. Experimental coupling constants, chemical shift changes, and NOEs were compared with results from static energy minimizations and dynamic MMC simulations and show a good agreement. MMC simulations show that the (1-6) linkage is much more flexible than the (1-3) or the (1-4) linkages. The preferred conformations of the disaccharides (**1**) and (**3**) show only slight differences due to the additional *N*-acetyl group in (**3**). The conformational equilibrium of β(1-3) glycosidic bonds of **1** and **3** was not affected by attaching a β(1-6) linked GlcNAc unit to the GalNAc residue in **2** and **4**. However, experimental and theoretical data show that the β(1-6) linkages of the trisaccharides **2** and **4**, which carry an additional β(1-3) linked glycosyl residue, change their preferred conformations when compared with (**5**). The 6-branch also shows significant interactions with the benzyl aglycon altering the preferred conformation of the hydroxymethyl group of the GalNAc to a higher proportion of the gt conformer. The (1-6) linkage of **2**, **4**, and **6** can have two different families of conformations of which the lower energy state is populated only to about 20% of the time whereas the other state with a relative enthalpy of $\approx$4 kcal mol^{-1} is populated to 80%. This fact demonstrates that the two conformational states have different entropy contents. Entropy is implicitly included in MMC simulations but cannot be derived from energy minimizations.

Keywords: mucin, *O*-glycan cores, conformation.

Abbreviations: Bzl, benzyl; COSY, correlation spectroscopy; Gal, D-galactose; GalNAc, *N*-acetyl-D-galactosamine; GalNAc-ol, *N*-acetylgalactosaminitol; GlcNAc, *N*-acetyl-D-glucosamine; HOHAHA, homonuclear Hartmann-Hahn-spectroscopy; MMC, metropolis Monte Carlo; NOE, nuclear Overhauser enhancement; pNp, *p*-nitrophenyl; ROESY, rotating frame Overhauser enhancement spectroscopy; TOCSY, totally correlated spectroscopy.

GalNAcα-Ser/Thr-linked oligosaccharides (*O*-glycans) occur in many secreted and membrane bound glycoproteins. In mucins, six core structures of *O*-glycans are common [1]: core 1, Galβ(1-3)GalNAcα-; core 2, GlcNAcβ(1-6)-[Galβ(1-3)]GalNAcα-; core 3, GlcNAcβ(1-3)GalNAcα-; core 4, GlcNAcβ(1-6)[GlcNAcβ(1-3)]GalNAcα-; core 5, GalNAcα(1-3)GalNAcα-; and core 6, GlcNAcβ(1-6)-GalNAcα-. These core structures may be elongated in a

* To whom correspondence should be addressed.

variety of ways to produce often extremely heterogeneous chains. *O*-Glycans influence the chemical, physical and antigenic properties of glycoproteins and have protective and recognition functions. However, very little is known about the biological roles of specific structures.

Glycosyltransferases assembling *O*-glycans are highly specific for the oligosaccharide moiety of their substrates as well as the peptide portion [2]. It has been suggested that processing of *N*-glycans is in part directed by the recognition of the 3-dimensional conformations of oligosaccharide substrates [3]. Nothing is known about the recognition by glycosyltransferases acting on *O*-glycans of the 3-dimensional structures of their substrates. The conformations of partial structures occurring on *O*-glycoproteins have been investigated by several groups [4–11]. NMR studies of entire mucin molecules revealed the mobilities of the oligosaccharide groups [12, 13]. An extensive literature exists on the structural elucidation of mucin-type oligosaccharide structures including many NMR studies [1, 14–18]. These NMR data were obtained using oligosaccharides released by β-elimination thus containing reduced GalNAc-ol. A systematic conformational analysis by experimental and theoretical approaches of the common *O*-glycan core structures where GalNAc is present in the pyranose form has not yet been carried out.

Experimental information on the solution conformations of oligosaccharides can be obtained by quantitative evaluation of NMR data. Important parameters are obtained from NOE measurements, coupling constants and chemical shifts. Oligosaccharides show NOE contacts mainly across glycosidic bonds, whereas other structurally relevant NOE effects are rare [19, 20]. Therefore, a conformational analysis based solely on NMR experiments contains uncertainties in the determination of oligosaccharide conformations. This problem can be overcome by combining NMR data with theoretically derived conformations or conformational equilibria.

In this paper we investigated the 3-dimensional structure of five *O*-glycan cores. We applied the Metropolis Monte Carlo (MMC) method that is included in the newly developed force field program GEGOP [21, 22]. Statistical methods of computation like the MMC approach can be used to study the dynamics of biological molecules [22]. Calculations of dynamic equilibria are better suited to deriving time averaged physical properties like NOEs than energy minimization techniques. Local minima as well as information from the calculation of the dynamic behaviour were used to describe the conformations of five core structures which commonly occur in *O*-glycoproteins, cores 1 to 4, and 6, and an elongated core 2 structure (Fig. 1). We have chosen benzyl glycosides for these studies since they have GalNAc in the α-pyranose configuration and have proved to be excellent enzyme substrates [23–26]. We also present the complete ^{1}H-NMR assignments for the synthetic benzyl glycosides **1**–**5** and **7**. The chemical shifts and coupling constants of the *p*-nitrophenyl glycoside **6** have been presented before [27]. NOE data derived from 2D-ROESY and 1D steady state experiments as well as coupling constants were obtained for compounds **1**–**6** to compare the solution conformations to the results of the theoretical calculations.

Materials and methods

Materials

Compound **7**, GalNAcα-Bzl, was purchased from Sigma. Compounds **1** to **5** (Fig. 1) were synthesized as described before: Galβ(1-3)GalNAcα-Bzl (**1**, [28]); GlcNAcβ(1-6)-[Galβ(1-3)]GalNAcα-Bzl (**2**, [29]); GlcNAcβ(1-3)GalNAcα-Bzl (**3**, [30]); GlcNAcβ(1-6)[GlcNAcβ(1-3)]GalNAcα-Bzl (**4**, [31]); GlcNAcβ(1-6)GalNAcα-Bzl (**5**, [32]). Compound **6** was kindly donated by Dr J. Dennis, Mount Sinai Hospital, Toronto [27].

^{1}H-NMR spectroscopy

Compounds **1**–**7** were exchanged and lyophilized twice in 99.9% ^2H_2O (Aldrich) and twice in 99.96% ^2H_2O (Merck, Sharpe and Dohme) and dried under high vacuum before dissolving in 99.96% ^2H_2O under argon atmosphere. Experiments were carried out at 300 K on Bruker AM 300 and AM 500 spectrometers equipped with Aspect 3000 computers at the Toronto Carbohydrate Research Center, University of Toronto, and a Bruker WM 400 spectrometer with an Aspect 2000 computer at the Institut für Organische Chemie, Universität Hamburg, FRG. Acetone was used as an external standard (chemical shift set at 2.225 ppm relative to 2,2-dimethyl-2-silapentane-5-sulfonate, DSS). Spectral assignments were obtained from 2D H,H-COSY [25] and 2D HOHAHA [33, 34] experiments. The 2D HOHAHA experiment was run with the MLEV-17 composite pulse sequence [35] to obtain an efficient spinlock and a mixing time of 85 ms. 1D NOE experiments were performed under steady state conditions with an irradiation time of 5 s at a decoupler power of 45 W. Reference FID were subtracted directly after each scan. 2-Dimensional rotating frame NOE spectroscopy (2D ROESY) experiments [36] were carried out using a 10°–60°–140° composite pulse for the exciting pulse. The exact 90° pulse length (about 70 μs) was determined prior to each experiment. Spectra were recorded using a mixing time of 200 ms, 512 increments in t_1 and 32 scans for each t_1 experiment. Before Fourier transformation, the time domain matrix was zero-filled to 2048 points in t_1 and the final matrix of 2K × 2K data points yielded a resolution of 2.4 Hz per point. Spectral widths were 2.5–4.5 kHz. All spectra were weighted prior to processing with squared phase-shifted sine bell functions. A μVAX II data station and Dennis Hare's (Infinity Systems, Seattle, WA, USA) FTNMR program were used for processing the 2D NMR data.

Force field calculations

All calculations were performed with the GEGOP program [21]. The energy of the oligosaccharides was computed using an improved HSEA force field [5]. The modifications include a new potential for the rotation of hydroxymethyl groups and the inclusion of hydrogen atoms of hydroxy groups. The coordinates of the monosaccharide units were taken from x-ray or neutron diffraction studies. The positions of C–H hydrogen atoms cannot be determined with sufficient precision from x-ray structure analyses and were thus placed according to geometrical criteria using the HGEN program (B. Meyer, unpublished). Bond angles, bond lengths, and the conformation of the pyranose rings of the monosaccharide residues were kept constant during the GEGOP calculations. The glycosidic bond angle was set to 117°. The program optimizes the torsional angles ϕ (H1′-C1′-O1′-Cx) and ψ (C1′-O1′-Cx-Hx) of the glycosidic linkage, where x denotes the aglyconic linkage site, the angle ω (O6-C6-C5-O5) of the exocyclic bonds as well as the rotation of H-C-O-H fragments and the rotation of *N*-acetyl groups. The torsion angle ψ in (1-6) linkages is defined by C1′-O1′-C6-C5. In the geometry optimizations all possible rotamers of the (1-6) linkages were used as starting conformations. All computations were carried out using the free oligosaccharide structures excluding the aglycons as well as using the benzyl oligosaccharides.

Simulations of the dynamics of the molecules were performed utilizing the Metropolis Monte Carlo algorithm [22, 37, 38]. Monte Carlo simulations with 300 000 steps were run for each structure, during which 90 000–120 000 different conformations were generated and stored. The equilibrium state of MMC calculations is normally reached after 5000 to 10 000 MMC steps. The temperature for the Metropolis criterion was normally chosen to be 300 K. One simulation for **6** was run at 600 K. All dihedral angles were changed simultaneously in order to sample the conformational space efficiently. The maximum intervals for the variations of the dihedral angles per step were chosen between 12° and 45° to obtain an acceptance rate of 30–50%. For energy minimizations, suitable starting conformations were chosen from highly populated areas of the dihedral angle distribution plots calculated from the Monte Carlo runs and minimized with the GEGOP program utilizing the Davidon–Fletcher–Powell algorithm [39].

Theoretical NOEs were calculated as ensemble averages of all Monte Carlo steps using the full relaxation matrix method [40] for $\langle r^{-3} \rangle^2$. The relaxation matrix was updated at each step of the calculation. The correlation coefficient τ_c for all compounds was set to $\tau_c = 10^{-10}$ s, the temperature was set to 300 K. Calculations with $\tau_c = 10^{-11}$ s showed that the relative NOEs are not significantly affected by a variation of the correlation time within this range.

All calculations were performed on DECsystems 5500. Molecular graphics were performed using the Sybyl software package (Tripos Assoc.) on a Silicon Graphics 4D/220 GTX computer. The traces, scatter plots, and contour plots visualizing the data from the Monte Carlo simulations have been plotted using the PV-WAVE (Precision Visuals) program.

Results

NMR spectroscopy

The ^{1}H-NMR spectral parameters were assigned from the 1D spectra in combination with 2D COSY and other 2D correlated spectroscopy spectra. Spectral parameters for **1–6** (Fig. 1) were compared with the NMR spectrum of GalNAcα-Bzl (**7**) and literature data [41]. The assignments of the H-6R and H-6S protons were made as described by Ohrui *et al.* [42]. The chemical shifts and coupling constants of benzyl glycosides **1–5** and **7** and *p*-nitrophenyl glycoside (**6**) are shown in Table 1. These assignments were obtained from 1D ^{1}H-NMR spectra, 2D COSY, long-range COSY, triple quantum-filter COSY, ROESY or C-H correlation spectra. The CH_3 signals for the *N*-acetyl groups could not be assigned unequivocally. An example of 1D and long range COSY spectrum of **4** recorded at 400 MHz is shown in Fig. 2. The 500 MHz ^{1}H-NMR spectra (1D spectra, 2D H,H-COSY and partial 2D ROESY) of tetrasaccharide **6** together with the proton assignments (chemical shift values and coupling constants) have been reported before [27]. In addition to the NOE data of **6**, we have included chemical shifts and most of the coupling constants of **6** for reference.

core 1	Gal′ β(1-3) GalNAc α- Bzl	1
	GlcNAc″ | β(1-6)	
core 2	Gal′ β(1-3) GalNAc α-Bzl	2
core 3	GlcNAc′ β(1-3) GalNAc α- Bzl	3
	GlcNAc″ | β(1-6)	
core 4	GlcNAc′ β(1-3) GalNAc α-Bzl	4
	GlcNAc″ | β(1-6)	
core 6	GalNAc α-Bzl	5
	Gal‴ β(1-4) GlcNAc″ | β(1-6)	
	Gal′ β(1-3) GalNAc α-pNp	6
	GalNAcα-Bzl	7

Figure 1. Structures of synthetic compounds **1–7**.

Table 1. ^{1}H-NMR chemical shifts (ppm) and coupling constants J (Hz), given in parentheses, of compounds **1–7** (^2H_2O, 400 MHz, acetone = 2.225 ppm).

Residue	*Proton (J)*	**1**	**2**	**3**	**4**	**5**	**6**[a]	**7**
GalNAc	H1 ($J_{1,2}$)	4.97 (3.8)	4.97 (3.8)	4.91 (3.8)	4.89 (3.8)	4.97 (3.8)	5.81 (3.6)	4.96 (3.8)
	H2 ($J_{2,3}$)	4.31 (11.2)	4.31 (11.1)	4.22 (11.1)	4.23 (11.0)	4.13 (11.1)	4.57 (10.5)	4.12 (11.1)
	H3 ($J_{3,4}$)	4.04 (3.0)	4.02 (3.0)	3.97 (3.0)	3.94 (3.0)	3.89 (3.2)	4.29 (3.1)	3.91 (3.2)
	H4 ($J_{4,5}$)	4.25 (−0.8)	4.23 (−0.8)	4.23 (−0.6)	4.20 (−1.0)	3.96 (−0.8)	4.31 (−0.6)	3.99 (−0.8)
	H5 ($J_{5,6}$)	4.06 (6.3)	4.15	4.04 (6.3)	4.13	4.08	4.18	4.02 (6.2)
	H6R ($J_{5,6R}$)	3.74	3.70 (8.4)	3.73	3.70 (8.3)	3.68 (7.2)	3.75 (7.9)	3.75
	H6S ($J_{5,6S;6R,6S}$)	3.74	4.07 (3.2; −10.8)	3.73	4.07 (3.2; −11.0)	4.05 (4.1; −10.2)	4.00 (3.9; −11.2)	3.75
	Ac	1.96	1.96[b]	1.98[b]	1.95[b]	1.97[b]	2.01[b]	1.96
	C_6H_5	7.45 (*m*)	7.447 (*m*)	7.455 (*m*)	7.433 (*m*)	7.445 (*m*)	8.300 (*m,m-H*) 7.269 (*m,o-H*)	7.425 (*m*)
	CH_2 ($J_{CH_2,CH_2'}$)	4.77 (−11.8)	4.71 (−11.4)	4.75 (−11.8)	4.69 (−11.4)	4.69 (−11.5)		4.76 (−11.9)
	CH_2,	4.57	4.50	4.55	4.47	4.51		4.55
Gal-β(1-3)	H1′ ($J_{1,2}$)	4.45 (7.8)	4.43 (7.8)				4.54 (7.8)	
	H2′ ($J_{2,3}$)	3.51 (10.0)	3.50 (10.0)				3.56 (9.9)	
	H3′ ($J_{3,4}$)	3.61 (3.4)	3.60 (3.3)				3.65 (3.4)	
	H4′ ($J_{4,5}$)	3.90 (−1.0)	3.89 (−0.6)				3.92 (−0.6)	
	H5′	3.65	3.63				3.69	
	H6R′ ($J_{5,6R}$)	3.76 (7.8)	3.76 (7.8)				3.80[c]	
	H6S′ ($J_{5,6S;6R,6S}$)	3.73 (4.3; −11.8)	3.71 (4.4; −11.9)				3.77[c]	
GlcNAc β(1-3)	H1′ ($J_{1,2}$)			4.56 (8.4)	4.53 (8.4)			
	H2′ ($J_{2,3}$)			3.69 (10.2)	3.67 (10.2)			
	H3′ ($J_{3,4}$)			3.53 (8.6)	3.51 (8.6)			
	H4′ ($J_{4,5}$)			3.45 (9.4)	3.44 (9.6)			
	H5′			3.40	3.38			
	H6R′ ($J_{5,6R}$)			3.76 (5.0)	3.74 (4.8)			
	H6S′ ($J_{5,6S;6R,6S}$)			3.88 (2.1; −12.4)	3.87 (2.2; −12.2)			
	Ac′			1.99[b]	1.97[b]			
GlcNAc β(1-6)	H1″ ($J_{1,2}$)		4.54 (8.4)		4.54 (8.6)	4.52 (8.5)	4.47 (8.3)	
	H2″ ($J_{2,3}$)		3.74 (10.1)		3.74 (10.2)	3.73 (10.1)	3.61[c]	
	H3″ ($J_{3,4}$)		3.54 (8.4)		3.52 (8.5)	3.55 (8.7)	3.51[c]	
	H4″ ($J_{4,5}$)		3.44 (9.6)		3.43 (9.6)	3.44 (9.6)	3.67[c]	
	H5″		3.47		3.47	3.48	3.50	
	H6R″ ($J_{5,6R}$)		3.73 (−3.0)		3.73 (5.4)	3.74 (5.3)	3.64[c]	
	H6S″ ($J_{5,6S;6R,6S}$)		3.93 (1.8; −12.2)		3.93 (2.0; −12.4)	3.94 (1.8; −12.4)	3.93[c]	
	Ac″		1.97[b]		1.98[b]	1.98[b]	1.90[b]	
Gal-β(1-4)	H1‴ ($J_{1,2}$)						4.39 (7.8)	
	H2‴ ($J_{2,3}$)						3.54 (10.0)	
	H3‴ ($J_{3,4}$)						3.66 (3.2)	
	H4‴ ($J_{4,5}$)						3.92 (−0.6)	
	H5‴						3.71	
	H6R‴						3.77[c]	
	H6S‴						3.74[c]	

[a] pNp-derivative, recorded at 500 MHz.
[b] Signals of acetyl groups may be interchanged.
[c] Coupling constants could not be obtained due to overlapping signals.

The intra ring coupling constants clearly show $^4C_1$ conformations for each of the sugar residues. In comparison with **7**, the introduction of β-D-Gal or β-D-GlcNAc units into the 3-position of α-D-GalNAc in **1–3** leads to downfield shift effects for H-2 (0.19 and 0.10 ppm), H-3 (0.13 and 0.06 ppm) and H-4 (0.26 and 0.24 ppm) of the GalNAc residue. These effects are more significant in Galβ(1-3)-GalNAc (**1**) than in GlcNAcβ(1-3)GalNAc (**3**). The extra residue at the 3-position has no effect on the δ-values of the exocyclic CH_2OH group.

Introduction of a β(1-6) linked GlcNAc unit in **5** produces a large deshielding effect of 0.30 ppm on the GalNAc's H-6S resonance and only minor shift effects for H-5 and H-6R. In **2** and **4**, H-5 of GalNAc shows downfield shifts of 0.13 and 0.11 ppm and H-6S exhibits a downfield shift of 0.32 ppm. In contrast, H-6R of GalNAc shows slight upfield shifts of −0.05 ppm in **2** and **4** and of −0.07 ppm in **5**. These shielding effects suggest that oxygen atoms of the GlcNAcβ(1-6) residue are a significant portion of the time in close proximity to H-6S of the GalNAc residue. The

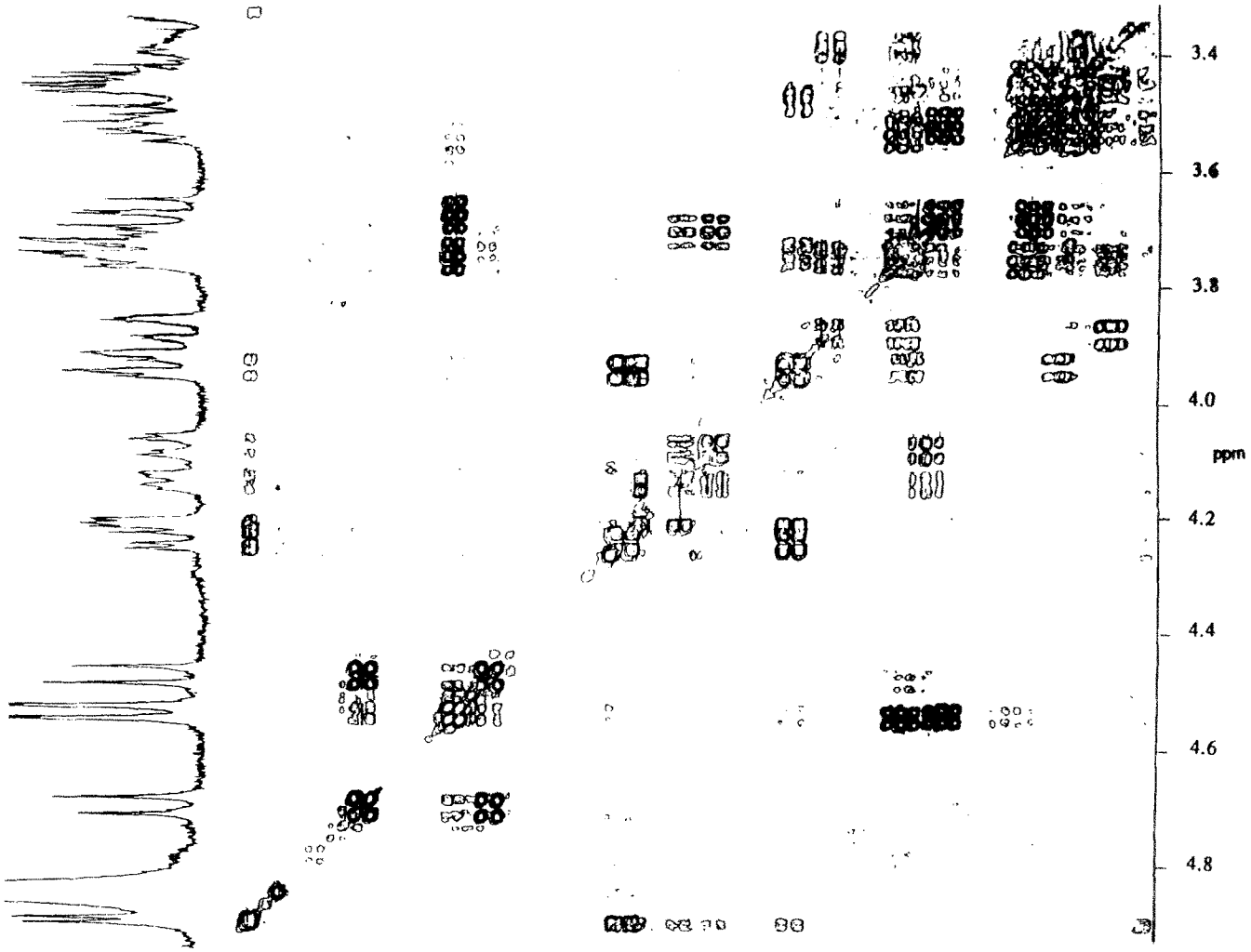

Figure 2. 400 MHz ^{1}H-NMR spectra (1D and 2D long-range COSY) of compound **4**.

change of the coupling constant C5-C6 suggests that rather H-6R should exhibit a downfield shift from being a higher portion of the time in close proximity to O4 of the GalNAc residue, whereas H-6S should show a high field effect from the atoms of its own pyranose ring.

No significant shift effects or changes in coupling constants are observed for the β(1-3) units upon addition of a β(1-6) linked GlcNAc residue to GalNAc, i.e., **2** versus **1** and **4** versus **3**. This indicates that the 6-branch in **2** and **4** has no significant influence on the conformation of the 3-branch. However, trisaccharides **2** and **4** exhibit coupling constants $J_{\text{H-5,H-6R}} = 8.3$ and $J_{\text{H-5,H-6S}} = 3.2$ Hz in the GalNAc residue which are significantly different from the disaccharide **5** which has $J_{\text{H-5,H-6R}} = 7.2$ and $J_{\text{H-5,H-6S}} = 4.1$ Hz. Thus, the β(1-3) unit obviously has a small influence on the distribution of the C5-C6 rotamers of GalNAc. This observation is in accordance with the general fact that the conformation of the β(1-3) linkage is less flexible than the conformation of the β(1-6) linkage; a change in the conformation of the β(1-6) linkage requires much less energy than a change in the conformation of the β(1-3) linkage (see below). The larger downfield shift of the GalNAc H-5 signal in trisaccharides **2** and **4** compared with **5** results probably also from a change in the conformation of the C5-C6 bond caused by interactions between the β(1-3) linked galactose and GlcNAc units and the β(1-6) linked GlcNAc residue.

Due to the *p*-nitrophenyl group attached to the GalNAc residue in tetrasaccharide **6**, the chemical shifts of the GalNAc protons differ significantly from those of the benzyl compounds **1**–**5**. Compared to **2**, shift differences are especially high for H-1, H-2, and H-3 with +0.92 ppm, +0.26 ppm and +0.27 ppm, respectively. The effects on the proton chemical shifts of the remaining residues are much smaller but still noticeable. A more remarkable fact, however, is that the *p*-nitrophenyl group produces an exchange of the chemical shifts of the H-1 protons of the Galβ(1-3) and GlcNAcβ(1-6) residues (Table 1). This is a genuine effect of the *p*-nitrophenyl group, as is shown by comparison with H-1 chemical shifts of *p*-nitrophenyl-substituted core 1 and core 2 oligosaccharides [27] and benzyl and *o*-nitrophenyl derivatives [23]. The effect can be explained by the strong anisotropy of the *p*-nitrophenyl

residue that has a significant shielding effect on the H-1 proton of the β(1-6) linked residue.

Experimental NOEs

Experimental NOE values were obtained from 1D steady state experiments and 2D ROESY experiments (Table 2). Data of 1D steady state NOE experiments were generally acquired at 300 MHz and are given in percentage relative to the value for the irradiated proton. As can be seen from Table 2, oligosaccharides **1–6** give rise to positive 1D NOE effects which can be expected from molecules of this size. Because of the complexity of ^{1}H spectra, 1D NOE experiments for tetrasaccharide **6** were performed at 500 MHz, but still show positive NOE effects. Sufficient NOE data could be obtained for the β(1-3) and β(1-6) linkages of **6** but not for the β(1-4) linkage because the GlcNAc H-3″ and H-4″ protons (see Table 1 and Fig. 1 for definitions) show strong overlap of resonances. ROESY experiments showed an ROE only between Galβ-H-1‴ and H-4″ (Table 2) and no ROE between H-1″ and H-3″.

Table 2. Experimental NOE interactions for chosen proton pairs in **1–6** as obtained from 1D difference NOE spectra and 2D ROESY experiments (in parentheses). Given are 1D NOE values (%) in percentages relative to the saturated peak (100%) and values for 2D ROESY crosspeaks classified as strong (s), medium (m) and weak (w) effects. The protons are labelled as in Table 1.

Proton pair	**1**	**2**	**3**	**4**[a]	**5**	**6**
β(1-3)-linkage						
H1′/H2′	4.8	2.4	1.2	6.9[b]	1.2	
H1′/H3	8.8	4.0	3.9	4.3		1.5
H1′/H3′	13.1	6.7	2.8	12.1[c]		2.6
H1′/J5′	(m)	(m)	3.1	5.4		(s)
β(1-6)-linkage						
H1″/H2″		4.8		6.9[b]	2.8	4.3
H1″/H5		1.6		0.6	0.7	0.5
H1″/H6R		2.8		6.9[b]	2.3	0.0
H1″/H6S		1.4		2.2	4.0	0.7
H1″/H3″		5.4		12.1[c]	5.0	6.0
H1″/H5″		5.9		(w)	6.7	(s)
β(1-4)-linkage						
H1‴/H2‴						4.0
H1‴/H6S″						0.0
H1‴/H6R″						(w)
H1‴/H4″						(s)
H1‴/H3‴						2.0
H1‴/H5‴						(s)

[a] 1D NOE values were obtained by irradiating H1′ and H1″ simultaneously. The 1D NOE effects were assigned to H1′ and H1″, respectively, by comparison with nonirradiated spectra.
[b] Proton signals of H2′, H2″ and H6R in 1D NOE experiments could not be integrated separately.
[c] Proton signals of H3′, H3″ and H5″ in 1D NOE experiments could not be integrated separately due to overlap.

ROESY experiments were run for all compounds. In general, the ROESY crosspeaks were not pure Overhauser enhancements but contained dispersion parts of TOCSY-type peaks, which were present in all of the spectra. The NOE data were therefore used for quantitation. Several ROESY experiments were run for each sample, varying the site of the offset and the power of the spinlock pulse. Neuhaus and Keeler [43] as well as Bax and Davies [36] have shown that the dependence of crosspeak intensities on the offset can be used to distinguish between scalar coupling effects and ROEs (Table 2).

Irradiation of H-1′ of Galβ(1-3) in **1, 2,** and **6** yielded NOEs across the glycosidic linkage to GalNAc H-3 of the same magnitude as intra-residue NOEs to Gal H-3′ and H-5′. No H-1′/H-4 NOE could be detected in any of the structures, indicating that the preferred conformations of this glycosidic linkage have only protons H-1′ and H-3 in close proximity. This is in contrast to results by Bush and Feeney [41] obtained for the Galβ(1-3)GalNAc disaccharide unit in antifreeze glycoprotein, where they found H-1′/H-3 and H-1′/H-4 NOEs of the same magnitude. Homans *et al.* [44] also observed an NOE from Gal H-1′ to GalNAc H-4 in antifreeze glycoprotein structures that was smaller than the H-1′/H-3 NOE. NMR studies of peptide models containing only GalNAc as the carbohydrate moiety linked to Thr suggest hydrogen bonding between Thr and GalNAc [45]. However, the benzyl glycosides studied here do not have polar interactions with the aglycon.

In **2, 4, 5,** and **6,** NOEs across the β(1-6) linkages are small. Some of the signals have an intensity just above the noise level. The strongest NOE across the β(1-6) linkage of **2** is observed between H-1″ and H-6R. Smaller NOEs are observed to H-6S and H-5. The interglycosidic NOEs of the β(1-6) linkage in **5** show a stronger NOE to H-6S and a weaker one to H-6R. The results for **6** are ambiguous as the 1D NOE experiment shows a stronger NOE to H-6S and the ROESY experiment shows a stronger ROE to H-6R.

Metropolis Monte Carlo simulations

MMC calculations were run in order to assess the carbohydrate's flexibility and in order to gain understanding of possible conformational variations. MMC simulations result in a random walk through the conformational space of the molecule. However, because of the intrinsic characteristics of MMC simulations, preferentially regions with a low energy content are sampled. Thus, MMC simulations can cover the conformational space of a molecule very efficiently. Additionally, these simulations weigh the conformations according to the Boltzmann statistics which implicitly includes the entropy in the evaluation of the preferred conformations. Thus, it is not the enthalpies of conformational states that are calculated by the force field that determine the statistical distribution between the states but rather the free energy.

Table 3. Results of Metropolis Monte Carlo simulations of structures **1**–**6**. The average values of dihedral angles ϕ, ψ, and ω and their variations during the MMC simulation are given as the average value with the variation in parentheses. The lowest energy (E_{min}), the energy ranges (E_{range}) covered and the acceptance rate (Acc. rate) are given.

Parameter	**1**	**2**	**3**	**4**	**5**	**6**
$\phi(\beta(1\text{-}3))$	55(113)	−56(94)	−56(95)	57(89)		56(107)
$\psi(\beta(1\text{-}3))$	−4(117)	−3(121)	−4(117)	−3(131)		−3(117)
$\phi(\beta(1\text{-}6))$		56(92)		57(97)	58(133)	62(105)
$\psi(\beta(1\text{-}6))$		−170(254)		178(259)	170(260)	130(245)
$\phi(\beta(1\text{-}4))$						55(89)
$\psi(\beta(1\text{-}4))$						0(88)
$\phi(\alpha(1\text{-}OB_21)$	−50(110)	−52(102)	−50(116)	−51(97)	−50(118)	−53(118)
$\psi(\alpha(1\text{-}OB_21)$	180(283)	−157(280)	173(289)	−161(293)	−172(299)	−151(281)
ω(GalNAc)	84(>360)	81(>360)	−112(>360)	38(>360)	113(>360)	91(348)
ω(Gal(1-3))	152(>360)	−65(>360)				97(>360)
ω(GlcNAc(1-3))			−131(>360)	60(>360)		
ω(Gal(1-4))						54(>360)
ω(GlcNAc(1-6))		28(>360)		−75(>36)	−146(>360)	−177(>360)
Acc. rates as %	32	38	25	36	30	30
E_{min} (kcal mol^{-1})	−0.6	−7.5	−1.8	−9.7	0.3	−9.1
E_{range} (kcal mol^{-1})	13.9	23.7	13.2	20.9	15.8	23.5

A summary of the relevant parameters of the MMC simulations describing the molecular structures is given in Table 3. The population profiles for all dihedral angles of **1**–**6** were analysed. An area of high population is found if the corresponding conformations have a low free energy (ΔG). The population density is displayed in scatter plots where each of the accepted conformations is represented as a dot in a ϕ/ψ diagram (Fig. 3). Alternatively, contour plots were used to display the population density where the population of each conformation is plotted relative to that of the highest populated conformation (Fig. 4). These contour plots were generated by applying the Boltzmann statistics to the calculated populations and thus represent ΔG values. The projection of the populations of individual angles are displayed as histograms. The scatter plots of the populations of the β(1-3) linkage (Fig. 3) indicate that the β(1-3) linkage has only one minimum. Most of the conformations for all (1-3) linkages are contained in a region of $15° < \phi < 80°$ and $-60° < \psi < 35°$. However, the populated ϕ,ψ-range of the β(1-3) linked GlcNAc of **3** is smaller than that of the β(1-3) linked Gal of **2** by the area of $15° < \phi < 5°$ and $-35° < \psi < -5°$. This is due to interactions between *N*-acetyl groups. The same effect is indicated on the energy contour plots (Fig. 3). It is evident that the boundary for significantly populated areas coincides approximately with the 3 kcal mol^{-1} contour line. Conformations with a higher energy content are not significantly populated. The distribution of dihedral angles of the β(1-3) linkage is not dependent on the presence of the 6-linked branch. Thus, from the population of the (1-3) linkage no conclusion on possible interactions between the (1-6) linked and the (1-3) linked branch could be drawn.

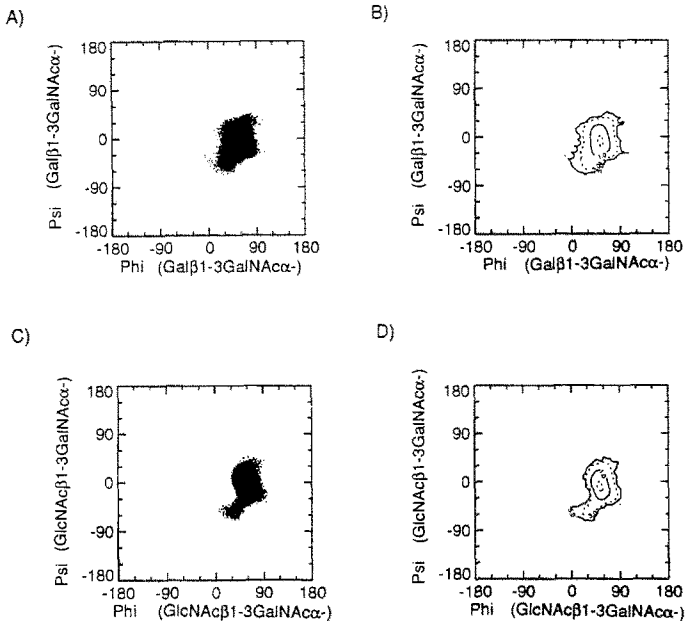

Figure 3. Scatter plots of the dihedral angle distributions (A, C) of a 300 000 step Metropolis Monte Carlo run for the β(1-3) linkages and conversion of these population distributions into free energies (B, D). (A) ϕ/ψ-plot of the Galβ(1-3)GalNAc linkage in Galβ(1-3)-GalNAcα-Bzl, **1**. (B) Free energy contour plot derived from populations shown in (A). (C) ϕ/ψ plot of the GlcNAcβ(1-3)-GalNAc linkage in GlcNAcβ(1-3)GalNAcα-Bzl, **3** and (D) free energy contour plot derived from populations shown in (C).

The β(1-6) linkage shows a significantly higher flexibility than the β(1-3) linkage. The energy contour plots of the low energy areas in ψ/ω space for **2**, **4**, **5**, and **6** shown in Fig. 6 are derived from corresponding population plots by

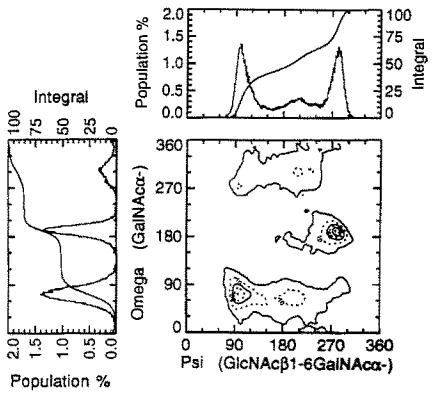

Figure 4. Contour plots of the relative population of the (1-6) linkage in ψ/ω space for GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-Bzl, **2**. The populations are given relative to the highest population area (see text) in %. Contour lines are drawn at 90, 70, 50, 30, 10, and 0.1%. Projections of the populations on to the ψ and ω axes are shown above and to the left of the contour plot, respectively. Note the asymmetrical distribution of the staggered conformations.

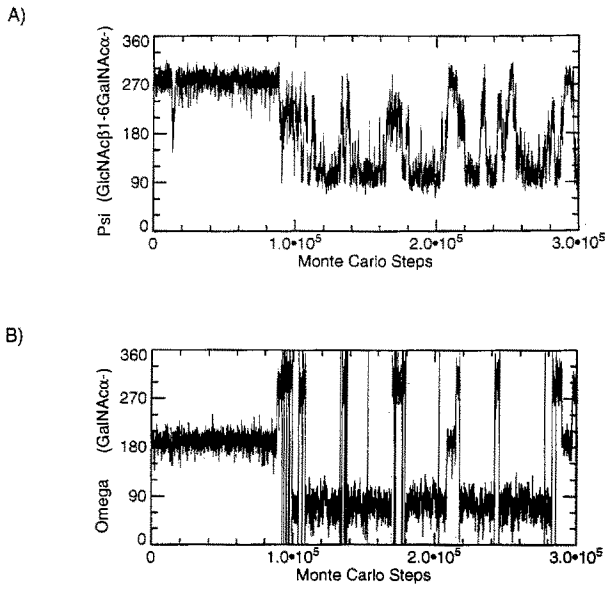

Figure 5. Traces of dihedral angles of the conformations obtained in a 300 000 step MMC simulation. (A) Trace of dihedral angle ψ of the GlcNAcβ(1-6)GalNAc linkage in GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-Bzl, **2**. (B) Trace of the dihedral angle ω of the GlcNAcβ(1-6)GalNAc linkage in **2**.

applying Boltzmann statistics. Fig. 6(B) shows the population plot in ψ/ω space of **4** which demonstrates that a significant portion of the total conformational space is actually occupied by the molecule. A similar situation is found for the other oligosaccharides with a (1-6) linkage, **2**, **5**, and **6** (Fig. 6A, C, D). A direct comparison of the energy contour plots of trisaccharides **2** and **4** indicates similarities between these two molecules, although the β(1-3)-linked residue is different in **2** and **4**. However, slight but distinct

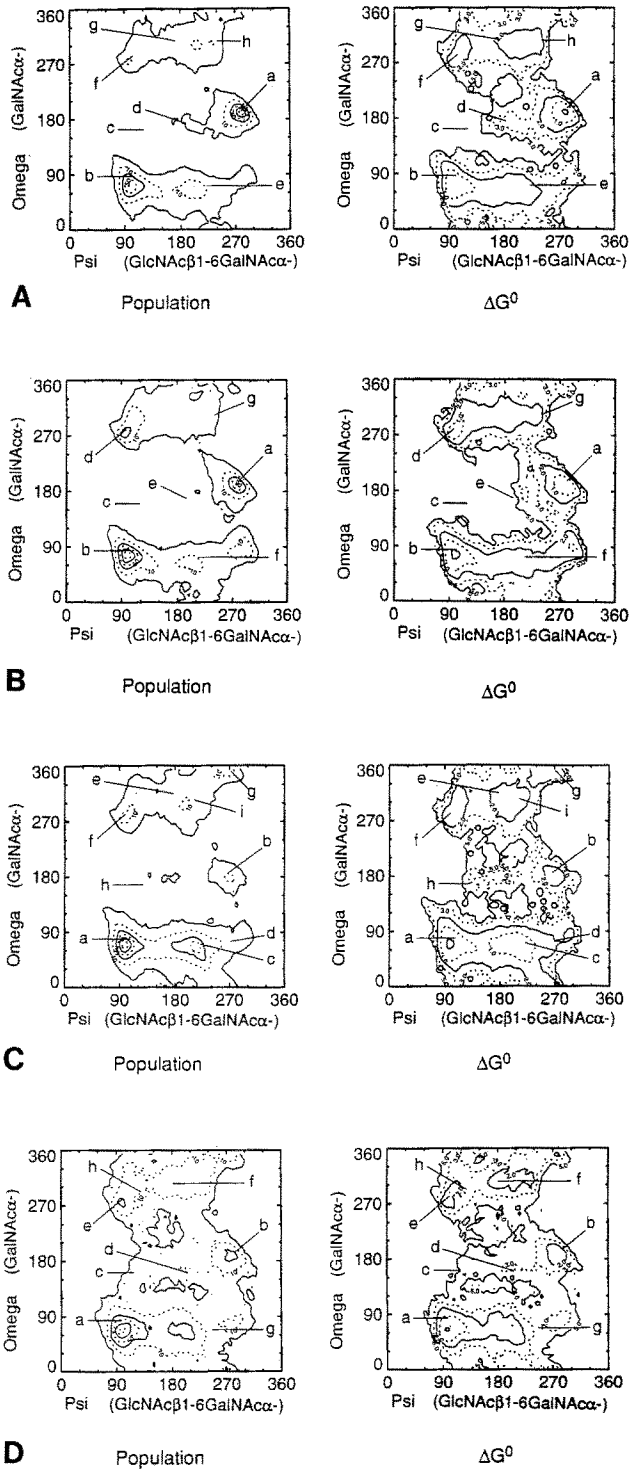

Figure 6. Comparison of the flexibility of the (1-6) linkages of GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-Bzl, **2** (A), GlcNAcβ(1-6)[GlcNAcβ(1-3)]GalNAcα-Bzl, **4** (B), GlcNAcβ(1-6)GalNAcα-Bzl, **5** (C), and Galβ(1-4)GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-pNp, **6** (D) demonstrated by the free energy contour plots in ψ/ω space. The contour lines shown are 0.1, 1, 3 and 5 kcal mol^{-1} of relative energy. The free energy is obtained from the populations calculated in MMC simulations. Areas indicated by small characters in the population and free energy plots refer to the positions of local minima corresponding to Table 5.

differences were found. In **2** the lowest energy conformation is found in the tg area ($\omega \approx 180°$) indicated as 'a' whereas the corresponding lowest energy area of **4** is found in the gt area ($\omega \approx 60°$) indicated as 'b'. Furthermore, **4** shows a lower energy in the gg area ($\omega \approx -60°$) than **2**. This indicates that there actually is an interaction between the (1-6) branch and the (1-3) branch. The fact that this effect is visible only in a change of the preferred conformation of the 6-branch is indicative of the significantly more shallow energy well of the (1-6) linkage when compared with the (1-3) linkage. Compound **5** has the least restricted (1-6) linkage because of the lack of a substituent in the 3 position of GalNAc indicated by the largest area covered in the energy contour map.

Traces of the dihedral angles were used in order to assure that the molecule was covering the whole conformational space. The presence of different conformational states of one dihedral angle can be recognized by transitions between distinct average values of this parameter (Fig. 5). A statistically significant analysis of the preferred conformations of a dihedral angle that can adopt more than one conformational state requires about 20 transitions during the MMC simulation.

The simulation of the tetrasaccharide **6** with a temperature parameter of 300 K did not provide enough transitions for the ω and ψ angles of the (1-6) glycosidic linkage. Thus, the simulation was repeated with a temperature parameter of 600 K, which showed frequent transitions in the traces of the ω and ψ angles. The conversion of the populations to the free energy show that the (1-6) linkage of tetrasaccharide **6** occupies a much smaller area of the conformational space than those of trisaccharides **2** and **4**. The conformational space of the (1-6) linkage seems to be mostly restricted if the 6-branch is in a tg conformation. This is indicated by a significantly lower relative population of the tg conformer of **6** compared with **2** of 17% and 34%, respectively. The higher temperature used in the calculation of **6** should have increased the relative population of the tg conformation if no additional restrictive forces would be operative. The relative populations of the ψ angle is also shifted from about equal populations of $\psi = -90°$ and $\psi = 90°$ for **2** with a temperature parameter of 300 K to a ratio of 1:4 for **6** at 600 K, which indicates that there is a relatively strong preference for a major conformer in **6**.

The projections of the populations for **2** on to the ω and ψ axes show that the population of the dihedral angle ω is not symmetrically distributed around the staggered conformations ($-60°$, $60°$, $180°$) but rather the maximum population is found at values that are shifted by 10–15°. Furthermore, it becomes clear that the ψ angles do not show a distribution of the population that is concentrated at one or two minima. About 33% of the conformers are found at values spreading between 120° and 240° (Fig. 4).

Traces of energy and acceptance rate are used as indicators for the stability of the MMC simulation. If only one energy well is occupied, the trace of the energy and the acceptance rate are normally stable after a short induction period necessary to reach equilibrium. However, if more than one energy well is occupied, both the acceptance rate and the energy may show repeated jumps. In this case one has to ensure that these jumps occur often enough to allow for statistical analysis of the dynamics of the conformational equilibrium. Correlated transitions in the acceptance rate and the energy of the molecule are used to identify areas in the conformational hyperspace that have a different shape, i.e., that have a different steepness of the energy wells. A low acceptance rate is caused by a narrow energy well and a high acceptance rate is correlated with a shallow energy well. A narrow energy well has a lower entropy (higher degree of order) than a wide energy well.

We find varying acceptance rate and energy in the MMC simulations with a temperature parameter of 300 K for **2** (Fig. 7), **4**, and **6**. Two different and highly populated regions are present in **2** that have an energy difference of ≈ 4 kcal mol^{-1}. The low acceptance rate is correlated with the low energy and the high acceptance rate is correlated with the high energy conformer. However, the lower energy conformation is only populated for $\approx 20\%$ of the time whereas the high energy conformation has a population of $\approx 80\%$. This feature indicates that the difference in entropy between the two conformations strongly disfavours the conformer with the lower enthalpy. Normally, in the interpretation of the results of energy minimizations, the Boltzmann distribution, $\Delta G = -RT \ln K$, is applied by substituting the free energy (ΔG) with the enthalpy that actually has been calculated. Using this approximation for the trisaccharide **2**, the Boltzmann distribution predicts that we should find less than 0.1% of the higher energy

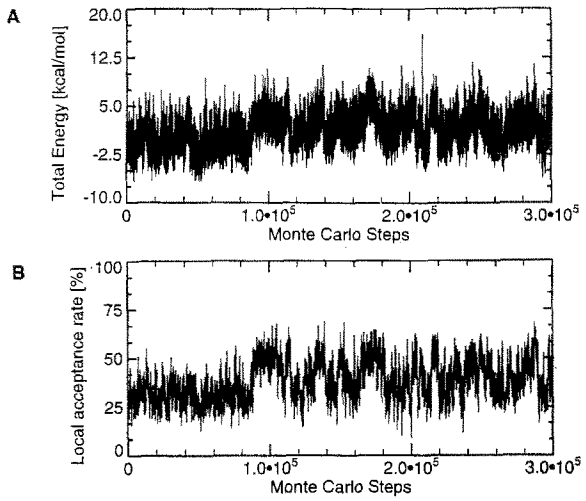

Figure 7. Traces of (A) the energy in kcal mol^{-1} and (B) acceptance rate obtained from a 300 000 step Metropolis Carlo simulation of GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-Bzl, **2**. Note the correlated transitions between high acceptance rate/high energy regions and low acceptance rate/low energy regions.

Table 4. Rotamer distributions in % of the C5-C6 bonds as derived from ^{1}H-NMR coupling constants in Hz[a] as well as calculated populations obtained from MMC simulations.

	Experimental					*Calculated from MMC*		
	$J_{5,6R}$	$J_{5,6S}$	gt	gg	tg	gt	gg	tg
GalNAcα unit								
2	8.4	3.2	65	25	10	51	15	34
4	8.3	3.2	65	25	10	57	21	22
5	7.2	4.1	50	30	20	79	17	4
6	7.9	3.9	58	25	17	63[b]	19[b]	18[b]
Galβ(1-3) unit								
1	7.8	4.3	54	24	22	71	20	9
2	7.8	4.4	54	23	23	83	11	6
GlcNAcβ(1-3) unit								
3	5.0	2.1	35	62	3	40	56	4
4	4.8	2.2	33	63	4	37	58	5
GlcNAcβ(1-6) unit								
4	5.4	2.0	40	59	1	42	56	2
5	5.3	1.8	39	61	−1	56	42	2

[a] The equations of Gerlt and Youngblood [46] were used for the calculations of the rotamer distributions: $1.3P_{gg} + 2.7P_{gt} + 11.7P_{tg} = J_{5,6S}$; $1.3P_{gg} + 11.5P_{gt} + 5.8P_{tg} = J_{5,6R}$; $P_{gg} + P_{gt} + P_{tg} = 1$; where: P_{gg} = population of the gg rotamer, P_{gt} = population of the gt rotamer, P_{tg} = population of the tg rotamer.
[b] Populations obtained from a MMC simulation at 600 K.

conformation. In fact, we find it populated at ≈80%. The difference in shape between these two conformations can be seen from the ΔG° plot of Fig. 6(A). The low energy conformation has a very narrow and steep well whereas the high energy conformation has a more shallow well, qualitatively indicating the difference in entropy between the two states. The region in the energy contour plot of **2** (Fig. 6A) representing the tg conformer ($\psi \sim -80°$, $\omega \sim -170°$) contains the conformations with the lowest free energy. The conformational space of the gt area does not show conformations with free energies as low as the tg area but is much wider spread in the ψ/ω diagram, thus allowing for a much higher total population of all gt conformers than all the tg conformers (Table 4).

Energy minimizations

Energy minimizations were started from conformations selected from the highly populated areas of the MMC simulations. The local minima found in the energy (enthalpy) minimizations have been marked in the free energy contour plots (Fig. 6). It is evident that the local minima from the enthalpy minimizations fall mostly into regions that also form local minima in the free energy contour plots. This correlation is by no means obvious because the local minima represent only the enthalpy of the molecule whereas the energy contours are based on the free energy of the molecules. The dihedral angles and energies of the conformations of **1**–**6** are summarized in Table 5. One minimum energy conformation is calculated for disaccharides **1** and **3**. The torsion angles of the β(1-3) linkages in **1** and **3** show the expected values of $\phi' = 60°$ and $\psi' = -10°$. In **2**, **4**, and **6**, the β(1-3) bond showed also only one minimum with $\phi' = 60°$ and $\psi' = -10°$.

Oligosaccharides **2**, **4**, **5**, and **6**, which all contain a β(1-6) linkage, show several local minima each originating mostly from local minima for the ψ and ω angles, respectively (Fig. 6 and Table 5). The two conformations of lowest energy **2**a, b, **4**a, b, **5**a, b, and **6**a, b have the ψ angle in the highly populated regions at −80° and 100° found in the MMC simulations. Most of the ψ angles of the other local minima are located in the region from 120° to 240°, which accounts for 33% of the total population of the ψ angle. This region does not have a marked minimum energy (high population) region. A significant diversity of conformations is calculated within a narrow energy band for **2**, **4**, **5**, and **6**. The three most favourable conformations of **2** and **4** are similar in shape, i.e. **2**a ≈ **4**a, **2**b ≈ **4**b, and **2**c ≈ **4**c. However, conformation **6**a is similar to **2**b and **4**b, respectively, and **6**b is similar to **2**a and **4**a. Thus, in **2** and **4** a tg conformer of the branched GalNAc has the lowest energy whereas the gt conformer is more favourable in **6** and **5** (Table 5). The reason for the stability of the gt conformer is probably the interaction with the benzyl aglycon. We found in the MMC simulations that it is essential to include the benzyl aglycon in the calculations. The calculated preferred conformations of the 6 branch were not in agreement with the experimental data if the aglycon was not included in the calculation. Thus, the interactions between the 6 branch and the benzyl aglycon seem to be represented quite well in the calculations *in vacuo* compared with the situation found experimentally in aqueous solution.

Discussion

The results of theoretical and experimental conformational analyses are in general agreement, indicating the validity of the methods applied. For **1**–**6** the preferred 3-dimensional structures can be derived from a combination of the experimental and the theoretical results described above. The conformation of an oligosaccharide is described by the conformation of the individual glycosyl residues and by the conformation of the glycosidic linkages. Experimentally, the conformation of the pyranose rings and the hydroxymethyl groups can be derived from coupling constants; the conformation of the glycosidic linkages can be derived from a combination of NOEs, shift effects and coupling constants.

Table 5. Dihedral angles and energies of the minimum energy conformations of **1–6** obtained from energy minimizations.

	E	β(1-3)	β(1-6)	β(1-4)	α(OBz)	ω_{GaN}	ω_{Gal3}	ω_{GN3}	ω_{GN6}	ω_{Gal4}
1	−1.3	56/7			−56/−116	71	71			
2a	−8.1	57/−10	48/−77		−54/−109	−164	62		−67	
2b	−6.7	61/6	62/101		−59/−137	87	74		59	
2c	−5.7	54/−28	61/120		−51/−95	163	65		−61	
2d	−3.7	58/−12	57/−177		−53/−101	175	71		−66	
2e	−3.7	57/−7	58/−138		−53/−145	71	69		−66	
2f	−3.6	57/−11	60/103		−41/−68	−77	68		−63	
2g	−3.5	56/−10	56/172		−43/−75	−52	69		−65	
2h	−3.1	57/−11	50/−120		−45/−84	−53	69		−65	
3	−2.3	56/−13			−54/−114	70		−65		
4a	−9.7	60/−9	49/−79		−54/−108	−163		53	−67	
4b	−7.4	58/−7	64/102		−55/−141	84		−63	−64	
4c	−7.4	61/−25	61/119		−51/−93	160		54	−61	
4d	−5.3	52/−1	60/97		−42/−68	−80		57	−62	
4e	−5.2	52/−3	59/−166		−53/−104	169		57	−66	
4f	−5.0	54/−4	62/−146		−58/−170	73		57	−66	
4g	−4.8	52/1	50/−119		−49/−96	−56		57	−66	
5a	−3.3		62/107		−52/−139	79			−64	
5b	−2.8		48/−94		−44/−87	−175			−66	
5c	−2.3		58/−139		−53/−145	71			−65	
5d	−2.3		55/−89		−60/64	75			−66	
5e	−2.2		57/177		−42/−73	−45			−63	
5f	−2.2		59/99		−41/−67	−78			−61	
5g	−1.7		51/−88		−48/−90	−9			−65	
5h	−1.3		57/128		−50/−94	167			−64	
5i	0.1		178/−148		−43/−81	−55			−64	
6a	−10.7	60/5	65/98	55/1	−58/−141	84	71		68	55
6b	−9.0	58/−10	47/−75	56/1	−43/−80	−165	60		69	−66
6c	−8.5	59/−28	67/119	60/10	−51/−97	162	66		67	−71
6d	−5.5	56/−10	52/−154	56/2	−44/−77	161	70		69	−67
6e	−5.4	57/−14	59/101	55/1	−41/−68	−78	69		69	−67
6f	−4.7	56/−10	57/−179	56/2	−41/−70	−54	69		69	−66
6g	−2.9	55/10	58/−114	56/1	−56/−172	69	68		69	161
6h	−2.8	57/−10	166/129	56/7	−38/−67	−67	68		74	−67

Intra ring coupling constants show that the pyranose rings in **1–6** adopt the expected $^4C_1$ conformation. The conformational equilibrium of the preferred conformations of the hydroxymethyl groups is derived experimentally by using the equation of Gerlt and Youngblood [46] (Table 4). The *N*-acetylglucosamine residues show a relative population of the gg to the gt conformation of 62:38 which is in agreement with literature data [47]. The calculated populations obtained from the MMC simulations of 56:40 (gg:gt) are in good agreement with the experimental data. The β(1-3) linked galactose residue shows an experimentally derived population of its hydroxymethyl group of gt:gg:tg = 54:23:23. This is in accordance with published NMR data [48, 49]. The calculated population is gt:gg:tg = 77:16:7 which indicates that free hydroxymethyl groups of galactose are calculated with a population of the tg conformer that is too low by ≈16%. The C5-C6 bond of *N*-acetylgalactosamine is part of the (1-6) linkage in **2**, **4**, **5**, and **6**. The experimentally determined population of the gt conformer of the *N*-acetylgalactosamine residue is ≈10% higher and that of the tg conformer is ≈10% lower compared to the corresponding populations of the β(1-3) linked galactose. The calculated populations deviate by up to 24% from the experimentally determined values. This corresponds to the energy of the gt conformation being calculated too low by 0.6 kcal mol^{-1}. The populations of the hydroxymethyl groups of **1** and **2** agree with previously published data for selectively deuterated core 1 and core 2 structures [42, 49] and other analogous compounds [50].

The glycosidic linkages can be divided into two groups

Table 6. Experimental (steady state NOEs and ROESY) and calculated (GEGOP) relative interglycosidic NOEs (experimental/calculated values) for protons in **1–6** upon irradiation of the respective H1. For the GEGOP computations, t_c, was set to 1.0×10^{-10} s and the temperature was 300 K. **1–5** were measured and calculated at 300 MHz; **6** was measured and calculated at 500 MHz.

Proton pair	**1**	**2**	**3**	**4**	**5**	**6**
β(1-3)-linkage						
H1′/H3	0.67/0.66	0.60/0.64	0.66/0.61	[a]/0.58		0.58/0.64
H1′/H4	0.0/−0.04	0.0/−0.04	0.0/−0.03	[a]/−0.03		0.0/−0.03
β(1-6)-linkage						
H1″/H4		0.0/0.04		[a]/0.02	0.0/0.0	0.0/0.01
H1″/H5		0.14/0.09		[a]/0.07	0.06/0.01	0.09/0.03
H1″/H6R		0.25/0.08		[a]/0.06	0.20/0.10	0.0/0.10
H1″/H6S		0.12/0.06		[a]/0.09	0.34/0.07	0.11/0.07
β(1-4)-linkage						
H1‴/H4″						(s)/0.69
H1‴/H6R″						(w)/0.09
H1‴/H6S″						(m)/0.04

[a] Experimental relative NOE data for **4** could not be obtained due to overlapping of signals; s, strong ROE; m, medium ROE; w, weak ROE. Protons were labelled as in Table 1.

that differ by flexibility, i.e., (1-3) and (1-4) linkages have significantly less flexibility than (1-6) linkages. We were able to obtain interglycosidic NOEs for the (1-3) linkages. The calculated and the experimental interglycosidic relative NOEs show excellent agreement for the (1-3) linkage (Table 6). Furthermore, we did not observe any NOEs to H4 of the *N*-acetylgalactosamine which is also in agreement with the calculations that predict a very small negative NOE. The signals that are important for the determination of the interglycosidic conformation of the (1-4) linkage could not be resolved in 1D NOE spectra due to overlap (Table 1). Thus, only the ROEs are given in Tables 2 and 6. The ROEs correlate qualitatively with the calculated NOEs.

The conformation of the (1-6) linkage is determined by a combination of the rotation around the C5-C6 bond and the rotation of the C1-O1 and O1-C6 bonds. The conformation of the C5-C6 bonds has already been discussed above. The conformation of the C1-O1 and the O1-C6 bonds is determined by the interglycosidic NOEs from H1″ to H6R and H6S. The long range NOEs to H5 and H4 of the *N*-acetylgalactosamine residue are dependent on all three dihedral angles ϕ, ψ, and ω. The agreement between experiment and calculation for the (1-6) linkages is not as good as the agreement found for the (1-3) linkages. The calculated values are significantly lower than the experimental values, indicating that the distances between the H1″ and the H6R, H6S protons are on the average about 12–20% shorter than calculated.

NOE data were also calculated from the geometries of the local minimum energy conformations of **6** (Table 7). A comparison of these effects with the calculated equilibrium NOE values from the MMC simulation demonstrates that several local minimum energy conformations contribute with a different weight to the overall NOE effect. Neither the global minimum nor the local minima alone can account for experimentally determined NOE. However, combinations of the local minima can account for the experimental NOE. It seems obvious that conformation **6b** does not contribute significantly to the ensemble average value. Conformations **6**a, c–g have NOEs that are qualitatively compatible with the experimentally found data. No weighting of these conformers is possible on the basis of the NOEs because the system of equations is underdetermined. It becomes clear that it is not possible for (1-6) linkages to use local minima in order to predict NOE values quantitatively. Dynamic simulations are necessary to approach that goal.

Table 7. Calculated NOE effects for the β(1-6) linkage for local minimum energy conformations of compound **6** computed from the respective minimum geometries. Given are NOE values relative to (H3″ + H5″). Protons were labelled as in Table 1.

Proton pair	*E*/(kcal mol^{-1})							
β(1-6)-linkage	−10.7	−9.0	−8.5	−5.5	−5.4	−4.7	−2.9	MMC
H1″/H4	−1	36	0	1	1	−1	0	1
H1″/H5	2	51	0	1	0	−1	2	3
H1″/H6R	−5	10	−2	33	−5	19	22	10
H1″/H6S	22	−2	20	−8	29	0	−5	7

The oligosaccharide conformations of benzyl glycosides found in this study (Figs 8, 9) are similar to those of core 1 linked to antifreeze glycoprotein [9, 41, 44] and core 2 [42]

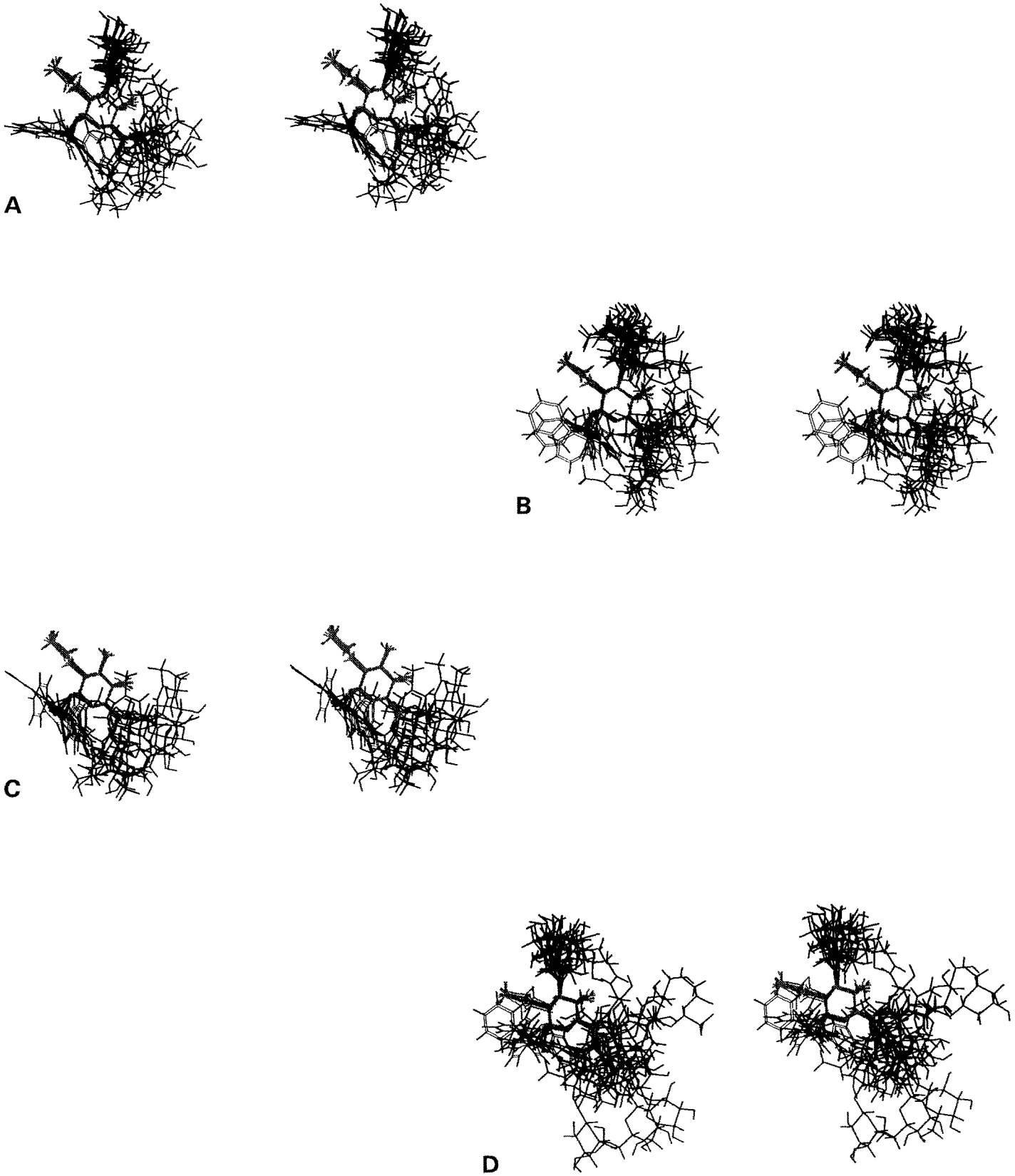

Figure 8. Superimposed stereo plots of ten randomly picked conformations out of 300 000 conformations from the MMC simulations. The conformers were positioned that the central α-D-GalNAc units are overlayed. (A) Superimposed stereo plots of GlcNAcβ(1-6)-[Galβ(1-3)]GalNAcα-Bzl, **2**. (B) Superimposed stereo plots of GlcNAcβ(1-6)[GlcNAcβ(1-3)]GalNAcα-Bzl, **4**. (C) Superimposed stereo plots of GlcNAcβ(1-6)GalNAcα-Bzl, **5**. (D) Superimposed stereo plots of Galβ(1-4)GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-Bzl, **6**.

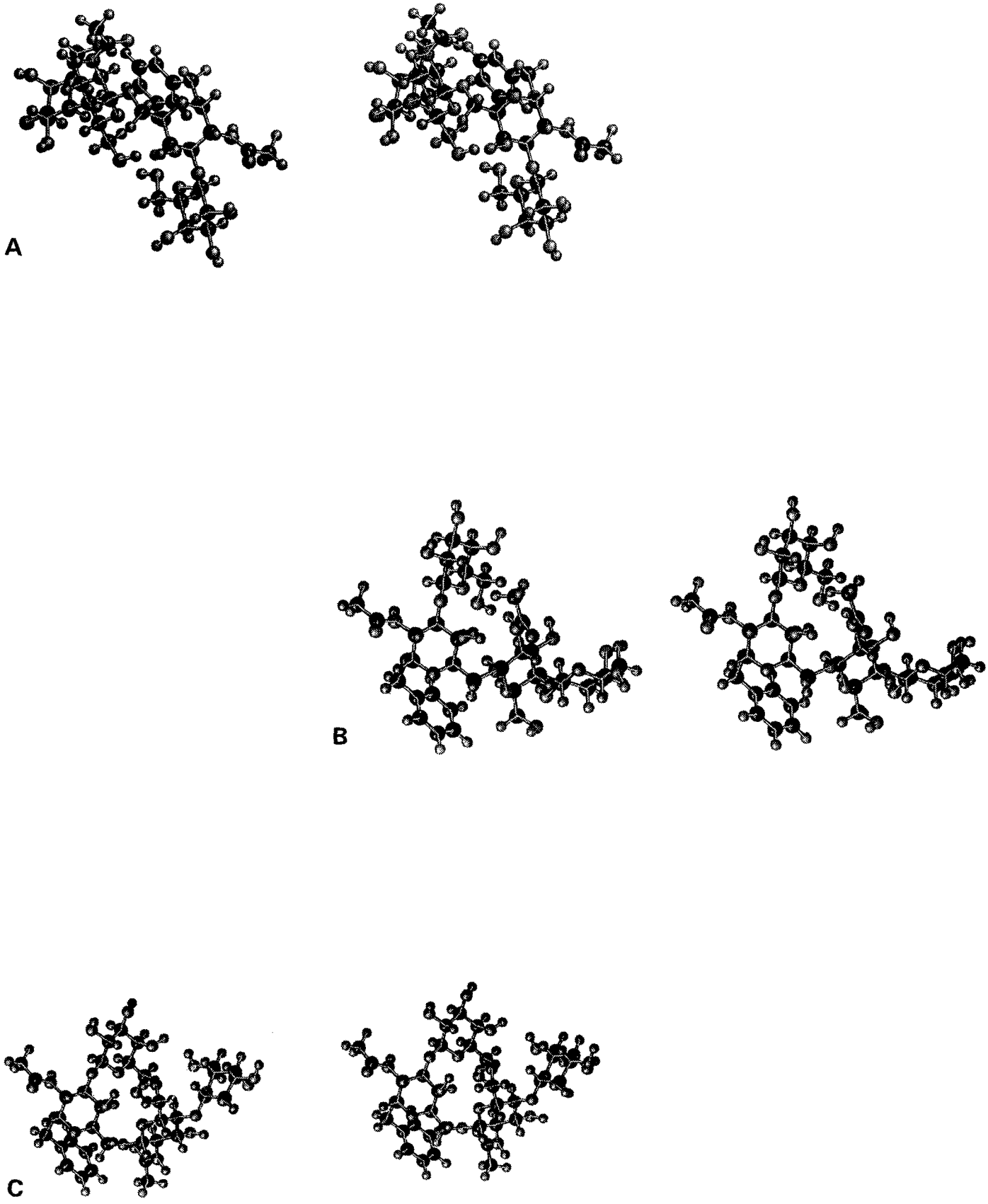

Figure 9. Stereo plots of the three local minima (A) **6**a, (B) **6**b, and (C) **6**c calculated for Galβ(1-4)GlcNAcβ(1-6)[Galβ(1-3)]GalNAcα-Bzl, **6**.

oligosaccharides studied previously by others, indicating that benzyl glycosides are appropriate models.

The data presented suggest that compounds **1**–**6** expose all of those hydroxyls required for further addition of *N*-acetylglucosamine or galactose residues by glycosyltransferases, i.e., the 6-hydroxyls of *N*-acetylgalactosamine in **1** and **3**, the 3-hydroxyls of galactose in **1**, **2**, and **6**, the 3- and 4-hydroxyls of *N*-acetylglucosamine residues in **2**, **3**, **4**, and **5**, as well as the 3-hydroxyl of *N*-acetylgalactosamine in **5**. A possible regulation of *O*-glycan biosynthesis may be based on a combination of accessibility and differential recognition by various competing glycosyltransferases of the entire or partial carbohydrate structure of the substrate. Future studies should determine whether these enzymes indeed preferentially recognize certain oligosaccharide conformations.

Acknowledgements

The authors thank Dr J. W. Dennis and coworkers for supplying the tetrasaccharide compound **6**. This research was supported by grants from the Canadian Cystic Fibrosis Foundation to Harry Schachter, The Hospital for Sick Children, Toronto and from the Medical Research Council of Canada (to IB), a NATO grant for international collaboration, a postdoctoral fellowship by the Deutsche Forschungsgemeinschaft (to AP-K), grants from the National Institutes of Health (1-P41-RR05351-01), Digital Equipment Corporation and the Advanced Computational Methods Center of the University of Georgia (to BM and RS-P), and grant No. CA 35329 from the National Cancer Institute (to KLM).

References

1. Schachter H, Brockhausen I (1992) In *Glycoconjugates, Composition, Structure and Function* (Allen H, Kisailus EC, Eds), pp 263–332. New York: Marcel Dekker.
2. Brockhausen I, Möller G, Merz G, Adermann K, Paulsen H (1990) *Biochemistry* **29**:10206.
3. Carver JP, Cummings D (1987) *Pure Appl Chem* **59**:1465.
4. Lemieux RU, Bock K, Delbaere LTJ, Koto S, Rao SVR (1980) *Can J Chem* **58**:631.
5. Thøgersen H, Lemieux RU, Bock K, Meyer B (1982) *Can J Chem* **60**:44.
6. Duben AJ, Bush CA (1983) *Arch Biochem Biophys* **225**:1.
7. Rao BN, Dua VK, Bush CA (1985) *Biopolymers* **24**:2207.
8. Bush CA, Yan Z-Y, Rao BN (1986) *J Am Chem Soc* **108**: 6168.
9. Rao BN, Bush CA (1987) *Biopolymers* **26**:1227.
10. Nishida Y, Hori H, Meguro H (1988) *Agric Biol Chem* **52**: 887.
11. Bechtel B, Wand AJ, Wroblewski K, Koprowski H, Thurin J (1990) *J Biol Chem* **265**:2028.
12. Shogren R, Gerken TA, Jentoft N (1989) *Biochemistry* **28**:5525.
13. Gerken TA, Jentoft N (1987) *Biochemistry* **26**:4689.
14. Vliegenthart JFG, van Halbeek H, Dorland L (1981) *Pure Appl Chem* **53**:70.
15. Klein A, Carnoy C, Lamblin G, Roussel P, van Kuik JA, de Waard P, Vliegenthart JFG (1991) *Eur J Biochem* **198**: 151.
16. Klein A, Carnoy C, Wieruszeski, Strecker G, Strang A-M, van Halbeek H, Roussel P, Lamblin G (1992) *Biochemistry* **31**:6152.
17. Chai W, Hounsell E, Cashmore GC, Rosankiewicz JR, Feeney J, Lawson AM (1992) *Eur J Biochem* **207**:973.
18. Chai W, Hounsell E, Cashmore GC, Rosankiewicz JR, Bauer CJ, Feeney J, Feizi T, Lawson AM (1992) *Eur J Biochem* **203**:257.
19. Meyer B (1990) *Topics Current Chem* **154**:141.
20. Poppe L, von der Lieth C-W, Dabrowski J (1990) *J Am Chem Soc* **112**:7762.
21. Struike-Prill R, Meyer B (1990) *Eur J Biochem* **194**:903.
22. Meyer B, Zsiska M, Struike-Prill R (1993) *Computer Simulations in Condensed Matter Physics IV* (Landau DP, Mon KB, Schuttler HB eds), Berlin: Springer-Verlag.
23. Brockhausen I, Williams D, Matta KL, Orr J, Schachter H (1983) *Can J Biochem Cell Biol* **61**:1322.
24. Brockhausen I, Matta KL, Orr J, Schachter H (1985) *Biochemistry* **24**:1866.
25. Brockhausen I, Matta KL, Orr J, Schachter H, Koenderman AHL, van den Eijnden DH (1986) *Eur J Biochem* **157**:463.
26. Yazawa S, Abbas SA, Madiyalakan R, Barlow JJ, Matta KL (1986) *Carbohydr Res* **149**:241.
27. Yousefi S, Higgins E, Doaling Z, Pollex-Krüger A, Hindsgaul O, Dennis J (1991) *J Biol Chem* **265**:1772.
28. Flowers HM, Shapiro D (1965) *J Org Chem* **30**:2041.
29. Piscorz CF, Abbas SA, Matta KL (1984) *Carbohydr Res* **126**:115.
30. Abbas SA, Barlow JJ, Matta KL (1983) *Carbohydr Res* **112**: 201.
31. Abbas SA, Barlow JJ, Matta KL (1983) *Carbohydr Res* **113**: 63.
32. Kumar A, Wagner G, Ernst RR, Wüthrich K (1980) *Biochem Biophys Res Commun* **96**:1156.
33. Davis DG, Bax A (1985) *J Am Chem Soc* **107**:7197.
34. Summers MF, Marzilli LG, Bax A (1986) *J Am Chem Soc* **108**:4285.
35. Bax A, Davis DG (1985) *J Magn Reson* **63**:207.
36. Bax A, Davis DG (1985) *J Magn Reson* **65**:355.
37. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) *J Chem. Phys* **21**:1087.
38. Peters T, Meyer B, Struike-Prill R, Somorjai R, Brisson J-R (1993) *Carbohydr Res* **238**:49.
39. Fletcher R, Powell MJD (1963) *Computer J* **6**:163.
40. Brisson J-R, Carver JP (1983) *J Biol Chem* **258**:1431.
41. Bush CA, Feeney RE (1986) *Int J Peptide Protein Res* **28**: 386.
42. Ohrui H, Nishida Y, Hori H, Meguro H (1988) *J Carbohydr Chem* **7**:711.
43. Neuhaus D, Keller J (1986) *J Magn Reson* **68**:568.
44. Homans SW, DeVries AL, Parker SB (1985) *FEBS Lett* **183**:133.

45. Maeji NJ, Inoue Y, Chujo R (1987) *Int J Peptide Protein Res* **29**:699.
46. Gerlt JA, Joungblood VA (1980) *J Am Chem Soc* **102**:7433.
47. Marchessault RH, Perez S (1979) *Biopolymers* **18**:2369.
48. Nishida Y, Hori H, Ohrui H, Meguro H (1987) *Carbohydr Res* **170**:106.
49. Ohrui H, Nishida Y, Higushi H, Hori H, Meguro H (1987) *Can J Chem* **65**:1145.
50. Breg J, Kroon-Batenburg LMJ, Strecker G, Montreuil J, Vliegenthart JFG (1989) *Eur J Biochem* **178**:727.
51. Streefkerk DG, De Bie MJA, Vliegenthart JFG (1973) *Tetrahedron* **29**:833.