Journal of Biomolecular NMR, 5 (1995) 79–94
ESCOM

J-Bio NMR 271

Conformations and internal mobility of a glycopeptide derived from bromelain using molecular dynamics simulations and NOESY analysis

Johannes P.M. Lommerse[a], Loes M.J. Kroon-Batenburg[b], Jan Kroon[b], Johannis P. Kamerling[a] and Johannes F.G. Vliegenthart[a,*]

[a]*Department of Bio-Organic Chemistry, Bijvoet Center and* [b]*Department of Crystal and Structural Chemistry, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands*

Received 24 October 1994
Accepted 20 February 1995

Keywords: Glycopeptide; Oligosaccharide; Molecular dynamics; Glycosidic linkage; Conformation

Summary

The conformation and internal flexibility of a glycopeptide Manα1-6(Xylβ1-2)Manβ1-4GlcNAcβ1-4(Fucα1-3)GlcNAcβ1-N(Asn-Glu-Ser-Ser), prepared from pineapple stem bromelain, have been analyzed using a combination of molecular dynamics (MD) simulations in water with NOESY ^{1}H NMR spectroscopy. Theoretical NOESY cross-peak intensities were calculated by the CROSREL program on the basis of models, obtained from MD simulations, using a full relaxation matrix approach. Special attention was paid to the description of internal flexibility of the hexasaccharide moiety by the use of generalized order parameters, in combination with the application of an individual rotation correlation time for each monosaccharide residue. The tetrapeptide moiety appeared to be very mobile during the MD simulations, which was confirmed by the absence of NOE cross peaks. For the oligosaccharide part a model was developed to estimate characteristic times for large reorientational motions around the glycosidic linkages, associated with conformational transitions. For the Manα1-6Man and the Fucα1-3GlcNAc linkages such a flexibility was found with a characteristic time of 2 ns. In contrast, the Xylβ1-2Manβ1-4GlcNAcβ1-4GlcNAc part of the glycan appears to be relatively rigid.

Introduction

Conformational analysis of oligosaccharide chains has mainly been focused on the determination of a static three-dimensional structure. Recently, flexibility has been recognized to play an important role as part of the conformational behavior of oligosaccharide chains in solution (Cumming and Carver, 1987; Hricovíni et al., 1992; Rice et al., 1993; Rutherford et al., 1993). This feature complicates the conformational analysis, because the experimental data obtained via the usual approaches might be affected by internal motions. Furthermore, the experimental data can often not be interpreted in terms of a single oligosaccharide chain conformation (De Waard et al., 1990).

^{1}H NMR NOESY spectra yield conformational data of oligosaccharide chains in solution, but these data are in terms of average conformations. Data concerning individual conformations and flexibility can be obtained theoretically by employing molecular dynamics (MD) simulations. In this study experimental and theoretical data were combined in order to analyse conformations of a glycopeptide derived from bromelain (E.C. 3.4.22.4), a proteinase occurring in the pineapple stem. The primary structure (Bouwstra et al., 1990) and a spatial model of the studied glycopeptide are presented in Fig. 1.

Nonrestrained MD simulations in aqueous solution have been performed from which theoretical models of the glycopeptide were obtained. For the MD simulations, the tetrapeptide and the carbohydrate moieties were constructed in close proximity of each other in the initial glycopeptide structures, in order to allow interactions between these moieties within reasonable computer time. Long MD runs were performed to obtain average conformations, and conformational probes were extracted from short MD runs.

NOESY buildup intensities were calculated for the average conformations and for the conformational probes, and compared with measured NOESY buildup

*To whom correspondence should be addressed.

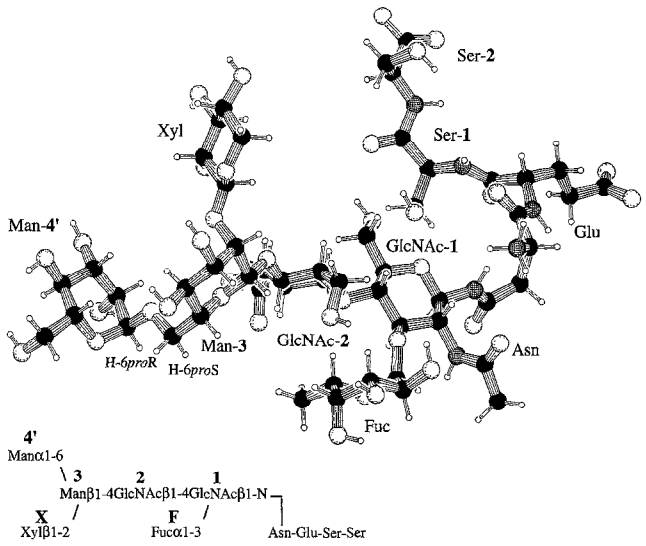

Fig. 1. The primary structure and a spatial model of the glycopeptide of pineapple stem bromelain.

series. The recently developed CROSREL program (Leeflang and Kroon-Batenburg, 1992) made it possible to reconstruct NOESY buildup data from the models, taking flexibility into account. The analysis was focused on the dihedral angles of the various glycosidic linkages in the hexasaccharide moiety and on the dihedral angle of the exocyclic oxymethyl group at the Manα1-6Man linkage.

Methods

^{1}H NMR spectroscopy

The glycopeptide was exchanged twice in 99.9% D_2O (MSD Isotopes, Montreal) with intermediate lyophilisation, and was finally dissolved in 99.96% D_2O (MSD Isotopes). The 500 MHz 2D ^{1}H-^{1}H NOESY spectra were recorded at 300 K on a Bruker AM-500 spectrometer using fixed mixing times of 75, 150, 250 or 350 ms. The spectral width was 3333 Hz in each dimension. The four 512×4K data matrices acquired were processed on a μVAX/VMS cluster using the TRITON NMR software (Bijvoet Center, Department of NMR Spectroscopy, Utrecht University, Utrecht). Multiplication with a $\pi/3$ shifted sine bell and zero-filling were applied prior to Fourier transformation; 512×2K data matrices were obtained. The 2D NOESY spectra were baseline-corrected by a fit with a straight line in ω_1. Integration of cross-peak intensities was performed with a rectangular area around the peak maximum and was corrected for the local baseline area by subtracting an average of integrated identical rectangular areas around the cross peak. The relative background error of peak integration (A_{noise}) was estimated to be 0.005.

The proton signals of the oligosaccharide moiety of the glycopeptide have been assigned previously (Bouwstra et al., 1990). For the stereospecific assignments of the H-6*proS* and H-6*proR* signals of Man-**3** (Fig. 1), the chemical shift data and NOE intensities were compared with those of Manα1-6Manβ1-OMe (Hori et al., 1988) and those of *N*-acetyllactosamine and oligomannose type N-glycans containing the Manα1-6Manβ element (Brisson and Carver, 1983). It could be inferred that the H-6 signal resonating at the highest field and having the largest interresidual NOE intensity with Man-**4'** H-1 is H-6*proS*. Similar assignments were performed for H-6*proS* and H-6*proR* of GlcNAc-**1**, GlcNAc-**2** and Man-**4'** (Nishida et al., 1987).

HSEA calculations

For the oligosaccharide part of the glycopeptide, HSEA calculations were performed using the GEGOP program (Stuike-Prill and Meyer, 1990) on a μVax/VMS computer. The HSEA force field takes into account nonbonded interactions as expressed by the Kitaigorodsky potential energy function, and applies an energy term for the exo-anomeric effect. No solvent molecules are included.

The monosaccharide rings were kept rigid according to coordinates taken from X-ray data. The angles of the glycosidic bonds were set at 117°. The dihedral angles of the glycosidic linkages (ϕ,ψ) and of the C5-C6 bonds (ω) were varied. The ϕ and ψ angles have been defined according to the IUPAC conventions (IUPAC-IUB, 1983); for the glycosidic linkage Monosaccharide A ($1 \rightarrow x$) Monosaccharide B holds: $\phi = \theta(O5_A\text{-}C1_A\text{-}Ox_B\text{-}Cx_B)$ and $\psi = \theta(C1_A\text{-}Ox_B\text{-}Cx_B\text{-}C(x-1)_B)$. In this study ω is defined by θ(O6-C6-C5-O5).

The initial low-energy conformations of the hexasaccharide were combined with the peptide conformation as part of the protein 1LDB (apo-L-lactate dehydrogenase, E.C. 1.1.1.27, Protein Data Bank (PDB)). Energies of the obtained glycopeptide structures were calculated in the GEGOP program, using the ECEPP/2 force field (Némethy et al., 1983) with respect to the peptide moiety, and it was shown that these glycopeptide conformations were possible from an energetic point of view.

The HSEA ϕ,ψ isoenergy contour maps were calculated with increments of 5° in ϕ and ψ of the disaccharides and were plotted at intervals of 1.5 kcal/mol with respect to the calculated global energy minimum of each disaccharide. Using these 'rigid maps' (Ha et al., 1988; French, 1989; Tran et al., 1989), energetically favorable conformations can be derived for each glycosidic linkage within the glycopeptide, provided that only interactions between adjacent monosaccharides occur in the glycopeptide (Imberty et al., 1990).

Molecular dynamics

Molecular dynamics simulations were carried out on local μVax/VMS, Convex 120 and Silicon Graphics 4D/35 computers and on a CRAY Y-MP supercomputer (NCF computer facilities, University of Amsterdam, Amsterdam)

TABLE 1
GROMOS ATOM TYPES AND PARTIAL CHARGES FOR A BROMELAIN-DERIVED GLYCOPEPTIDE

Atom(s)	Atom type	Partial charge
G2(C2), G1(C2)	CS1	0.00
M4'(C2,C3,C4), M3(C3,C4), G2(C3), X(C2,C3,C4), F(C2,C3,C4)	CS1	0.15
M4'(C5), M3(C2,C5), G2(C4,C5), G1(C3,C4,C5), F(C5)	CS1	0.16
G1(C1)	CS1	0.20
M4'(C1), M3(C1), G2(C1), X(C1), F(C1)	CS1	0.40
M4'(C6), G2(C6), G1(C6)	CS2	0.15
X(C5), M3(C6)	CS2	0.16
N(C^{α}), E(C^{α}), S1(C^{α}), S2(C^{α})	CH1	0.00
N(C^{β}), E(C^{β}), E(C^{γ})	CH2	0.00
S1(C^{β}), S2(C^{β})	CH2	0.15
G2(C8), G1(C8), F(C6)	CH3	0.00
E(C^{δ}), S2(C)	C	0.27
G2(C7), G1(C7), N(C), N(C^{γ}), E(C), S1(C)	C	0.38
M4'(HO2,HO3,HO4,HO6), M3(HO3,HO4), G2(HO3,HO6)	HO	0.40
G1(HO6), X(HO2,HO3,HO4), F(HO2,HO3,HO4), S1(H^{γ}), S2(H^{γ})	HO	0.40
G2(H(N2)), G1(H(N2)), N(H^{δ}), E(H), S2(H), S1(H)	H	0.28
N(H1,H2,H3)	H	0.50
M4'(OH2,OH3,OH4,OH6), M3(OH3,OH4), G2(OH3,OH6)	OA	−0.55
G1(OH6), X(OH2,OH3,OH4), F(OH2,OH3,OH4), S2(O^{γ}), S1(O^{γ})	OA	−0.55
M4'(O1,O5), M3(O1,O5), G2(O1,O5), G1(O5), X(O1,O5), F(O1,O5)	OS	−0.36
G2(O7), G1(O7), N(O^{δ}), N(O), E(O), S1(O)	O	−0.38
E(O^{ε}_1,O^{ε}_2), S2(O_1,O_2)	OM	−0.635
G2(N), G1(N), N(N^{δ}), E(N), S2(N), S1(N)	N	−0.28
N(N)	NL	−0.50
HO(water)	HW	0.415
OH(water)	OW	−0.83
Na	NA	1.00
Cl	CL	−1.00

The parameters listed in the table were used for MD simulations in water.
M4' = α-D-Man-**4**'; M3 = β-D-Man-**3**; G2 = β-D-GlcNAc-**2**; G1 = β-D-GlcNAc-**1**; X = β-D-Xyl; F = α-L-Fuc; N = L-Asn; E = L-Glu; S1 = L-Ser-**1**; S2 = L-Ser-**2**. C7 is the C=O carbon of the *N*-acetyl group; C8 is the CH_3 carbon of the *N*-acetyl group; O7 is the C=O oxygen of the *N*-acetyl group.

using the standard force field of the GROMOS program package (Koehler et al., 1987; Van Gunsteren, 1987). GROMOS does not contain a special term to describe the exo-anomeric effect. Jeffrey et al. (1972) demonstrated that a large part of this effect is caused by dipole–dipole interactions. These electrostatic interactions are included well in the GROMOS force field (and are often absent in other force fields), giving reasonable energies at the exo-anomeric bond. Moreover, the accuracy of the ab initio calculations on which the exo-anomeric potentials are based, is very limited (Jeffrey et al., 1978; Wiberg and Murcko, 1989). The adequacy of GROMOS has been demonstrated in a number of papers, therefore we decided not to modify the force field (Koehler et al., 1987,1988; Kroon-Batenburg et al., 1993; Van Eijck et al., 1993).

GROMOS atom types and partial charges of the various atoms of the glycopeptide are listed in Table 1. Positions of hydrogen atoms of extended atoms were calculated after the simulation, using 'ideal' tetrahedral geometries and a C-H bond length of 0.11 nm. Neutral charge groups were applied with a cutoff radius of 1.2 nm. The SPC model (Berendsen et al., 1981) was used to describe the water molecules. The initial conformations of the hexasaccharide were obtained from the HSEA calculations.

The glycopeptide was placed in a truncated octahedron, with a minimum distance of 0.8 nm to the edge of the box. Subsequently, the box was filled with 1050–1300 water molecules, one chloride ion and two sodium ions. Periodic boundary conditions were applied. A cutoff radius of 0.8 nm was used. All bond lengths were constrained and the water molecules were kept rigid using the SHAKE algorithm (Van Gunsteren and Berendsen, 1977).

Before the simulations were started, an energy minimization was performed on the whole system. Initial velocities were obtained from a Maxwellian distribution at 300 K and continued with time steps of 0.002 ps. No constraints were applied, in order to give the glycopeptide 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 by loose separate coupling of solvent and solute to temperature baths with relaxation

TABLE 2
AVERAGE VALUES FOR THE DIHEDRAL ANGLES OF THE GLYCOSIDIC LINKAGES IN THE LONG MD TRAJECTORIES (A, B, C) AND IN THE CONFORMATIONAL PROBES (D1–D10)

Trajectory	t (ps)	M4'(1-6)M3			X(1-2)M3		M3(1-4)G2		G2(1-4)G1		F(1-3)G1	
		ϕ	ψ	ω	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
A	260	117	180	–6	–85	–102	–74	108	–67	125	–102	–111
B	460	104	–177	64	–78	–97	–74	110	–57	128	–145	–141
C	460	117	–179	3	–75	–102	–83	106	–70	126	–125	–126
D1	40	166	177	–52	–62	–93	–61	110	–72	124	–92	–105
D2	70	159	180	–56	–82	–103	–70	114	–70	130	–157	–147
D3	30	156	177	–59	–89	–105	–57	116	–55	132	–141	–140
D4	30	147	176	–49	–88	–108	–59	109	–49	132	–138	–136
D5	50	147	177	–41	–85	–111	–92	84	–64	133	–133	–134
D6	30	76	175	51	–94	–102	–71	99	–76	129	–154	–146
D7	30	83	–161	52	–80	–78	–85	101	–64	134	–143	–134
D8	70	90	–172	62	–69	–89	–97	99	–78	119	–94	–103
D9	35	119	–173	65	–68	–94	–88	103	–75	121	–94	–104
D10	35	142	–142	68	–132	–140	–81	107	–76	122	–90	–99

The abbreviations for the monosaccharide residues are given in Table 1.

times (τ_t) of 0.5 and 2.5 ps, respectively. An equilibrium of the total energy was reached after about 10 ps in each run. The first 20 ps were considered to be equilibration time of the system and were not used for the analysis.

Three long MD simulations, A, B and C, were performed for 280, 480 and 480 ps, respectively, and six short runs for 100 ps each. From these short runs, 10 parts of trajectories were chosen, each at least 30 ps long, in which no transitions at any of the glycosidic linkages occurred (D1–D10, representing 10 different conformational probes). The average dihedral angles of the MD trajectories A, B, C and D1–D10 are listed in Table 2.

Theory of oligosaccharide flexibility

The program CROSREL (Leeflang and Kroon-Batenburg, 1992) applies a full relaxation matrix analysis in order to calculate the NOEs of a system of resonating nuclei in a given 3D structure. The NOE cross-relaxation rate σ_{ij} is described by the spectral density functions $J_n(\omega)$:

$$\sigma_{ij} = C[6J_2(\omega) - J_0(\omega)] \tag{1}$$

with $C = 0.1\gamma^4(h/2\pi)^2(\mu_0/4\pi)^2$.

Motions of vectors between nuclei can be described by a rotation correlation function C(t). In the case of a rigid molecule, which tumbles isotropically with a rotation correlation time τ_o, C(t) can be described by:

$$C(t) = r_{ij}^{-6}\exp(-t/\tau_0) \tag{2a}$$

and the spectral density function becomes:

$$J_n(\omega) = r_{ij}^{-6}\frac{\tau_0}{1+(n\omega\tau_0)^2} \tag{2b}$$

where r_{ij} is the fixed distance between nuclei i and j, and ω is the Larmor frequency at a given magnetic field.

However, oligosaccharides can usually not be considered as rigid molecules. In addition to the overall motion of the molecule with rotation correlation time τ_o, other important motions occur in these molecules, especially at the glycosidic linkages (Cumming and Carver, 1987; Rice et al., 1993). In this study different approaches were applied, accounting for internal motions in an isotropically tumbling oligosaccharide chain based on a more exact description of the spectral density function J(ω).

The rotation correlation function C(t) can be written as a product of overall and internal rotation correlation functions:

$$C(t) = C_o(t) \cdot C_i(t) \tag{3}$$

where $C_i(t)$ describes reorientational motions of a particular vector $\mathbf{r}_{ij}$ with respect to a molecule-fixed system of axes. According to a model-free approach (Lipari and Szabo, 1982; Clore et al., 1990), the rotation correlation function becomes:

$$C(t) = \langle r_{ij}^{-3}\rangle^2\exp\{-t/\tau_o\} \times [S^2 + (1-S_f^2)\exp\{-t/\tau_f\} + (S_f^2-S^2)\exp\{-t/\tau_s\}] \tag{4a}$$

and its spectral density function:

$$J_n(\omega) = \langle r_{ij}^{-3}\rangle^2\left[S^2\frac{\tau_0}{1+(n\omega\tau_0)^2} + (1-S_f^2)\frac{\tau_{0f}}{1+(n\omega\tau_{0f})^2} + (S_f^2-S^2)\frac{\tau_1}{1+(n\omega\tau_1)^2}\right] \tag{4b}$$

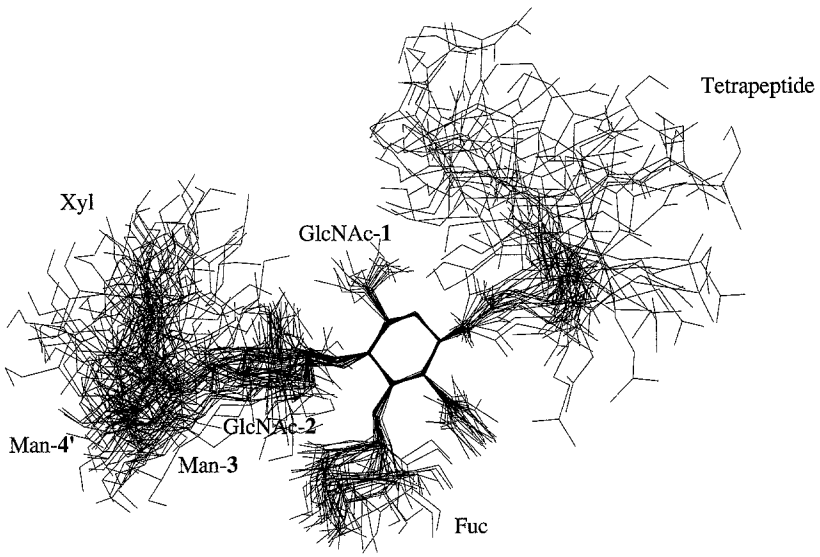

Fig. 2. Snapshots of structures of the glycopeptide, taken at intervals of 20 ps of the simulated trajectory of MD run A. The different conformations are superimposed on the GlcNAc-**1** residue. Note that six *gg* and seven *gt* conformations for the Manα1-6Man linkage are present.

in which motions on different time scales with their individual order parameters are included. τ_o = overall rotation correlation time of the whole molecular system; τ_f = internal rotation correlation time for fast motions of a pair of nuclei; τ_s = internal rotation correlation time for slow motions of a pair of nuclei; S_f^2 = order parameter for fast internal motions; S_s^2 = order parameter for slow internal motions; $S^2 = S_f^2 \cdot S_s^2$; $\tau_{of}^{-1} = \tau_o^{-1} + \tau_f^{-1}$; and $\tau_i^{-1} = \tau_o^{-1} + \tau_s^{-1}$. Strictly speaking, the r^{-3} terms should be included in S^2 (Lipari and Szabo, 1982). As will be seen in Eq. 5, the correlation function $C_i(t)$ from which S^2 is obtained will be normalized, so that S^2 becomes dimensionless. Therefore, r^{-3} terms are explicitly included. The best approach is to use $\langle r^{-3} \rangle^2$ instead of $\langle r^{-6} \rangle$, because the averaging takes place on the fast picosecond MD time scale, which is usually much faster than the time scale of overall tumbling.

The generalized order parameter S_x^2 (x = s,f) of a pair of nuclei is a measure for the spatial restriction of the internal reorientational motion with correlation time τ_x. If $S_x^2 = 1$, the vector between two nuclei is rigidly fixed to the molecule and if $S_x^2 = 0$, the vector between two nuclei is completely unrestricted. S^2 can be determined from NMR experiments using ^{15}N or ^{13}C relaxation rates (Clore et al., 1990). Owing to the limitation of a small amount of material, not enriched with ^{13}C, these relaxation rates cannot be obtained for the glycopeptide. Furthermore, the S^2 values are related to motions of the C-H and N-H vectors, and should not be used as S^2 values for the reorientational motions of other (e.g. ^1H-^1H) vectors.

The generalized order parameters S_f^2 can be obtained from normalized rotation correlation function plots, which are calculated from MD data. The internal rotation correlation function for a pair of protons i and j has been described by Olejniczak et al. (1984):

$$C_i(t) = \sum_{n=-2}^{n=2} \left\langle \frac{Y_n^2(\Theta_{ij}^{mol}(t) \cdot \Phi_{ij}^{mol}(t)) \cdot Y_n^{2*}(\Theta_{ij}^{mol}(0) \cdot \Phi_{ij}^{mol}(0))}{r_{ij}^3(t) \cdot r_{ij}^3(0)} \right\rangle \quad (5)$$

where Y_n are the second-order spherical harmonic functions, and Θ^{mol} and Φ^{mol} the polar angles defining the orientation of the ^1H-^1H vector with respect to the molecule-fixed axes. During the first picoseconds, the internal rotation correlation function will decrease quickly due to the fast internal motions (τ_f), until a plateau is reached at which $\exp(-t/\tau_f) \approx 0$ (see also Fig. 4). Then the function will gradually decrease further by the slow internal motions (τ_s). The plateau value of the normalized internal rotation correlation function gives the order parameter (S_f^2).

In practice, the complete form of Eq. 4b cannot be used, because not all parameters can be determined reliably. Although the MD simulations give access to S_f^2, τ_f and $\langle r_{ij}^{-3} \rangle$, because the underlying motions typically occur on the MD time scale, the S_s^2 is not accessible from these simulations. Therefore, three parameters, S_s^2, τ_s and τ_o, are left to be determined. NOE data that do not contain structural information of interest can be used to fit S_s^2, τ_s and τ_o (Eq. 4b). In view of the number of NOE data (only four) for each pair of protons, obtained at the same Larmor frequency ω, the unknown parameters cannot be obtained independently. Therefore, four approximations were applied, which are described below in order of an increasing number of simplifications.

Method IV Slow internal motions of oligosaccharide chains occur due to conformational changes at each glycosidic linkage k. Fast internal motions are the result of motions within a monosaccharide ring or motions within an energy well on the ϕ,ψ energy surface. Conformational

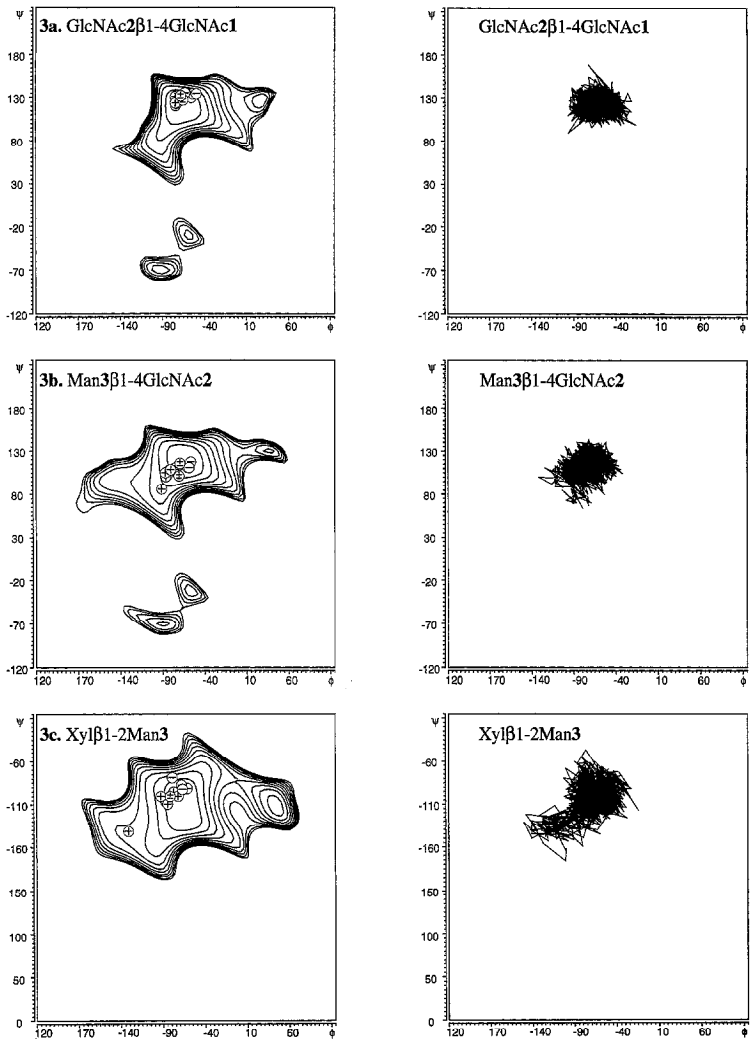

Fig. 3. Left column: HSEA energy contour plots of the disaccharide elements in the glycopeptide. Iso-energy contours are plotted at intervals of 1.5 kcal/mol with respect to the observed global energy minimum. Circles in the plots are positioned on the dihedral angle averages of the long MD trajectories A, B, C and of the conformational probes D1–D10. Within these circles, the signs +, − or ±, respectively, indicate that the calculated interresidual NOE intensities across the glycosidic linkage are too large, too small or reasonably fit ($R_w < 25\%$) the experimental data. The signs of the glycosidic linkage of Manα1-6Man are based on the Man-**4'** H-1,Man-**3** H-6*proS* NOE intensities. Right column: combined ϕ,ψ dihedral angles of all MD trajectories of the glycosidic linkages of the glycopeptide. Each picosecond of the total simulated time after equilibration (1560 ps) has been plotted. For the Manα1-6Man linkage, data are separated into *gg* (ω = −60) and *gt* (ω = 60) conformations.

transitions result in considerable changes in the orientation of an involved pair of nuclei (i,j). Therefore, its order parameter S_s^2 will tend to become (close to) zero. If it is assumed that $S_s^2 \approx 0$ and furthermore that $\tau_f \ll \tau_s$ and $\tau_f \ll \tau_o$, then Eq. 4a can be simplified to:

$$C(t) = \langle r_{ij}^{-3} \rangle^2 S_f^2 \exp(-t/\tau_1) \qquad (6a)$$

and the spectral density function becomes:

$$J_n(\omega) = \langle r_{ij}^{-3} \rangle^2 S_f^2 \frac{\tau_1}{1+(n\omega\tau_1)^2} \qquad (6b)$$

The S_f^2 and r_{ij} values can be obtained from MD simulations. The τ_1 value of conformationally independent pairs of protons, e.g. the intraresidual ^{1}H-^{1}H vectors of monosaccharide k, can be determined by a fit of calculated NOE intensities to the experiment using Eq. 6b. In this relation, the symbol τ_k is defined as the local rotation correlation time for (i,j)-vectors in the monosaccharide k, being in-between O-glycosidic linkages k and k+1 (GlcNAc-**1** has k = 0).

A fit of calculated NOE intensities using Eq. 6b cannot be applied for conformationally dependent vectors. However, if $S_{s,k}^2$ of a vector across the glycosidic linkage is small, the rotation correlation time is almost the same as that of an intraresidual vector of the most mobile residue involved (smallest τ_k).

In order to calculate all NOE intensities of a theoretical model using Method IV, first the local rotation correlation times (τ_k) within the molecule have to be determined.

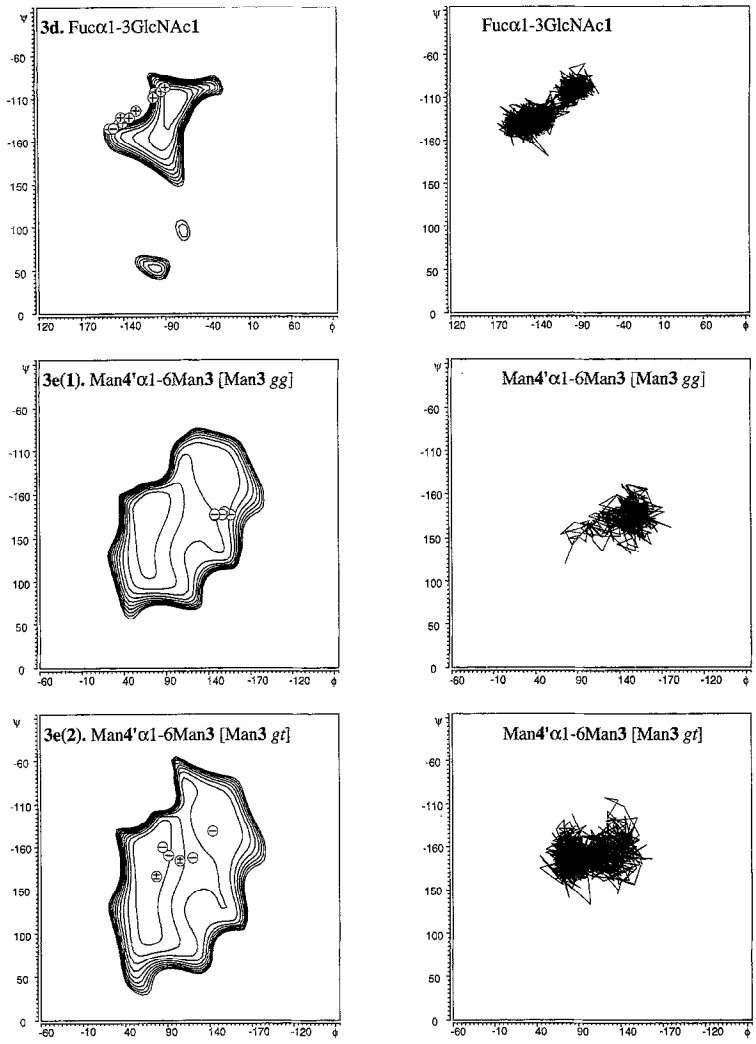

Fig. 3. (continued).

Method III If the generalized order parameters are not available, one could implicitly include the fast internal motions in the rotation correlation times of the different monosaccharide residues. The fit of the calculated NOE intensities within monosaccharide residue k with the experimental data will result in an effective rotation correlation time, $\tau_{eff,k}$. The $\tau_{eff,k}$ value describes all motions of monosaccharide residue k:

$$J_n(\omega) = \langle r_{ij}^{-3} \rangle^2 \frac{\tau_{eff,k}}{1+(n\omega\tau_{eff,k})^2} \qquad (7)$$

Like in Method IV, the interresidual vectors are described by $\tau_{eff,k}$ of the most mobile residue. This approach is equivalent with application of the initial NOE buildup rate analysis for structural analysis of oligosaccharides (Kumar et al., 1981; Scarsdale et al., 1986), using implicitly effective rotation correlation times for the various monosaccharide residues. By fitting an intraresidual NOE buildup curve of a ^{1}H-^{1}H pair having a known interproton distance to the experimental curve, the initial NOE buildup rate is obtained, which represents the effective rotation correlation time $\tau_{eff,k}$. Subsequently, this rate can be used for the determination of distances between related protons, which is especially important for interresidual pairs. The initial rate analysis does not account for spin-diffusion effects. In this study, spin-diffusion effects are included by the use of the full relaxation matrix approach.

Method II If it is assumed that only the overall rotation correlation time τ_o is needed for the description of the slow motions of the molecule and that the generalized order parameters account for (fast) internal motions, then Eq. 6b becomes:

$$J_n(\omega) = S_f^2 \langle r_{ij}^{-3} \rangle^2 \frac{\tau_0}{1+(n\omega\tau_0)^2} \qquad (8)$$

This approach is often applied for the calculation of NOE intensities within large (bio)molecules, in order to deal with fast internal motions (Koning et al., 1991).

Method I The most simple approach is based on the

TABLE 3
RELATIVE MAGNETIZATION OF THE ANOMERIC PROTONS AT MIXING TIME $\tau_m = 0$, CALCULATED BY EXTRAPOLATION OF THE TOTAL MAGNETIZATION IN THE NOESY SPECTRA

Proton	M_0	Relative error (%)
Man-**4'** H-1	4.89	2.2
Xyl H-1	5.41	7.9
Man-**3** H-1	4.87	7.1
GlcNAc-**2** H-1	4.75	6.9
Fuc H-1	3.60	2.5
GlcNAc-**1** H-1	3.71	4.0

assumption that the molecule is in a single, rigid, average conformation, experiencing only one rotation correlation time describing the overall motion:

$$J_n(\omega) = \langle r_{ij}^{-3} \rangle^2 \frac{\tau_0}{1+(n\omega\tau_0)^2} \qquad (9)$$

It has to be noted that this average rigid model does not completely comply to rigidity, because average distances from MD trajectories are used.

Results and Discussion

The tetrapeptide moiety

It was not possible to apply a NOESY analysis of the tetrapeptide part of the glycopeptide, due to the lack of NOE cross peaks. This is caused by fast reorientational motions within the tetrapeptide moiety, described by a correlation time τ_c, leading to NOE buildup rates of about zero ($\omega\tau_c \approx 1$). The mobility of the peptide moiety could be inferred from a ROESY analysis of the C^β protons of glutamic acid ($\tau_c \approx 0.3$ ns; data not shown). MD simulations of the glycopeptide demonstrate this large flexibility as well. In Fig. 2, 13 snapshots of glycopeptide structures taken from MD run A are given, superimposed on the GlcNAc-**1** residue. The peptide moiety extends in a large area of space with low density, when compared to the oligosaccharide chain. Because of the absence of strong interactions between peptide and carbohydrate chain, the flexibility of the former is hardly reduced.

The MD simulations frequently show a close proximity of the peptide moiety to the oligosaccharide chain, as inferred from several hydrogen bonds. These hydrogen bonds, defined by (i) a distance between the donor hydrogen and the acceptor of less than 0.24 nm; and (ii) a bond angle of more than 120° (Van Eijck and Kroon, 1989), are observed at various instances spread over the trajectories. Typical examples are the hydrogen bond between asparagine and GlcNAc-**1** (N N^δ-H ··· **G1** O7), formed for 5.5% of the total simulated time, and that between Ser-**1** and GlcNAc-**1** (**S1** O^γ-H ··· **G1** O6), formed for 13.9% of the time.

NOE analysis

The CROSREL program was applied to transform structural information of the oligosaccharide part, obtained from MD simulations, into NOE intensities. In this way the complete long MD trajectories A, B and C, as well as the conformational probes D1–D10, generated by the six short MD simulations, were investigated.

The average values of the dihedral angles of the glycosidic linkages (Table 2) have been marked in the HSEA iso-energy contour plots of the disaccharide elements in the glycopeptide, as depicted in Figs. 3a–e, left column. The variations of the dihedral angles of all MD trajectories have been plotted in Figs. 3a–e, right column. Average distance matrices, each consisting of 62 glycopeptide protons, were calculated for the MD trajectories by $r = \langle r^{-3} \rangle^{-1/3}$ averaging. The distance matrices were used as theoretical models for the oligosaccharide chain.

TABLE 4
S_f^2 VALUES OBTAINED FROM ROTATION CORRELATION FUNCTION PLOTS OF MD TRAJECTORIES

^{1}H-^{1}H pair		$S_f^{2\,a}$			$\langle S^2 \rangle^b$
		A	B	C	
Man-**4'** H-1	Man-**4'** H-2	0.68 (0.22)	0.55 (0.10)	0.70 (0.10)	0.64
Man-**4'** H-1	Man-**3** H-6*proR*	0.57 (0.07)	0.61 (0.09)	0.65 (0.02)	0.61
Man-**4'** H-1	Man-**3** H-6*proS*	0.42 (0.16)	0.58 (0.06)	0.58 (0.02)	0.53
Xyl H-1	Xyl $H\text{-}5_{ax}$	0.71 (0.07)	0.62 (0.06)	0.56 (0.04)	0.63
Xyl H-1	Man-**3** H-2	0.62 (0.09)	0.62 (0.07)	0.54 (0.01)	0.59
Man-**3** H-1	Man-**3** H-2	0.85 (0.04)	0.78 (0.05)	0.84 (0.01)	0.82
Man-**3** H-1	GlcNAc-**2** H-4	0.78 (0.05)	0.77 (0.03)	0.67 (0.02)	0.74
GlcNAc-**2** H-1	GlcNAc-**2** H-5	0.81 (0.06)	0.82 (0.03)	0.79 (0.09)	0.81
GlcNAc-**2** H-1	GlcNAc-**1** H-4	0.78 (0.11)	0.85 (0.08)	0.81 (0.02)	0.81
Fuc H-1	Fuc H-2	0.81 (0.07)	0.81 (0.04)	0.79 (0.07)	0.80
Fuc H-1	GlcNAc-**1** H-3	0.62 (0.17)	0.70 (0.11)	0.61 (0.11)	0.64
GlcNAc-**1** H-1	GlcNAc-**1** H-5	0.85 (0.03)	0.82 (0.04)	0.84 (0.03)	0.84

[a] Values between brackets are standard deviations: s.d. $= \sqrt{\Sigma(x_i - \langle x \rangle)^2/(N-1)}$.
[b] Mean S^2 values, used for the conformational probes (D1–D10).

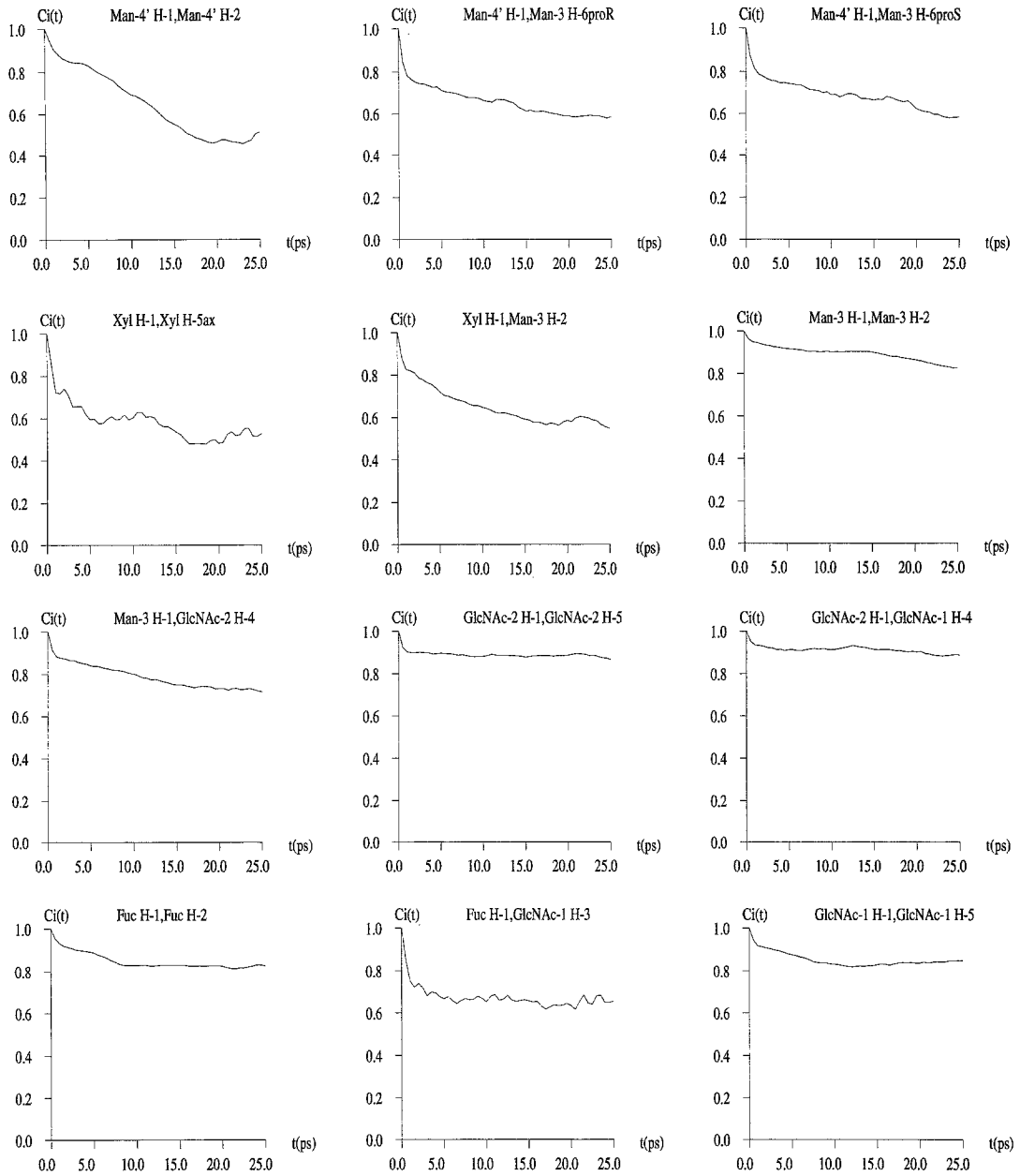

Fig. 4. Rotation correlation function plots of selected ^{1}H-^{1}H pairs. The functions are calculated from the trajectory of run B from 20 to 70 ps, of which 25 ps are plotted.

The CROSREL program offers the possibility to employ both cross- and diagonal peak intensities by $\mathbf{M}_0$ scaling. A preliminary approach using $\mathbf{M}_0$ scaling, applying an extrapolation of the total quantity of magnetization as an exponential function of the mixing time (τ_m), resulted in different magnitudes of the magnetization at $\tau_m = 0$ for the various anomeric protons (Table 3). The spread in $\mathbf{M}_0$ values is probably caused by interference of overlapping small sample impurities and/or large base-line errors at the diagonal. Because $\mathbf{M}_0$ scaling could not be applied, Obs scaling was used instead, using only the cross-peak intensities. Starting with a reasonable choice for an overall τ_o, an overall scaling factor (o.s.f.) was determined from a fit with observed NOE intensities (see below) for each theoretical model. Subsequently, a new τ_o (or τ_k's in the case of Methods III and IV) was determined. Reoptimalizations of the o.s.f. only changed it by 1–2%. Therefore, the same o.s.f. was used in all calculation methods.

In Table 4 the S_f^2 values of the most important ^{1}H-^{1}H pairs have been listed for the MD trajectories A, B and C, derived from normalized internal rotation correlation function plots (Fig. 4). The S_f^2 values were only calculated for proton pairs having their protons separated by 5 Å or less. S_f^2 values of proton pairs with larger interproton distances, actually having little influence on the calculation of the NOE intensities, were assigned the value 1, since these values tend to approach 1. The MD trajectories were divided into parts of 50 ps. Average S_f^2 values

TABLE 5

THE OVERALL SCALING FACTORS, ^{1}H-^{1}H PAIRS USED FOR FITTING OVERALL ROTATION CORRELATION TIMES AND INDIVIDUAL ROTATION CORRELATION TIMES OF EACH MONOSACCHARIDE RESIDUE FOR THE VARIOUS MD TRAJECTORIES

Traj.	o.s.f.	Method applied and intraresidual ^{1}H-^{1}H pairs used (see text)											
		I	II	III					IV				
				M4' H-1,2	X H-1,5$_{ax}$	M3 H-1,2	G2 H-1,5	G1 H-1,5	F H-1,2	M4' H-1,2	X H-1,5$_{ax}$	M3 H-1,2	G2 H-1,5	G1 H-1,5	F H-1,2
A	7.391	0.67	0.73	0.50	0.66	0.75	0.85	0.81	0.59	0.55	0.72	0.80	0.93	0.87	0.62
B	7.595	0.67	0.77	0.51	0.69	0.79	0.79	0.84	0.59	0.62	0.83	0.88	0.84	0.92	0.63
C	7.396	0.67	0.77	0.51	0.67	0.76	0.84	0.87	0.59	0.56	0.81	0.82	0.97	0.97	0.65
D1	7.552	0.68	0.77	0.52	0.66	0.75	0.94	0.78	0.59	0.58	0.77	0.82	1.05	0.86	0.64
D2	7.644	0.67	0.76	0.53	0.67	0.75	0.88	0.79	0.57	0.59	0.77	0.80	0.96	0.87	0.63
D3	8.086	0.69	0.77	0.54	0.71	0.79	0.81	0.73	0.61	0.61	0.82	0.85	0.89	0.79	0.67
D4	7.970	0.70	0.79	0.53	0.75	0.78	0.86	0.76	0.62	0.61	0.91	0.84	0.95	0.83	0.67
D5	7.959	0.70	0.79	0.51	0.71	0.80	0.80	0.78	0.64	0.61	0.92	0.86	0.87	0.86	0.69
D6	7.713	0.67	0.77	0.52	0.73	0.82	0.88	0.79	0.58	0.58	0.81	0.88	0.98	0.86	0.63
D7	8.142	0.69	0.78	0.49	0.64	0.87	0.93	0.69	0.63	0.60	0.87	0.98	1.05	0.75	0.69
D8	7.134	0.66	0.75	0.51	0.67	0.78	0.96	0.86	0.58	0.56	0.73	0.80	1.06	0.95	0.62
D9	7.335	0.67	0.76	0.54	0.75	0.86	0.97	0.83	0.58	0.57	0.79	0.85	1.07	0.86	0.62
D10	7.957	0.71	0.79	0.57	0.69	0.82	0.94	0.78	0.59	0.64	0.86	0.91	1.03	1.11	0.63

Overall scaling factors (o.s.f.) were calculated using Method I. Overall rotation correlation times (ns) were fitted using ^{1}H-^{1}H pairs with Methods I and II, and individual rotation correlation times (ns) of each monosaccharide unit were obtained using Methods III and IV.

were calculated from the internal rotation correlation functions of these parts, taking average values between 10 and 20 ps. Similar S_f^2 values for the same ^{1}H-^{1}H pairs were obtained from the three long MD trajectories (see Table 4). S_f^2 turns out to be largely independent of the trajectory. Therefore, average S_f^2 values of the long MD trajectories A, B and C were used for the calculation of NOE intensities (Methods II and IV) of the conformational probes D1–D10.

The following six intraresidual ^{1}H-^{1}H NOE intensities were used for the determination of rotation correlation times from a fit with the experimental data: Man-**4'** H-1,2; Xyl H-1,5$_{ax}$; Man-**3** H-1,2; GlcNAc-**2** H-1,5; GlcNAc-**1** H-1,5 and Fuc H-1,2. The calculated overall rotation correlation times τ_o (Methods I and II), the effective rotation correlation times $\tau_{eff,k}$ (Method III), and the local rotation correlation times τ_k (Method IV) are listed in Table 5. They are mainly determined by intraresidual ^{1}H-^{1}H distances and depend only slightly on conformation. The largest variations of the local rotation correlation times are found for the inner-core monosaccharide residues Man-**3**, GlcNAc-**2** and GlcNAc-**1**, due to deformation of the monosaccharide rings in several oligosaccharide chain conformations. Apparently, deformed inner-core monosaccharide rings only slowly relax to average optimal geometries as obtained after the long MD runs A, B and C.

In order to evaluate the effect of a leakage rate R_L (Leeflang and Kroon-Batenburg, 1992), Method IV was extended by using $R_L = 1.0\ s^{-1}$ (Method IV + R_L).

In the CROSREL program a weighted residual (R_w) factor is used (Leeflang and Kroon-Batenburg, 1992), in order to evaluate the quality of the fit of a set of calculated NOE intensities with the experimental data. The R_w value summarises in one figure the fit of the calculated NOE values (NOE_{ij}^{calc}) to all or a selected group of experimental values of the measured NOESY spectra (NOE_{ij}^{exp}):

$$R_w = \left\{ \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\tau_m}^{m} w_{ij}(\tau_m)\cdot\left[NOE_{ij}^{calc}(\tau_m) - NOE_{ij}^{obs}(\tau_m)\right]^2}{\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\tau_m}^{m} w_{ij}(\tau_m)\cdot\left[NOE_{ij}^{obs}(\tau_m)\right]^2} \right\}^{1/2} \quad (10)$$

with the weight factors $w_{ij}(\tau_m)$:

$$w_{ij}(\tau_m) = \frac{1}{A_{noise} + \left|NOE_{ij}^{obs}(\tau_m)\right|} \quad (11)$$

$R_w = 0$ in case of an exact fit. The R_w values of the six intraresidual ^{1}H-^{1}H pairs indicate the ability of the four methods to accurately fit theoretical models to the experimental data. These R_w values are listed in Table 6. In comparison with the isotropic, rigid model (Method I), an improved fit for the six intraresidual ^{1}H-^{1}H NOE intensities is obtained when taking into account fast internal motions by using S_f^2 (Method II). Fitting rotation correlation times for each intraresidual NOE intensity (Methods III and IV) results in the best overall R_w value. The difference in the quality of these fits between Methods III

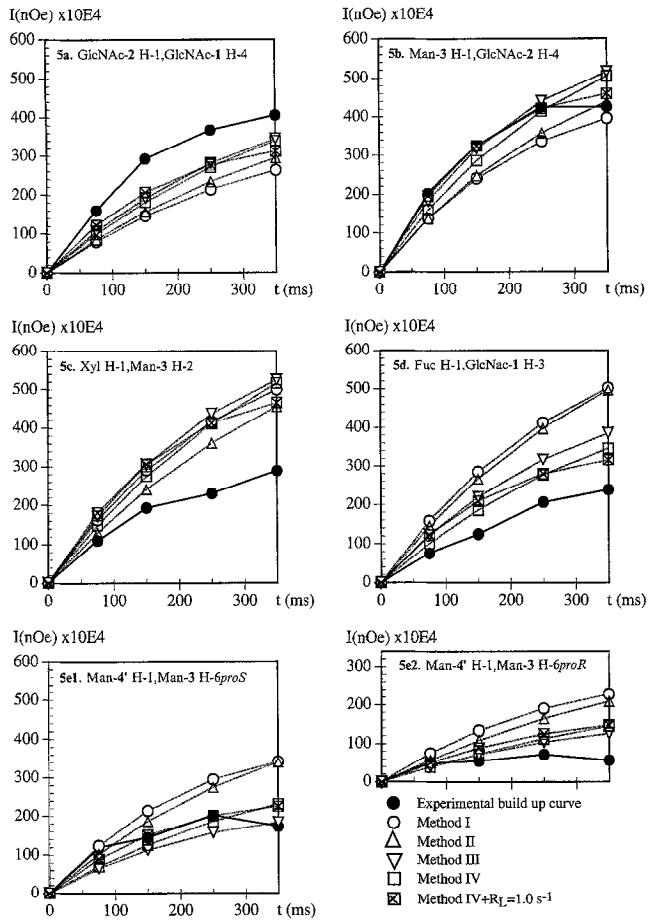

Fig. 5. Buildup curves of interresidual NOE intensities, both experimental and calculated from data of MD run B, applying the different calculation methods.

and IV is explained by the shape of the calculated NOE buildup curves (Fig. 5): with S_f^2 values, the calculated rotation correlation times are larger and the NOE buildup curves become less bent, resulting in a worse fit. When a leakage rate is included (Method IV+R_L), a strong levelling off appears, improving the fit.

The calculated interresidual NOE intensities are dependent on the conformation of the oligosaccharide chain. Only those interresidual ^{1}H-^{1}H pairs were analyzed that could be measured reliably: Man-**4'** H-1,Man-**3** H-6*proS*; Man-**4'** H-1,Man-**3** H-6*proR*; Xyl H-1,Man-**3** H-2; Man-**3** H-1,GlcNAc-**2** H-4; GlcNAc-**2** H-1,GlcNAc-**1** H-4 and Fuc H-1,GlcNAc-**1** H-3. NOEs related to Fuc H-5 were not used, because the Fuc H-5 signal is very close to the suppressed HDO signal in the NMR experiments. The Man-**3** H-1,GlcNAc-**2** H-6*proR* and GlcNAc-**2** H-1,GlcNAc-**1** H-6*proR* NOEs were not used, because the rotamer distributions of the GlcNAc-**2** and GlcNAc-**1** hydroxymethyl groups could not be simulated reliably.

The fit of the interresidual NOE intensities with the experimental data gives an indication for the validity of the respective model. This can be expressed by the R_w values of the interresidual ^{1}H-^{1}H pairs, which are listed in Table 7 for each calculation method and each theoretical model. The R_w values of the individual interresidual ^{1}H-^{1}H pairs, applying Method IV, are listed in Table 8 for each theoretical model. From Tables 7 and 8 it can be seen that the individual conformational probes D1–D10 fit the experimental data less well than do the longer MD trajectories A, B and C. Furthermore, as an example, in Fig. 5 both the experimental and the calculated NOE buildup curves of MD run B are given for the analyzed interresidual ^{1}H-^{1}H pairs.

The glycosidic linkages

The dihedral angles of the glycosidic linkages of the GlcNAc**2**β1-4GlcNAc**1** element only changed in a limited range ($\Delta\phi \pm 20$, $\Delta\psi \pm 15$) around $\phi,\psi = -70,125$ during the

TABLE 6

R_w VALUES FOR THE FIT OF SIX INTRARESIDUAL ^{1}H-^{1}H PAIRS[a] OF THE OLIGOSACCHARIDE MOIETY OF THE GLYCOPEPTIDE FOR THE DIFFERENT MD TRAJECTORIES USING THE VARIOUS CALCULATION METHODS

Trajectory	Method				
	I	II	III	IV	IV+R_L
A	0.352	0.320	0.103	0.126	0.078
B	0.339	0.263	0.102	0.124	0.078
C	0.356	0.342	0.097	0.128	0.080
D1	0.345	0.317	0.099	0.125	0.076
D2	0.330	0.296	0.095	0.121	0.077
D3	0.258	0.230	0.093	0.120	0.077
D4	0.286	0.258	0.094	0.122	0.075
D5	0.272	0.230	0.092	0.116	0.086
D6	0.359	0.311	0.100	0.123	0.079
D7	0.331	0.295	0.104	0.128	0.082
D8	0.406	0.358	0.096	0.132	0.085
D9	0.391	0.355	0.101	0.123	0.084
D10	0.314	0.296	0.094	0.115	0.091

[a] The six ^{1}H-^{1}H pairs are listed in Table 5.

MD simulations (Fig. 3a, right), suggesting that this part of the molecule is rather stable. The micro-environment at this linkage can be compared with methyl β-cellobioside (Kroon-Batenburg et al., 1993), or with Lewis^x (Le^x) type sequences (Wormald et al., 1991; Miller et al., 1992; Rutherford et al., 1994) which contain Galβ1-4[Fucα1-3]GlcNAcβ units, or with the GalNAcβ1-4[Fucα1-3]GlcNAcβ structure (Bergwerff et al., 1993). For these compounds similar ϕ,ψ values have been obtained from calculations employing different force fields in combination with experimental NOESY or ROESY data. Furthermore, the appearance of only a single conformation in these studies points to rigidity at this linkage.

The interresidual GlcNAc-**2** H-1,GlcNAc-**1** H-4 NOE intensities show that the intensities calculated from run B fit the experimental values reasonably well (Fig. 5a). All calculation methods which include flexibility within the molecule (Methods II, III, IV and IV + R_L) result in better fits than does the average rigid model (Method I).

An important alternative to the simulated sampling around the minimum ϕ,ψ = −70,125 would be a simulation centered around ϕ,ψ = −100,−60 (see Fig. 3a, left). This conformation has a short GlcNAc-**2** H-1,GlcNAc-**1** H-3 distance and, in principle, this conformation can also give rise to the observed cross peak, because the H-3 and H-4 signals of GlcNAc-**1** overlap (Bouwstra et al., 1990). The intensity of the GlcNAc-**2** H-1,GlcNAc-**1** H-3 cross peak will be larger than that of the GlcNAc-**2** H-1, GlcNAc-**1** H-4 cross peak, so the occurrence of this conformation is possible during a short period of time. However, in methyl β-cellobioside it appeared that this conformation is probably nonexistent (Kroon-Batenburg et al., 1993).

The energy contour plot of the Man**3**β1-4GlcNAc**2** linkage is similar to that of the GlcNAc**2**β1-4GlcNAc**1** linkage (Figs. 3a and b, left), and the same ϕ,ψ = −70,125 conformation is exclusively observed during the MD simulations (Fig. 3b, left). This conformation might be stabilized by the hydrogen bond G**2** O3-H···M**3** O5 (34.2% of the total simulated time, see also Homans, 1990; Imberty et al., 1990; Kroon-Batenburg et al., 1993). This linkage appears to be equally rigid as the GlcNAc**2**β1-4GlcNAc**1** linkage, which can be inferred from the slow internal rotation correlation time, as will be explained later. In contrast to the situation at GlcNAc-**1**, the H-3 and H-4 signals of GlcNAc-**2** do not overlap, so that the interresidual cross peak across the Man**3**β1-4GlcNAc**2** linkage arises exclusively from the Man-**3** H-1,GlcNAc-**2** H-4 contact. The Man-**3** H-1,GlcNAc-**2** H-4 NOE intensities perfectly fit the experimental data using Methods III, IV or IV + R_L (Fig. 5b), confirming the existence of a single accessible region of conformational space.

The glycosidic linkage of Xylβ1-2Man**3** is more flexible during the simulated time than the two linkages described above (Fig. 3c, right). The ϕ,ψ energy contour plot of this linkage (Fig. 3c, left) shows that ϕ and ψ can change in a large region around the global minimum (ϕ,ψ = −70, −105), at the cost of a small increase in energy. This flexibility is represented by small S_f^2 values across the glycosidic linkage ($S_f^2 \approx 0.6$, see Table 4).

Additionally, the possibility of other conformations has been investigated. The HSEA energy contour plot only points to another conformation near ϕ,ψ = 50,−120, but this conformation turns out to be unstable when simulated in water using the GROMOS force field.

The simulated region near ϕ,ψ = −132,−140 (conformational probe D10) is only visited for a short period of time and results in calculated NOE intensities that are too large (Table 8). Therefore, this conformation can be discarded.

TABLE 7

R_w VALUES FOR THE FIT OF SIX INTERRESIDUAL ^{1}H-^{1}H PAIRS[a] OF THE OLIGOSACCHARIDE MOIETY OF THE GLYCOPEPTIDE FOR THE DIFFERENT MD TRAJECTORIES USING THE VARIOUS CALCULATION METHODS

Trajectory	Method				
	I	II	III	IV	IV + R_L
A	0.889	0.616	0.560	0.373	0.371
B	0.667	0.577	0.434	0.394	0.383
C	1.048	0.848	0.740	0.576	0.566
D1	1.046	0.926	0.707	0.633	0.633
D2	0.837	0.697	0.543	0.492	0.489
D3	1.218	1.002	0.939	0.841	0.843
D4	1.352	1.139	1.065	0.866	0.865
D5	1.484	1.302	1.245	1.024	1.028
D6	0.656	0.516	0.699	0.574	0.570
D7	0.857	0.741	0.770	0.709	0.635
D8	1.157	1.084	0.904	1.061	0.965
D9	1.316	1.185	1.003	0.877	0.863
D10	2.325	2.169	1.527	1.597	1.564

[a] The six ^{1}H-^{1}H pairs are listed in Table 8.

TABLE 8
R_w VALUES[a] FOR THE FIT OF INDIVIDUAL INTERRESIDUAL NOE INTENSITIES ACROSS THE GLYCOSIDIC LINKAGES, APPLYING METHOD IV, FOR THE VARIOUS MD TRAJECTORIES

Trajectory	Man4' H-1-Man3 H-6*proS*	Man4' H-1-Man3 H-6*proR*	Xyl H-1-Man3 H-2	Man3 H-1-GlcNAc2 H-4	GlcNAc2 H-1-GlcNAc1 H-4	Fuc H-1-GlcNAc1 H-3
A	–0.384	–/+0.409	–/+0.152	–/+0.154	–/+0.155	+0.896
B	–/+0.257	+0.841	+0.679	–/+0.144	–0.284	+0.419
C	–0.426	+0.777	+0.305	+0.400	+0.344	+1.263
D1	–0.758	+0.991	–0.498	–0.373	+0.322	+1.247
D2	–0.628	+1.797	+0.623	–0.234	–/+0.128	–0.256
D3	–0.555	+2.014	+1.319	–0.551	–0.603	+0.750
D4	–0.448	+1.650	+1.143	–0.384	–0.643	+1.489
D5	–0.499	+1.313	+1.077	+0.340	–0.421	+2.370
D6	–/+0.303	–/+0.425	+1.295	–/+0.145	+0.331	–0.137
D7	–0.463	+1.025	–0.435	+0.875	–0.551	+0.918
D8	–0.390	–/+0.649	–0.334	+1.189	+1.123	+1.617
D9	–0.425	+1.972	–0.290	+0.973	+0.678	+1.302
D10	–0.575	+7.518	+1.764	+0.592	+0.381	+0.707

[a] Minus, plus or minus/plus indicates that a calculated NOE buildup curve is smaller than, larger than or goes through the experimental measured curve, respectively.

The average values of the dihedral angles for runs A and B are nearly identical (Table 2). However, whereas the interresidual Xyl H-1,Man-3 H-2 NOE intensities of run A fit the experimental data very well, those of run B do not (Table 8, Fig. 5c). This difference demonstrates that a limited simulation time might lead to inadequate sampling.

The MD simulations revealed two stable conformations for the glycosidic linkage Fucα1-3GlcNAc**1** at average values of $\phi,\psi = -90,-100$ and $\phi,\psi = -140,-135$, respectively (see Fig. 3d, right). These minimum energy conformations coincide roughly with the two lowest energy conformations for the free disaccharide, as proposed by Imberty et al. (1991) ($\phi,\psi = -70,-95$ and $-145,-145$, respectively). The small differences between the ϕ,ψ values of the glycopeptide element and those of the free disaccharide might be due to interference of GlcNAc-**2**. The transitions between the observed main conformations are partially mediated by hydrogen bond formation between N-H of GlcNAc-**1** and O2 of Fuc (30%). A correlation exists between the conformations of the Fucα1-3GlcNAc**1** and GlcNAc**2**β1-4GlcNAc**1** linkages: ϕ of GlcNAc**2**β1-4GlcNAc**1** shifts from –60 to –80 when the ϕ,ψ combination of the Fucα1-3GlcNAc**1** linkage changes from –140,–135 to –90,–100. The Fucα1-3GlcNAc**1** linkage appears to be rather flexible, as concluded on the basis of the small S_f^2 values (Table 4), the observed conformational transitions and the time-scale analysis of these transitions.

Two interresidual NOEs relevant for this linkage were observed, i.e., a strong Fuc H-1,GlcNAc-**1** H-3 NOE and a weak (10–30% of the former) Fuc H-5,GlcNAc-**2** H-2 NOE. The calculated NOE intensities of Fuc H-1,GlcNAc-**1** H-3 for run B, being mostly in the $\phi,\psi = -140,-135$ region, are too large in comparison with the experimental data (Fig. 5d). Nevertheless, run B gives clearly better results than runs A or C (see Table 8). On the basis of this NOE, it can be concluded that the conformations in the region $\phi,\psi = -140,-135$ are important. The Fuc H-5,GlcNAc-**2** H-2 contact is nearly absent in this region and seems to be too strong in the $\phi,\psi = -90,-100$ region. This suggests that both conformations are populated in the glycopeptide. The slow internal rotation correlation time indicates that mobility is present with a characteristic time of 2 ns (see below), which can probably be associated with the motion between the $\phi,\psi = -140,-135$ and $-90,-100$ conformations. The MD simulations should last several times this correlation time in order to obtain an adequate sampling of the conformational space for this linkage.

In view of these findings, it is of interest to compare the results with reports on compounds in which the GlcNAc-**2** residue is replaced either by Gal (Le^x-type sequences: Wormald et al., 1991; Miller et al., 1992; Rutherford et al., 1994) or by GalNAc (Bergwerff et al., 1993). In these studies, only one stable conformation near $\phi,\psi = -90,-100$ has been reported for the Fucα1-3GlcNAc element. The various results are based on Fuc H-1, GlcNAc H-3 and Fuc H-5,Gal/GalNAc H-2 NOE or ROE contacts, in combination with MD simulations without the explicit inclusion of water molecules. The present results are at variance with these results in two respects. First, the measured Fuc H-5,GlcNAc-**2** H-2 contact is weak, which is in agreement with Bergwerff et al. (1993) for Fuc H-5,GalNAc H-2, and also with Van Dam et al. (1994) for Fuc H-5,Gal H-2 in poly-$Lewis^x$, but in disagreement with Miller et al. (1993) and Rutherford et al. (1994), who found medium to strong contacts for Fuc H-5,Gal H-2 (about 85% of the Fuc H-1,GlcNAc H-3 intensity). Secondly, all previous calculations have

been performed in vacuo, resulting in a strong preference for a 'stacked' interaction between Fuc and Gal or GalNAc (due to dispersion interactions). The present study includes water molecules explicitly. It is seen that the importance of the $\phi,\psi = -90,-100$ conformation is reduced, leading to a significant population in the $\phi,\psi = -140,-135$ region. The difference between the present and most other force field calculations could also be caused by the treatment of the anomeric effect, which is more pronounced in other force fields. Only verification against many more experimental data and different molecular species will allow final conclusions on this aspect.

In general, the fits of interresidual NOE intensities with the experimental data are about equal when using Methods III or IV, because the S_f^2 values of the inter- and intraresidual ^{1}H-^{1}H pairs for a monosaccharide residue are about the same in most cases. However, the S_f^2 values of the intraresidual Fuc H-1,H-2 pair and of the interresidual Fuc H-1,GlcNAc-**1** H-3 pair are quite different (Table 4). The fast motions are rather restricted within the monosaccharide residue as compared to the same motions across the glycosidic linkage. Therefore, the calculated NOE intensities are smaller and in better agreement with the measured NOE intensities when using Method IV compared to Method III. An important consequence of this observation is that the initial NOE build-up rate analysis cannot be applied to determine the interresidual Fuc H-1,GlcNAc-**1** H-3 distance. This analysis implicitly assumes that the effective rotation correlation times of the intra- and interresidual ^{1}H-^{1}H pairs for a monosaccharide residue are equal. In this case the effective rotation correlation time of the intraresidual Fuc H-1,H-2 pair will not be the same as that of the interresidual Fuc H-1,GlcNAc-**1** H-3 pair. The initial rate analysis will predict a Fuc H-1,GlcNAc-**1** H-3 distance that is too long.

Because of the extra bond connecting Man-**4'** with Man-**3** at the Man**4'**α1-6Man**3** linkage, the Man-**4'** residue is expected to be the most mobile (Homans et al., 1987). The flexibility during the simulated time is indeed apparent in the ϕ,ψ plots (Fig. 3e, right), and can also be inferred from the small S_f^2 values on the picosecond time scale (Table 5). The conformational transitions between regions around minimum-energy ϕ,ψ-conformations point to flexibility on a longer time scale.

As reported earlier (Bouwstra et al., 1990), the ^{1}H NMR data of Man-**3** show small values for both $^3J_{H5,H6proS}$ (~1 Hz) and $^3J_{H5,H6proR}$ (2.1 Hz), indicating exclusively (>98%) the *gg* conformation ($\omega = -60°$). However, during the MD simulations seven conformational changes from *gg* to *gt* ($\omega = 60°$) or from *gt* to *gg* were observed. For example, in MD runs A and C the relative amounts of *gg* and *gt* were about equal (see Table 2). The *tg* ($\omega = 180°$) rotamer did not appear. It is important to note that the ω transitions occur simultaneously with transitions for ϕ from $\phi,\omega = 140,-60$ to 85,60 and vice versa. This dual change results in a limited change of the position of Man-**4'** with respect to Man-**3**, as can be inferred from superposition of the different conformations. For example, in simulation A, a conformational change from *gg* to *gt* appeared twice, whereas the superimposed conformations exhibit only a small change in the position of Man-**4'** (Fig. 2). In both the *gg* (ϕ,ψ in the region 140,–170) and *gt* (ϕ,ψ in the region 85,–175) conformations of the α1-6 linkage, Man-**4'** is close to GlcNAc-**2**, and the Man-**4'** O4 and O3 atoms are involved in several different weak hydrogen bonds with the GlcNAc-**2** O3 atom and *N*-acetyl group.

In the *gg* rotamer no conformations have been found in the region $\phi,\psi = 85,-175$. However, taking into account the latter region, calculated NOE intensities of Man-**4'** H-1,Man-**3** H-6*proS* and Man-**4'** H-1,Man-**3** H-6*proR* are in good agreement with the experimental data as found for the *gt* rotamer. In this context it should be noted that in the *gg* conformation of long MD simulations of Man**4'**α-1-6[Man**4**α1-3][Xylβ1-2]Man**3**β1-4GlcNAc**2**β1-4GlcNAc-**1**β1-OMe (Lommerse, 1994) the $\phi,\psi = 85,-175$ conformation is visited. There are no repulsive interactions in the $\phi,\psi,\omega = 85,-175,-60$ region, because Man-**4'** stretches out in free space. This can also be inferred from the HSEA surface. Therefore, it can be concluded that for the *gg* conformation of the glycopeptide the sampling is not adequate. This shows that, in general, flexible linkages should be simulated for periods that are at least a few times the characteristic time of conformational transitions, in order to obtain a complete sampling. This observation indicates that for the Man**4'**α1-6Man**3** linkage several nanoseconds of simulation are needed (see below). It is concluded that both $\phi,\psi = 85,-175$ and $\phi,\psi = 140,-170$ are populated and, from coupling constants, that exclusively the *gg* conformation is present. Additional studies are needed to explain the finding of both the *gg* and the *gt* conformation during the MD simulations.

The calculated NOE intensities of run B show that it is important to treat internal motions explicitly in order to obtain a reasonable agreement with the experimental data (Fig. 5e(1,2)). Therefore, a local treatment of reorientational motions (Methods III and IV) is necessary.

Flexibility of the carbohydrate chain and estimates of characteristic times of conformational transitions

Fast unrestricted motions and slow internal motions play important roles in an oligosaccharide. Method IV accounts for these motions by the generalized order parameter S_f^2 and by local rotation correlation times. This combination results in the best agreement of the calculated interresidual NOE intensities with the experimental data, as shown by the R_w values (Tables 6 and 7). The use of the leakage rate $R_L = 1.0\ s^{-1}$ does not afford an improved fit to experimental data, but it changes the

TABLE 9
CALCULATED SLOW INTERNAL ROTATION CORRELATION TIMES OF THE GLYCOSIDIC LINKAGES IN THE GLYCOPEPTIDE[a]

Glycosidic linkage	$\tau_{s,k}$ (ns)
Man4'α1-6Man3	2
Xylβ1-2Man3	17
Man3β1-4GlcNAc2	11
GlcNAc2β1-4GlcNAc1	42
Fucα1-3GlcNAc1	2

[a] Calculated rotation correlation times were obtained using local rotation correlation times according to Method IV, utilizing averaged values of MD runs A, B and C.

shape of the buildup curve. A clear levelling off appears, which is more consistent with the experimental data (Figs. 5a–e).

The GlcNAc-**1** residue has the largest local rotation correlation time τ_k. For residues that are situated at a position more remote from GlcNAc-**1**, τ_k becomes smaller (Table 4). These findings are consistent with a picture where GlcNAc-**1** is near the center of mass of the molecule and rotates with the overall rotation correlation time τ_0. Each connected residue will experience additional reorientational motion due to the glycosidic conformational changes. Assuming that the accompanying S_s^2 is almost zero, the GlcNAc-**2** residue will have a characteristic rotation correlation time of τ_1, where $\tau_1^{-1} = \tau_0^{-1} + \tau_{s,1}^{-1}$, and $\tau_{s,1}^{-1}$ is the rotation correlation time characteristic for conformational changes of the glycosidic linkage GlcNAc2β1-4GlcNAc1. The Man-**3** residue will have a local reorientational motion τ_2, with $\tau_2^{-1} = \tau_1^{-1} + \tau_{s,2}^{-1}$ or $\tau_2^{-1} = \tau_0^{-1} + \tau_{s,1}^{-1} + \tau_{s,2}^{-1}$, where $\tau_{s,2}^{-1}$ is the characteristic reorientational time for conformational changes of the linkage Man3β1-4GlcNAc2. Therefore, even if all $\tau_{s,k}$ values would be identical for all glycosidic linkages, the more peripheral residues have much smaller τ_k values. On the other hand, the successive τ_k values can be used to estimate the various $\tau_{s,k}$ values.

The $\tau_{s,k}$ values have been calculated for the glycopeptide, using $\tau_{s,k}^{-1} = \tau_k^{-1} - \tau_{k-1}^{-1}$ for the averages of runs A, B and C, and are listed in Table 9. The uncertainty in $\tau_{s,k}$ becomes larger when τ_{k-1} and τ_k are nearly equal ($\tau_{s,k}^{-1}$ rapidly goes to infinity). Therefore it is not possible to distinguish the larger $\tau_{s,k}$ values of the Xylβ1-2Man3β1-4GlcNAc2β1-4GlcNAc1 fragment, but it is clear that this sequence is significantly more rigid than the Man4'α1-6Man3 and Fucα1-3GlcNAc1 linkages on a longer time scale. The flexibility of the latter linkage is striking (see above), because for the structurally similar Le^x-type sequences, it is often concluded that this linkage is rigid. Interestingly, the Xylβ1-2Man3 linkage appears to be rather immobile on a longer time scale ($\tau_{s,k} = 17$ ns), in contrast to the local flexibility on a short time scale ($S_f^2 = 0.59$).

Conclusions

Previously, a single conformational model for the glycopeptide derived from pineapple stem bromelain has been proposed, based on an initial NOE buildup rate analysis in combination with HSEA ϕ,ψ energy surfaces (Bouwstra et al., 1990). In this paper, the flexibility within the glycopeptide was taken into account by utilization of molecular dynamics simulations and NOE full relaxation matrix calculations, including generalized order parameters and a local rotation correlation time for each monosaccharide residue. The present results show that the earlier reported conformation is only one of the possible conformations.

The glycosidic linkages of the Xylβ1-2Man3β1-4GlcNAc2β1-4GlcNAc1 part of the molecule are rather rigid and occupy predominantly one conformation, which is very similar to that of the previous model (Bouwstra et al., 1990). The Man4'α1-6Man3 and Fucα1-3GlcNAc1 linkages turn out to be flexible on a picosecond time scale as well as on a nanosecond time scale. With respect to the Fucα1-3GlcNAc1 linkage, the results are not in agreement with the data from conformational studies on Le^x-type sequences, which have closely related micro-environments at this linkage. The conclusions concerning the rigidity of the Le^x-type sequences were based on calculations in vacuum, possibly overrating the preference for a 'stacked' position of Fuc to Gal.

The model used in this paper allows the calculation of characteristic times of large reorientational motions around the glycosidic linkages, associated with transitions between regions around minimum-energy ϕ,ψ-conformations. For the Man4'α1-6Man3 and Fucα1-3GlcNAc1 linkages, these motions occur on a nanosecond time scale (typically 2 ns). No evidence for such motions was found for the other linkages.

Acknowledgements

This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO). We thank the authors of GROMOS for the use of their program package, Dr. B.P. van Eijck for his assistance to use this program and the National Computing Facilities (NCF) Foundation for performing part of the MD calculations on the CRAY Y-MP supercomputer.

References

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

Bergwerff, A.A., Van Kuik, J.A., Schiphorst, W.E.C.M., Koeleman, C.A.M., Van den Eijnden, D.H., Kamerling, J.P. and Vliegenthart, J.F.G. (1993) *FEBS Lett.*, **334**, 133–138.

Bouwstra, J.B., Spoelstra, E.C., De Waard, P., Leeflang, B.R., Kamerling, J.P. and Vliegenthart, J.F.G. (1990) *Eur. J. Biochem.*, **190**, 113–122.

Brisson, J.-R. and Carver, J.P. (1983) *Biochemistry*, **22**, 3680–3686.

Clore, G.M., Szabo, A., Bax, A., Kay, L.E., Driscoll, P.C. and Gronenborn, A.M. (1990) *J. Am. Chem. Soc.*, **112**, 4989–4991.

Cumming, D.A. and Carver, J.P. (1987) *Biochemistry*, **26**, 6664–6676.

De Waard, P., Boelens, R., Vuister, G.W. and Vliegenthart, J.F.G. (1990) *J. Am. Chem. Soc.*, **112**, 3232–3234.

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

Ha, S.N., Madsen, L.J. and Brady, J.W. (1988) *Biopolymers*, **27**, 1927–1952.

Homans, S.W., Pastore, A., Dwek, R.A. and Rademacher, T.W. (1987) *Biochemistry*, **26**, 6649–6655.

Homans, S.W. (1990) *Biochemistry*, **29**, 9110–9118.

Hori, H., Nishida, Y., Ohrui, H., Meguro, H. and Uzawa, J. (1988) *Tetrahedron Lett.*, **29**, 4457–4460.

Hricovíni, M., Shah, R.N. and Carver, J.P. (1992) *Biochemistry*, 31, 10018–10023.

Imberty, A., Gerber, S., Tran, V. and Pérez, S. (1990) *Glycoconjugate J.*, **7**, 27–54.

Imberty, A., Delage, M.-M., Bourne, Y., Cambillau, C. and Pérez, S. (1991) *Glycoconjugate J.*, **8**, 456–483.

IUPAC-IUB Joint Commission on Biochemical Nomenclature (1983) *Eur. J. Biochem.*, **131**, 5–7.

Jeffrey, G.A., Pople, J.A. and Radom, L. (1972) *Carbohydr. Res.*, **25**, 117–131.

Jeffrey, G.A., Pople, J.A., Binkley, J.S. and Vishveshwara, S. (1978) *J. Am. Chem. Soc.*, **100**, 373–379.

Koehler, J.E.H., Saenger, W. and Van Gunsteren, W.F. (1987) *Eur. Biophys. J.*, **15**, 197–210.

Koehler, J.E.H., Saenger, W. and Van Gunsteren, W.F. (1988) *J. Biomol. Struct. Dyns.*, **6**, 181–198.

Koning, T.M.G., Van der Marel, G.A., Van Boom, J.H. and Kaptein, R. (1991) *Biochemistry*, **30**, 3787–3797.

Kroon-Batenburg, L.M.J., Kroon, J., Leeflang, B.R. and Vliegenthart, J.F.G. (1993) *Carbohydr. Res.*, **245**, 21–42.

Kumar, A., Wagner, G., Ernst, R.R. and Wüthrich, K. (1981) *J. Am. Chem. Soc.*, **103**, 3654–3658.

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

Lipari, G. and Szabo, A. (1982) *J. Am. Chem. Soc.*, **104**, 4546–4559.

Lommerse, J.P.M. (1994) Ph.D. Thesis, Utrecht University, Utrecht, pp. 161–209.

Miller, K.E., Mukhopadhyay, C., Cagas, P. and Bush, C.A. (1992) *Biochemistry*, **31**, 6703–6709.

Némethy, G., Pottle, M.S. and Scheraga, H.A. (1983) *J. Phys. Chem.*, **87**, 1883–1887.

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

Olejniczak, E.T., Dobson, C.M., Karplus, M. and Levy, R.M. (1984) *J. Am. Chem. Soc.*, **106**, 1923–1930.

Rice, K.G., Wu, P., Brand, L. and Lee, Y.C. (1993) *Curr. Opin. Struct. Biol.*, **3**, 669–674.

Rutherford, T.J., Partridge, J., Dweller, C.T. and Homans, S.W. (1993) *Biochemistry*, **32**, 12715–12724.

Rutherford, T.J., Spackman, D.G., Simpson, P.J. and Homans, S.W. (1994) *Glycobiology*, **4**, 59–68.

Scarsdale, J.N., Yu, R.K. and Prestegard, J.H. (1986) *J. Am. Chem. Soc.*, **108**, 6778–6784.

Stuike-Prill, R. and Meyer, B. (1990) *Eur. J. Biochem.*, **194**, 903–919.

Tran, V., Buleon, A., Imberty, A. and Pérez, S. (1989) *Biopolymers*, **28**, 679–690.

Van Dam, G.J., Bergwerff, A.A., Thomas-Oates, J.E., Rotmans, J.P., Kamerling, J.P., Vliegenthart, J.F.G. and Deelder, A.M. (1994) *Eur. J. Biochem.*, **225**, 467–482.

Van Eijck, B.P. and Kroon, J. (1989) *J. Mol. Struct. (THEOCHEM)*, **195**, 133–146.

Van Eijck, B.P., Hooft, R.W.W. and Kroon, J. (1993) *J. Phys. Chem.*, **97**, 12093–12099.

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, Groningen.

Wiberg, K.B. and Murcko, M.A. (1989) *J. Am. Chem. Soc.*, **111**, 4821–4828.

Wormald, M.R., Edge, C.J. and Dwek, R.A. (1991) *Biochem. Biophys. Res. Commun.*, **180**, 1214–1221.