Structure and Dynamics of a Peptidoglycan Monomer in Aqueous Solution Using NMR Spectroscopy and Simulated Annealing Calculations

Hans Matter,[†] László Szilágyi,*[,‡] Péter Forgó,[‡] Željko Marinić,[§] and Branimir Klaić[§]

Contribution from TRIPOS GmbH, Martin-Kollar-Strasse 15, D-81829 München, Germany, Department of Organic Chemistry, L. Kossuth University, H-4010 Debrecen, Hungary, and Department of Organic Chemistry and Biochemistry, Ruder Bošković Institute, HR-10001 Zagreb, P.O. Box 1016, Croatia

Received August 8, 1996[⊗]

Abstract: The conformation of the peptidoglycan monomer (PGM) from *Brevibacterium divaricatum* was determined in aqueous solution using a combined approach by 2D NMR spectroscopy, restrained simulated annealing, and molecular dynamics (MD) calculations. MD simulations in water without experimental constraints provided insights into the structure and dynamics of this glycopeptide. Hierarchical cluster analyses for conformer classifications were performed using a global molecular shape descriptor (CoMFA steric fields). Principal component analysis was subsequently employed to extract orthogonal principal conformational properties. Correlated dihedral angle mobilities were identified using a dynamic cross correlation map. The calculation of radial distribution functions for all polar protons of the molecule leads to additional information about the solvation of PGM in a protic solvent, while autocorrelation functions for dihedral angle fluctuations were used to monitor dynamic processes in different regions. From simulated annealing, a set of 11 conformers was obtained, all characterized by a well-defined extended N-terminal peptide part additionally stabilized by the bound disaccharide; the C-terminal part, on the other hand, exhibits more conformational flexibility in agreement with experimental data and MD simulations. The disaccharide conformation is in agreement with the conformational minimum computed for the model disaccharide 3-*O*-Me-4-*O*-*β*GlcNAc-αMurNAc using various force fields. Not only the interglycosidic bond but also the glycopeptide linkage exists in a single, well-defined conformation, for which no conformational changes can be detected during the MD simulations. In contrast, conflicting experimental data for the *N*-acetyl group of GlcNAc could be explained using a conformer population analysis based on ROE intensities and coupling constants accounting for a conformational equilibrium with one dominantly populated rotamer.

1. Introduction

Peptidoglycans are components of the cell walls of all bacteria. The repeating unit, peptidoglycan monomer (PGM), obtained by digestion with lysozyme of the linear non-cross-linked peptidoglycan oligomer from a penicillin-treated *Brevibacterium divaricatum* mutant is a glycopeptide with the following structure: [2-deoxy-2-acetamido-3-*O*-(D-ethyl-1-carbonyl)-4-*O*-(2-acetamido-2-deoxy-*β*-D-glucopyranosyl)-D-glucopyranose]-L-alanyl-D-isoglutaminyl-[*meso*-α(L)-*ϵ*(D)-diaminopimeloyl]-D-alanyl-D-alanine (GlcNAc-MurNAc-L-Ala-D-iGln-m-Dap-D-Ala-D-Ala).[1] In contrast to other smaller peptidoglycan fragments, PGM shows no pyrogenic effect.[2]

Potential pharmacological properties of peptidoglycan fragments resulted in extensive investigations of their biological effects and stimulated synthetic and conformational studies.[3] Muramyl dipeptide (MDP, MurNac-L-Ala-D-iGln) is the smallest compound in this class to show immunomodulating properties.[4] On the basis of ^{1}H NMR studies, Fermandjian et al.[5,6] proposed that MDP possesses in DMSO solution an S-shaped structure composed of two adjacent *β*-turns. The first of them is stabilized by a hydrogen bond between AlaNH and the acetamido carbonyl of MurNAc whereas the second *β*-turn involves a hydrogen bond between the α-carboxamide of D-iGln and the D-Lac carbonyl. Harb et al.[7] found some evidence, using molecular dynamics calculations, for the proposed S-shaped conformation. On the other hand, NMR-restrained modeling calculations by Boulanger et al.[8] did not support the existence of a second *β*-turn whereas formation of the first one was confirmed for MDP in DMSO solution.

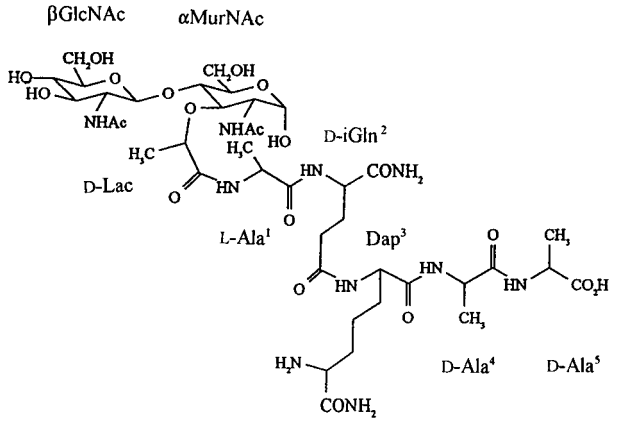
It is well-known that small linear peptides do not form definite structure elements in isotropic media (e.g., DMSO or water solution), unless they are stabilized by specific sequences,[9] by incorporation of nonstandard amino acids,[10] or by environmental effects due to specific solvents such as 2,2,2-trifluoroethanol (TFE),[11] cryomixtures or micelles.[12] However, glycosylation of amino acids can lead to the formation of local structural

* Author to whom correspondence should be addressed at L. Kossuth University (fax ++36 52-310-936; e-mail lszilagyi@tigris.klte.hu).
† TRIPOS GmbH.
‡ L. Kossuth University.
§ Ruder Bošković Institute.
⊗ Abstract published in *Advance ACS Abstracts*, February 15, 1997.

(1) (a) Keglević, D.; Ladešić, B.; Tomašić, J.; Valinger, Z.; Naumski, R. *Biochim. Biophys. Acta* **1979**, *585*, 273−281. (b) Klaić, B.; Ljubić, B.; Metelko, B.; Pongračić, M. *Carbohydr. Res.* **1983**, *123*, 168−172.
(2) Suskovic, B.; Vajtner, Z.; Naumski, R. *Tetrahedron* **1991**, *47*, 8407−8416.
(3) Baschang, G. *Tetrahedron* **1989**, *45*, 6331−6360.
(4) (a) Ellouz, F.; Adam, A.; Ciorbaru, R.; Lederer, E. *Biochem. Biophys. Res. Commun.* **1974**, *59*, 1317−1325. (b) Lefrancier, P.; Choay, J.; Derrien, M.; Lederman, I. *Int. J. Pept. Protein Res.* **1977**, *9*, 249−257.
(5) Fermandjian, S.; Perly, B.; Level, M.; Lefrancier, P. *Carbohydr. Res.* **1987**, *162*, 23−32.
(6) Sizun, P.; Perly, B.; Level, M.; Lefrancier, P.; Fermandjian, S. *Tetrahedron* **1988**, *44*, 991−997.
(7) Harb, V.; Mavri, J.; Kidrić, J.; Hadži, D. *Croat. Chem. Acta* **1991**, *64*, 551−559.
(8) Boulanger, Y.; Tu, Y.; Ratovelomanana, V.; Purisima, E.; Hanessian, S. *Tetrahedron* **1992**, *48*, 8855−8868.

S0002-7863(96)02776-X CCC: $14.00

Scheme 1. The Chemical Structure of the Peptidoglycan Monomer (PGM) (the α-anomeric form of the reducing end MurNAc residue is shown)



elements and rigidification of the peptide backbone in specific cases.[13] For PGM, there were some indications that certain structural elements may be dominant in solution (sufficient dispersion of peptide NH chemical shifts, occurence of some long-range and peptide-saccharide ROEs), which initiated this study to gain insight into structure and dynamics of this glycopeptide.

Preliminary investigation of of PGM in DMSO solution using one- and two-dimensional ^{1}H NMR methods has been reported.[14] Since the aqueous medium is closer to natural condition than DMSO, we present here results of a study into the conformation of PGM in H_2O. The conformational preferences were studied using a combination of 2D NMR spectroscopy, systematic searches, restrained simulated annealing (SA), and molecular dynamics (MD) calculations. Experimentally derived distance constraints from 2D ROESY spectra in H_2O were used for subsequent structure refinement. Molecular dynamics simulations of selected PGM conformers in water were performed to understand structural and dynamic properties of this molecule in aqueous solution. It was possible to identify conformational families using hierarchical cluster analysis techniques. As conformational descriptors, different local and global conformational measures were used and compared: (1) characteristic local torsion angles (peptide backbone and glycosidic bond torsions), (2) pairwise global root mean square (rms) deviation of backbone atoms, and (3) global molecular shape descriptors. All classifications presented here were generated using novel global molecular shape descriptors.

2. Methods

2.1. Experimental Procedures. Sample. The purity of PGM obtained from a penicillin-treated *Brevibacterium divaricatum* mutant was checked by thin-layer chromatography.[1] For the NMR measurements, two samples were prepared by dissolving amounts (12 mg) of this material in 0.5 mL of 99.96% D_2O or $H_2O(9)/D_2O(1)$ (v/v) mixture to give solutions of 25 mM total concentrations. The pH was adjusted to 4.6 using 0.1 M DCl or HCl solutions. A combination glass electrode was used to measure the pH directly in the 5 mm NMR tubes; the value read for the D_2O sample was corrected for the deuterium isotope effect.[15]

NMR Measurements. NMR experiments have been carried out using Bruker AC-400, DRX-500, Varian Gemini 300, or Unity Plus 500 spectrometers. The sample temperature was 298 K. Amide ^{1}H chemical shift temperature coefficients have been determined over a temperature range between 278 and 313 K using 5 K increments. Phase sensitive DQF-COSY and TOCSY (mixing times of 15, 30, and 60 ms) measurements have been performed for the sample in H_2O/D_2O with low-power preirradiation to suppress the residual water signal. No irradiation was applied during t_1. The raw data sets typically consisted of 2K × 512 complex data points. Rotating frame NOE (ROESY)[16,17] spectra in the phase sensitive mode (TPPI) were obtained using a CW spin lock of 3.57 kHz strength. For water suppression, the irradiation frequency was set to the water resonance during the relaxation delay and then shifted to a different frequency for the duration of the spinlock and data acquisition times. HOHAHA effects[17,18] and *J*-relayed crosspeaks[19,20] were identified by repeating the experiment using a different offset frequency for the spinlock field.[19] ROE buildup curves were constructed from experiments run with 40, 80, 120, 160, and 250 ms for the mixing time; values up to 160 ms were found to be in the linear regime. The crosspeak intensities were determined by volume integration from the base plane corrected ROESY spectrum recorded with 160 ms mixing time. The ^{13}C,^{1}H COSY experiment for ^{13}C NMR assignments was run for the sample in D_2O, long-range ^{13}C,^{1}H NMR correlations were obtained in H_2O/D_2O from a gradient HMBC[21] measurement using a selective ^{13}C read pulse on the carbonyl region, and ^{15}N NMR chemical shifts were determined from a gradient ^{15}N HSQC experiment.[22] ^{1}H chemical shifts and coupling constants were extracted from a resolution-enhanced 1D spectrum recorded at 750 MHz and from 500 MHz 1D TOCSY[23] spectra.

Quantitative information on interproton distances for the structure determination was obtained from analyzing the 2D ROESY spectrum of PGM in H_2O. The individually assigned crosspeaks were checked for the absence of artifacts, such as HOHAHA contributions[17,18] or relayed ROE/HOHAHA transfer,[19] before converting them into distance constraints using the isolated spin pair approximation (ISPA).[24] For a correct conversion of measured ROE integral volumes into geometric parameters, the offset effect was taken into account.[25] The distances between the 1,3- and 1,5-diaxial protons within the GlcNAc ring (with $^4C_1$ chair geometry)[26,27] were used for calibration of the offset-corrected integrals. These calibration peaks within one moiety of the molecule produced a set of internally consistent distances when compared with standard distances from the peptide part (e.g., D-iGln^2-Hα to Hβ1/Hβ2). Hence, the effect of different internal correlation times on the distance constraints could be neglected.

2.2. NMR Assignments. In aqueous solution, PGM consists of an equilibrium mixture of the α- and β-anomeric forms (ca. 2:1) at the

(9) (a) Kessler, H. *Angew. Chem., Int. Ed. Engl.* **1982**, *21*, 512−523. (b) Constantine, K. L.; Mueller, L.; Andersen, N. H.; Tong, H.; Wandler, C. F.; Friedrichs, M. S.; Bruccoleri, R. E. *J. Am. Chem. Soc.* **1995**, *117*, 10841−10854. (c) de Alba, E.; Jiménez, M. A.; Rico, M.; Nieto, J. L. *Folding Des.* **1996**, *1*, 133−144.

(10) (a) Vijayakumar, E. K. S.; Balaram, P. *Biopolymers* **1983**, *22*, 2133−2140. (b) Bisang, C.; Weber, C.; Inglis, J.; Schiffer, C. A.; van Gunsteren, W. F.; Jelesarov, I.; Bosshard, H. R.; Robinson, J. A. *J. Am. Chem. Soc.* **1995**, *117*, 7904−7915.

(11) Sönnichsen, F. D.; Eyk, J. E.; Hodges, R. S.; Sykes, B. D. *Biochemistry* **1992**, *31*, 8790−8798.

(12) (a) Macquaire, F.; Baleux, F.; Giaccobi, E.; Huynh-Dinh, T.; Neumann, J.-M.; Sanson, A. *Biochemistry* **1992**, *31*, 2576−2582. (b) Rizo, J.; Blanco, F. J.; Kobe, B.; Bruch, M. D.; Gierasch, L. M. *Biochemistry* **1993**, *32*, 4881−4894.

(13) (a) Gerz, M.; Matter, H.; Kessler, H. *Int. J. Pept. Protein Res.* **1994**, *43*, 248−257. (b) Liang, R.; Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1995**, *117*, 10395−10396.

(14) Klaić, B.; Domenick, R. L. *Carbohydr. Res.* **1990**, *196*, 19−27.

(15) Bates, R. G. *Determination of pH. Theory and Practice*; Wiley: New York, 1964.

(16) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811−813.

(17) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *63*, 207−213.

(18) Bax, A. *J. Magn. Reson.* **1988**, *77*, 134−147.

(19) Neuhaus, D.; Keeler, J. *J. Magn. Reson.* **1986**, *68*, 568−574.

(20) Farmer, B. T., II; Macura, S.; Brown, L. R. *J. Magn. Reson.* **1987**, *72*, 347−352.

(21) Hurd, R. E.; John, B. K. *J. Magn. Reson.* **1991**, *91*, 648−653.

(22) Davis, A. L.; Keeler, J.; Laue, E. D.; Moskau, D. *J. Magn. Reson.* **1992**, *98*, 207−216.

(23) Kessler, H.; Mronga, S.; Gemmecker, G. *Magn. Reson. Chem.* **1991**, *29*, 527−557.

(24) Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis*; VCH Publishers: Weinheim, 1989.

(25) Griesinger, C.; Ernst, R. R. *J. Magn. Reson.* **1987**, *75*, 261−271.

(26) Mo, F.; Jensen, L. H. *Acta Crystallogr., Sect. B* **1978**, *34*, 1562−1569.

(27) Szilágyi, L.; Forgó, P. *Carbohydr. Res.* **1993**, *247*, 129−144.

Table 1. NMR Chemical Shift Data[a−c]

	δ_H	δ_C	δ_N		δ_H	δ_C	δ_N
GlcNAc-1	4.58	100.53		D-iGln^2-α	4.38	52.84	
GlcNAc-2	3.77	56.10		D-iGln^2-β	2.02	27.17	
GlcNAc-3	3.62	73.65		D-iGln^2-β'	2.19		
GlcNAc-4	3.49	70.36		D-iGln^2-γ,γ'	2.44	31.56	
GlcNAc-5	3.46	76.18		D-iGln^2-γCONH	8.55 (8.58)	175.94 (175.91)	119.68 (120.06)
GlcNAc-6	3.79	61.20		D-iGln^2-αCONH$_2^Z$	7.21 (7.23)	175.16	107.90
GlcNAc-6′	3.98			D-iGln^2-αCONH$_2^E$	7.68 (7.72)		
GlcNAc-CONH	8.42	174.76	122.91	*m*-Dap^3-α	4.28	53.94	
GlcNAc-Me	2.09	22.15		*m*-Dap^3-β	1.85	30.56	
MurNAc-1	5.27 (4.68)	90.31 (95.17)		*m*-Dap^3-β'	1.78		
MurNAc-2	3.82 (3.77)	53.77 (55.91)		*m*-Dap^3-γ	1.45	20.66	
MurNAc-3	3.77 (3.63)	76.39 (79.57)		*m*-Dap^3-γ'	1.52		
MurNAc-4	3.89	75.48 (75.09)		*m*-Dap^3-δ,δ'	1.94	30.37	
MurNAc-5	3.88 (3.51)	71.17 (75.33)		*m*-Dap^3-ϵ	4.06	52.84	
MurNAc-6	3.84 (3.85)	59.86 (60.08)		*m*-Dap^3-αCONH	8.30 (8.32)	173.77	125.53
MurNAc-6′	3.73 (3.92)			*m*-Dap^3-ϵCONH$_2^Z$	7.37	172.10	109.22
MurNAc-CONH	7.98 (7.83)	174.20 (174.57)	122.53 (121.83)	*m*-Dap^3-ϵCONH$_2^E$	7.96		
MurNAc-Me	1.99 (1.97)	22.18 (22.41)		D-Ala^4-α	4.38	49.66	
D-Lac-α	4.58 (4.45)	77.59 (78.14)		D-Ala^4-β	1.41	16.70	
D-Lac-β	1.42 (1.41)	18.34 (18.19)		D-Ala^4-CONH	8.36 (8.38)	173.62	125.00
D-Lac-CO		175.87 (175.30)		D-Ala^5-α	4.15	51.08	
L-Ala^1-α	4.31 (4.34)	50.09 (49.88)		D-Ala^5-β	1.37	17.53	
L-Ala^1-β	1.48	16.76 (16.83)		D-Ala^5-CONH	7.94	179.79	128.40
L-Ala^1-CONH	8.44 (8.28)	175.08 (175.11)	126.45 (126.15)				

[a] Given for the dominant α-anomeric form; those for the minor β-anomeric form are in parentheses. [b] Chemical shift referencing: DSS (internal) for ^{1}H, TMS (external) for ^{13}C NMR and indirect referencing using[80] $\gamma_N/\gamma_H = 0.101\,329\,118$ for ^{15}N NMR. [c] See Methods for sample conditions.

reducing MurNAc end of the molecule.[28] ^{1}H NMR assignments have been reported for the α-anomeric form in DMSO solution,[14] whereas tentative ^{13}C NMR assignments have been deduced, by comparison with model compounds, for the same in D_2O solution.[28] We have achieved full assignments of ^{1}H, ^{13}C, and ^{15}N-signals of the major α-anomeric form in water relying on various 2D measurements. Partial assignments could only be deduced, because of severe overlap, for the minor β-anomeric form. ^{1}H NMR assignments have been based on DQF-COSY and 2D TOCSY experiments recorded with different mixing times (15, 30, and 60 ms). The doublet signals of the anomeric protons served as starting points to derive assignments for the disaccharide moiety. Individual side chain resonances for the peptide part have been identified as characteristic crosspeak patterns in the TOCSY maps. These assignments have been cross-checked using a series of 1D TOCSY measurements.[23] A three-bond Hα/C═O correlation in the gradient HMBC spectrum[21] and an NH/NH crosspeak in the ROESY map between two Ala residues helped to establish sequential assignments for the three Ala residues. Hε of *m*-Dap exhibited a four-bond coupling to ϵCONH$_2^Z$ in the DQ-COSY map. HSQC measurements[22] have furnished assignments for protonated carbons and amide ^{15}N NMR signals, whereas gradient-enhanced ^{13}C HMBC experiments[21] provided identification for the C═O signals. The NMR chemical shift data are collected in Table 1 while Table 2 lists ^3J(NH,Hα) values and temperature dependence of amide NH chemical shifts.

Most of the ^{1}H−^{1}H NOEs were close to zero in NOESY experiments at 400 MHz and $T = 298$ K. ROESY spectra,[16,17] on the other hand, contained sufficient number of crosspeaks for obtaining internuclear distance constraints to be used in the subsequent modeling calculations. After checking for artifacts and spin diffusion (see Section 2.1), 39 out of the 45 crosspeaks in the ROESY spectrum, recorded for a solution in H_2O/D_2O (9/1), have been converted into distance constraints for the major α-anomeric form (Scheme 1). Upper and lower distance limits were set to ±5% of the calculated distances, respectively. For nondiastereotopically assigned CH_2/NH_2 protons and methyl groups, 90 and 100 pm were added to the upper bounds as pseudoatom corrections, respectively. Distance constraints are summarized and compared to the computed values in Table 3. A smaller number of crosspeaks could be identified for the minor β-anomeric form as well, but no modeling calculation was attempted on the basis of this limited data set.

Table 2. Homonuclear Coupling Constants, Temperature Coefficients of Amide Protons and Calculated ϕ Dihedral Angles

residue	^3J(NH,Hα) [Hz]	dihedral angle (ϕ)[a,b]	$\Delta\delta/\Delta T$ [−ppb/K]
GlcNAc	9.6	−145±5 −98±10	11.8
MurNAc	7.5	−159±6 −83±4	4.3
L-Ala^1	5.4	−170±3 −70±2	13.8
D-iGln^2	7.8	+158±6 +84±4	13.7
m-Dap^3	6.8	−163±6 −78±3	10.5
		+90±6 +32±5	
D-Ala^4	6.8	+163±6 +78±3	10.4
D-Ala^5	6.9	+163±6 +78±3	12.5

[a] Estimated from ^3J(NH,Hα) values using the Karplus curve given by Bystrov.[81] Very similar values were obtained using the equation given by Pardi et al.[82] except for GlcNAc where the latter returned a single ϕ value of ca. −120°. Lower positive (for L configuration) vs negative (for D configuration) torsion angles have been disregarded except for *m*-Dap^3 where calculation results are indicative of left-handed α-helical conformation (see text). [b] Definitions of the ϕ angles for the acetamido groups in GlcNAc and MurNAc: C3−C2−N2−C(═O).

2.3. Computational Procedures. All modeling work was performed using the program package SYBYL,[29] versions 6.2 or 6.22, on Silicon Graphics workstations (INDY or Indigo-2, 64 MB main memory, 128 MB swap space). Starting structures for the glycopeptide and the disaccharide were modeled interactively using SYBYL. All energy calculations for the disaccharide as well as for the glycopeptide were based on the TRIPOS 6.0 force field[30] including Gasteiger−Hückel charges,[31] except where noted. All calculations were done on molecules without formal charges.

2.3.1. Systematic Search. Rigid conformational maps were computed using SYBYL systematic search applied to the disaccharide model 3-*O*-Me-4-*O*-βGlcNAc-αMurNAc to find preferred energy conformations for the β(1-4) glycosidic linkage. Glycosidic torsion angles ϕ and ψ were incremented in 2 or 5° steps, respectively. Conformations were sampled within an energy window of 100 kcal/mol including electrostatic interactions with a distance dependent dielectricity constant resulting ca. 800 or 1800 conformers for analysis. Relaxed conformational maps were computed using the standard

(28) Klaić, B. *Carbohydr. Res.* **1982**, *110*, 320−325.

(29) *SYBYL Molecular Modeling Package*; Versions 6.2 and 6.22; Tripos: St. Louis, MO, 1995

(30) Clark, M.; Cramer, R. D.; van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982−1012.

(31) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980**, *36*, 3219−3228.

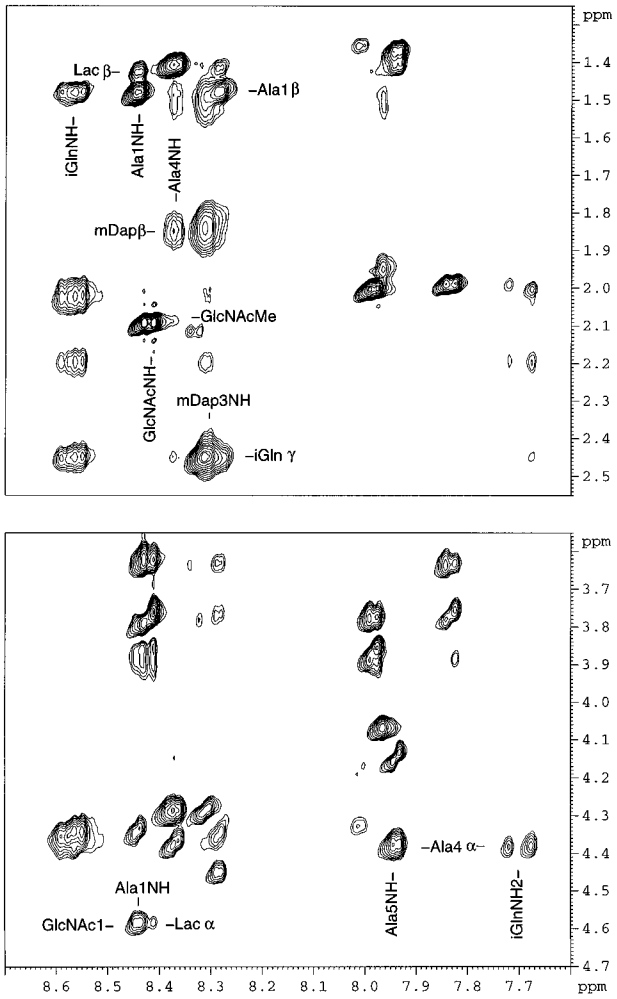


Figure 1. Partial ROESY map (τ_{mix} = 160 ms) of PGM, recorded at 500 MHz for a solution in H_2O/D_2O 9:1 (v/v), showing crosspeaks of the amide proton resonances. Selected peaks discussed in the text are marked.

TRIPOS 6.0 force field with and without the extended set of carbohydrate force-field parameters developed by S. Perez and coworkers (PIM parameters[32]) and the MM3(94) force field.[33] These maps were obtained using a systematic search procedure on the ϕ- and ψ-glycosidic angles with an increment of 10° and subsequent minimization of every conformation while holding the torsion angles of the bonds being searched fixed. Conformations in SYBYL were minimized using a quasi-Newton−Raphson procedure (BFGS), while for MM3(94) a block-diagonal Newton−Raphson method was utilized.

2.3.2. Restrained Simulated Annealing Calculations. The ROE-restrained simulated annealing (SA) calculations[34] were carried out from five different starting structures with a time step of 1.0 fs for the integration of the Newton equation of motion (using the Verlet algorithm[35,36]) for a duration of 280 ps each. The kinetic energy was included to the simulated *N,V,T* ensemble by coupling the entire system to a thermal bath.[37] The dielectricity function was constant and set to 1.

Both 6-membered carbohydrate rings were constrained in the $^4C_1$ chair geometry using torsion constraints based on X-ray data.[26,38] The ROE-derived distance constraints were applied as a biharmonic constraining function, and force constants for upper and lower boundaries were initially set to 20 kcal/(mol Å^2). The atomic velocities were applied following a Boltzmann distribution about the center of mass to obtain a starting temperature of 700 K. After simulating for 2 ps at this high temperature, the system temperature was reduced stepwise using an exponential function over a 5 ps period to reach a final temperature of 100 K. Resulting structures were sampled every 7 ps, minimized (conjugate gradient algorithm with termination criterium: gradient 0.05 kcal/(mol Å)), and stored in a separate database. This results in 5 × 40 cycles of 7 ps of a simulated annealing protocol, and the sampled 200 individual structures were analyzed subsequently. Before minimizing, the distance constraint force constants were set to a value of 3 kcal/(mol Å^2).

2.3.3. Molecular Dynamics Simulations in Water. The glycopeptide conformer with the lowest energy and constraint violation was soaked with water using the SPC model.[39] Periodic boundary conditions[40] were applied using a cube with an edge length of 34.94 Å containing ca. 1540 water molecules. Due to the application of the SHAKE[36] algorithm, the step size for the integration was set to 1 fs. Neighbor lists for the calculation of nonbonded interactions were updated every 25 fs, the actual calculation was carried out up to a radius of 10 Å without the use of switching functions. First, the solvent was allowed to relax by 500 steps and energy minimization with constrained positions of the solute atoms followed by 500 steps without any constraints. The MD simulations were performed without position or distance constraints with 4 ps for equilibration and 100 ps for analysis. Structures for analysis were sampled every 200 fs.

2.3.4. Analysis. To analyze the SA calculations, 100 conformers with the lowest averaged constraint violations and acceptable total energies were used and the following descriptors were calculated: total energy with and without electrostatic interactions, rms violations and energies of NOE-derived distance constraints, steric comparative molecular field analysis (CoMFA) fields, peptide backbone and glycosidic torsion angles, and pairwise rms deviation values of selected glycopeptide atoms. Acceptable structures were selected on the basis of maximum pairwise rms deviations.[41] For each structure, the pairwise rms deviations to all structures with lower target function values were computed and the maximum rms deviation plotted as a function of the constraint violation[41] (Figure 2, cf. section 3.2). Conformational families from simulated annealing and from MD trajectories were identified using local and global conformational descriptors: (1) local torsion angles (peptide backbone, glycosidic bond torsions), (2) pairwise global rms deviations of selected backbone atoms, and (3) global molecular shape descriptors.

Here, we propose the usage of comparative molecular field analysis[42] as a molecular shape descriptor[43] to analyze conformations. After a superposition rule based on an iterative fitting approach[44] for the different conformers was defined, the steric interaction energies between a probe atom and each structure are calculated at the surrounding points of a predefined grid. Hence, molecules or conformers are described by their surrounding Lennard−Jones-derived fields. The magnitude of the regions was defined to extend the ensemble of superimposed conformers by 4.0 Å along the principal axes of a Cartesian coordinate system. An sp^3-carbon served as probe atom, and a grid spacing of

(32) (a) Imberty, A.; Hardmann, K. D.; Carver, J. P.; Perez, S. *Glycobiology* **1991**, *1*, 456−484. (b) Imberty, A.; Perez, S. *Glycobiology* **1994**, *4*, 351−366.

(33) (a) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566. (b) Allinger, N. L.; Li, F. B.; Yan, L. Q.; Tai, J. C. *J. Comput. Chem.* **1990**, *11*, 868−895. (c) *MM3 1994 Force Field Manual*; Tripos, Inc.: St. Louis, MO, 1995; referencees cited on pp 34−38.

(34) (a) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1988**, *229*, 317−324. (b) van Gunsteren, W. F.; Berendsen, H. J. C. *Angew. Chem.* **1990**, *102*, 1020−1055.

(35) Ryckaert, J. P.; Cicotti, C.; Berendsen, H. J. C. *J. Comput. Chem.* **1977**, *23*, 327−343.

(36) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.

(37) Berendsen, H. J. C.; Dinola, A.; Haak, J. R.; Postma, J. P. M.; van Gunsteren, W. F. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(38) (a) Johnson, L. N. *Acta Crystallogr.* **1966**, *21*, 885−891. (b) Knox, J. R.; Murthy, N. S. *Acta Crystallogr., Sect. B* **1974**, *30*, 365.

(39) Berendsen, H. J. C.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces;* Pullmann, B., Ed.; Reidel Publishing Co.: Dordrecht, The Netherlands, 1981; pp 331−342.

(40) van Gunsteren, W. F.; Berendsen, H. J. C.; Hermans, J.; Hol, W. G. J.; Postma, J. P. M. *Proc. Nat. Acad. Sci. U.S.A.* **1983**, *80*, 4315−4319.

(41) Widmer, H.; Widmer, A.; Braun, W. *J. Biomol. NMR* **1993**, *3*, 307−324.

(42) Kubinyi, H., Ed. *3D-QSAR in Drug Design. Theory, Methods and Applications*; ESCOM: Leiden, The Netherlands, 1993.

(43) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959−5967.

(44) Nilges, M.; Clore, G. M.; Gronenborn, A. M. *FEBS Lett.* **1987**, *219*, 11−16.

Table 3. Comparison between Experimental and Calculated Interproton Distances (Å) from 2D ROESY Analysis (39 constraints for 11 conformers from Simulated Annealing calculations)[a]

atom 1	atom 2	r_{min}	r_{max}	r_{act}	$\langle\Delta r\rangle$	$\langle\Delta r\rangle_{rms}$	atom 1	atom 2	r_{min}	r_{max}	r_{act}	$\langle\Delta r\rangle$	$\langle\Delta r\rangle_{rms}$
L-Ala^1-NH	D-Lac-α	2.18	2.40	2.50	0.10	0.10	GlcNAc-3	GlcNAc-NH	2.26	2.50	2.43	0.00	0.00
D-Ala^4-α	D-Ala^5-NH	2.20	2.43	2.56	0.13	0.14	D-iGln^2-α	D-iGln^2-γ	2.98	4.19	3.02	0.00	0.00
D-Ala^4-α	D-iGln^2-αCONH$_2^Z$	2.61	3.79	3.52	0.00	0.00	D-iGln^2-α	D-iGln^2-β	2.47	2.73	2.55	0.01	0.03
m-Dap^3-α	D-iGln^2-γ	3.30	4.55	4.93	0.38	0.38	D-iGln^2-α	D-iGln^2-β'	2.47	2.73	3.04	0.31	0.31
MurNAc-3	D-Lac-β	3.82	5.22	4.35	0.00	0.00	*m*-Dap^3-ϵ	*m*-Dap^3-γ	3.08	4.31	3.09	0.13	0.17
D-iGln^2-γNH	L-Ala^1-β	3.91	5.32	3.96	0.09	0.12	D-iGln^2-γNH	D-iGln^2-β	2.32	2.56	2.64	0.08	0.09
m-Dap^3-αNH	D-iGln^2-γ	2.79	3.98	2.98	0.00	0.00	D-iGln^2-γNH	D-iGln^2-β'	2.70	2.98	3.33	0.41	0.42
D-Ala^4-NH	*m*-Dap^3-β	3.22	4.46	4.40	0.00	0.00	D-iGln^2-γNH	D-iGln^2-γ	2.98	4.19	4.47	0.28	0.28
L-Ala^1-NH	D-Lac-β	3.82	4.22	3.66	0.16	0.16	L-Ala^1-α	L-Ala^1-β	3.59	4.97	3.42	0.17	0.17
D-Lac-α	MurNAc-3	2.46	2.72	2.80	0.08	0.09	D-Ala^4-NH	D-Ala^4-β	3.64	5.03	3.58	0.10	0.13
GlcNAc-1	MurNAc-4	2.35	2.60	2.48	0.00	0.00	*m*-Dap^3-αNH	*m*-Dap^3-γ	2.99	4.20	2.92	0.07	0.08
L-Ala^1-α	MurNAc-2	2.71	3.00	3.19	0.19	0.19	*m*-Dap^3-αNH	*m*-Dap^3-β	2.78	3.97	3.76	0.00	0.00
D-Ala^4-NH	D-Ala^5-NH	2.33	2.58	2.66	0.08	0.10	D-Ala^5-NH	D-Ala^5-β	3.65	5.04	3.48	0.17	0.17
D-iGln^2-α	D-iGln^2-γNH	2.21	2.45	2.37	0.02	0.08	GlcNAc-5	GlcNAc-1	2.37	2.62	2.51	0.00	0.00
GlcNAc-1	GlcNAc-NH	2.25	2.49	2.78	0.29	0.29	GlcNAc-3	GlcNAc-1	2.47	2.73	2.66	0.00	0.00
m-Dap^3-αNH	*m*-Dap^3-α	2.18	2.41	2.30	0.00	0.00	D-iGln^2-γNH	D-iGln^2-αCONH$_2$	2.65	3.83	3.83	0.02	0.05
D-Ala^5-NH	D-Ala^5-α	2.51	2.77	2.33	0.18	0.18	D-iGln^2-β'	D-iGln^2-αCONH$_2$	2.83	4.03	4.32	0.37	0.39
m-Dap^3-ϵ	*m*-Dap^3-ϵCONH$_2$	2.95	4.16	3.53	0.00	0.00	*m*-Dap^3-δ	*m*-Dap^3-ϵCONH$_2$	3.59	4.87	3.97	0.00	0.00
MurNAc-3	MurNAc-NH	2.32	2.57	2.48	0.00	0.00	MurNAc-2	MurNAc-1	2.87	3.18	2.50	0.37	0.37
GlcNAc-2	GlcNAc-NH	2.18	2.41	2.95	0.54	0.54							

average restraint violation: 0.121 Å
average rms restraint violation: 0.130 Å

[a] Abbreviations: r_{min}, lower distance limit; r_{max}, upper distance limit; r_{act}, actual distance; $\langle\Delta r\rangle$, average violation; $\langle\Delta r\rangle_{rms}$, rms violation.

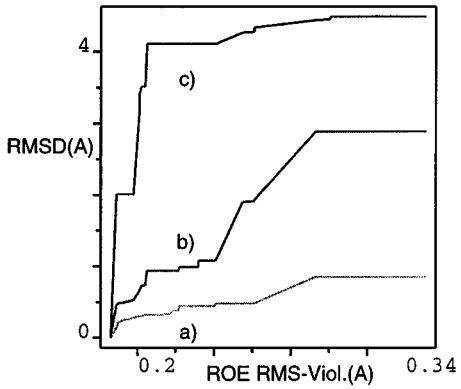


Figure 2. Maximum of the rms deviation between pairs of PGM conformers from simulated annealing calculations with a rms NOE violation smaller than a cut-off value. The sample of structures for rms deviation calculations gets progressively larger with increasing cut-off values for the NOE violations given on the *x*-axis. On the *y*-axis, the maximum rms deviation values are given, obtained as described in the text. The rms deviation values were obtained using three different set of atoms: (a) bottom, using all MurNAc heavy atoms atoms plus peptide backbone atoms of residues D-Lac, Ala^1, and D-iGln^2; (b) middle, using all heavy atoms for the disaccharide plus peptide residues D-Lac, Ala^1, and D-iGln^2; (c) top, using all heavy atoms of PGM.

2.0 Å was used. The resulting grid point matrices ("molecular steric fields") were analyzed using hierarchical cluster analysis.[45] A pairwise matrix of Euclidean distances between "molecular fields" is used for clustering. "Natural" cluster levels were identified by finding the largest distance between various cluster levels.

The usage of any local descriptor, such as torsion angles, normally does not reflect the global fold of a conformer properly, while global descriptors like global rms deviation or molecular steric fields can tackle this problem. Small changes of local descriptors can cause large positional differences in different regions of a molecule, thus molecules within a single cluster still tend to have relatively large pairwise backbone rms deviation violations. Thus, monitoring of global conformational changes is one advantage of CoMFA fields. Moreover, other conformational properties can be included into the molecular shape analysis to classify conformers (e.g., electrostatic, hydrophobic[46] hydrogen bond, or indicator fields). Another advantage of molecular fields is that orthogonal principal properties can be extracted and visualized to monitor conformational changes in a two- or three-dimensional graph.

Thus, principal component analyses[47] (PCA) were performed on molecular fields for the MD trajectory to contract the large number of highly correlated steric field points as variables to a few descriptive, orthogonal dimensions, so-called "principal conformational properties". The scores can be plotted against each other to visualize the changes in principal global conformational properties during the time course of the MD simulation. The first new coordinate describes the maximum variance among all possible directions, the second one the next largest variation among all directions orthogonal to the first one, etc. In PCA, the data matrix **X** (steric field energies x_{ik} for *i* conformers and *k* grid points) is decomposed to means (x_k), scores (t_{ia}), loadings (p_{ak}) and residuals (e_{ik}), with *a* denoting the number of significant model dimensions, as defined by equation 1:

$$x_{ik} = x_k + \sum\nolimits_a t_{ia} \cdot p_{ak} + e_{ik} \qquad (1)$$

The numerical value of *a* (i.e., the number of appropriate principal properties) is determined by cross-validation.[48]

Dihedral angle autocorrelation and radial distribution[49] functions were computed using standard procedures. Cross-correlation coefficients for normalized dihedral angle fluctuations were computed and displayed in a 2D dynamic cross-correlation map (DCCM[50]). Off-diagonal elements reveal correlated fluctuations between different parts of the molecule (cf. Figure10).

3. Results and Discussion

3.1. NMR Data. The majority of the observed ROEs arise from intraresidue dipolar contacts (Table 3). The relatively

(45) *Sybyl 6.2. Ligand-Based Design Manual*; Tripos: St. Louis, MO, 1995; pp 246−255 and references cited therein.

(46) Kellogg, G. E.; Semus, S. F.; Abraham, D. J. *J. Comput.-Aided Mol. Des.* **1991**, *5*, 545−552.

(47) (a) Malinowski, E. R.; Howery, D. G. *Factor Analysis in Chemistry*; Wiley: New York, 1980. (b) Cramer, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 1837−1849. (c) Stahle, L.; Wold, S. Multivariate Data Analysis and Experimental Design in Biomedical Research. In *Progress in Medicinal Chemistry*; Ellis, G. P., West, G. B., Eds.; Elsevier: Amsterdam, The Netherlands, 1988; pp 292−338. (d) Rännar, S.; Lindgren, F.; Geladi, P.; Wold, S. *Chemometrics* **1994**, *8*, 111−125.

(48) Wold, S. *Technometrics* **1978**, *20*, 379.

(49) (a) Rossky, P. J.; Karplus, M. *J. Am. Chem. Soc.* **1979**, *101*, 1913−1937. (b) Jorgensen, W. L.; Bigot, B.; Chandrasekar, J. *J. Am. Chem. Soc.* **1982**, *104*, 4584−4591.

(50) Swaminathan, S.; Harte, W. E.; Beveridge, D. L. *J. Am. Chem. Soc.* **1991**, *113*, 2717−2721.

small number of interactions between protons on different residues in the peptide chain is, in itself, an indication for the low probability of ordered structures to occur under the experimental conditions used. Moreover, the majority of the *inter*residue ROE contacts occur between neighboring residues as evidenced by $\alpha N(i, i + 1)$ and $\beta N(i, i + 1)$ sequential[51] crosspeaks (Table 3). In fact, all possible sequential contacts, either of αN- or βN-type, could be observed along the hexapeptide backbone starting with D-Lac; this latter being considered as a ψ-Ala N-terminal residue. The crosspeak between D-iGln-Hγ and *m*-Dap-(NH)α can be considered as an "αN"-type contact because (CO)γ, rather than the usual (CO)α, is involved into the backbone for D-iGln. Sequential crosspeaks like those mentioned are usually observed for small unstructured peptides in aqueous solution.[52] $NN(i, i + 1)$ ROEs such as the one observed here between Ala^4 and Ala^5 (Table 3) also frequently occur under such conditions for C-terminal residues with an unprotected COOH group.[52,53] On the other hand, three of the ROEs between the carbohydrate and peptide moieties (MurNAc-3/D-Lac-Hβ, MurNAc-3/D-Lac-Hα, and MurNAc-2/Ala^1-Hα, Table 3) are indicative of less conformational freedom at the N-terminal part of the hexapeptide chain. In addition, a strong interglycosidic ROE was detected between GlcNAc-1 and MurNAc-4 (Table 3).

Two of the remaining interresidue contacts along the peptide chain may indicate departure from the random coil state toward some preferred conformation; a long-range ROE appears between D-iGln-αNH_2 and $D\text{-}Ala^4$-Hα, and another one is seen between D-iGln-Hγ and m-Dap-Hα. The latter one is reminiscent of the $\alpha\alpha(i, i + 1)$ type of interaction which can be observed in regular sheet structures.[51] Amide ^{1}H chemical shifts display large temperature coefficients (Table 2), indicating practically total exposure to solvent. The single exception, showing a markedly decreased $\Delta\delta/\delta T$ value, is that of the MurNAc-NH; this may be indicative of the formation of a weak to moderate hydrogen bond.[54] Modeling calculations (vide infra) suggest one of the solvent H_2O molecules, rather than any of the acceptor atoms of the peptide chain, to be the most likely acceptor to this hydrogen.

3.2. Structure Selection. Acceptable conformers for analysis were selected from the set of 100 SA structures with the lowest averaged constraint violations and total energies. The dependence of the maximum rms deviation between pairs of 100 PGM conformers on the rms NOE violation cutoff is plotted in Figure 2. The rms deviation values were calculated using three sets of atoms: (a) all MurNAc heavy atoms plus backbone atoms from D-Lac, Ala^1, and $D\text{-}iGln^2$, (b) all heavy atoms for the disaccharide plus peptide residues D-Lac, Ala^1, and $D\text{-}iGln^2$, and (c) all heavy atoms. For a single family of conformations, this maximum rms deviation should rise continuously as a function of the ROE violations.[41] However, Figure 2 reveals some significant steps indicating that a new conformational state becomes accessible, which differs significantly from all structures with lower target function values.

An unbiased selection of NMR-derived structures now is based on the grouping of conformers with low-constraint violations and a relatively stable conformational space accessible. Therefore, taking into account only the first well-defined plateau in plots 2b and 2c, with a constraint violation between 0.18 and 0.20 Å, 11 structures were selected (Figure 3) which form a well-defined family with a rigid carbohydrate- and N-terminal peptide moiety (upper part of the plot) and alternate conformations in the C-terminal peptide part, in agreement with the experimental data. Allowing for a higher constraint violation (between 0.21 and 0.24 Å), a second plateau with 43 conformers can be identified. Again, the well-defined part of the molecule carrying the carbohydrate moiety is in contrast to the flexible C-terminal region (see Supporting Information).

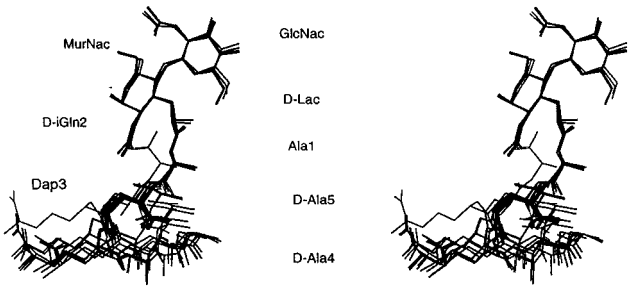


Figure 3. Stereoview of 11 best conformers obtained from simulated annealing calculations of PGM corresponding to the first plateau plotted in Figure 2. In general, protons are not displayed. All structures were superimposed using the fixed region defined by using all heavy atoms for the disaccharide plus peptide residues D-Lac, Ala^1, and $D\text{-}iGln^2$.

More backbone conformations can be identified which differ with respect to the "hinge" region around $D\text{-}iGln^2$ and, consequently, the orientation of the C-terminal residues. These structures can be grouped into three different clusters (cluster IDs: 2, 3, and 4). In total, five separate clusters with 15, 16, 14, 40, and 15 conformers were identified but only three (2, 3, and 4) coincide with the 43 conformers from the first two plateaus, while the first plateau only coincides with structures from cluster 4 (40 members). Two subgroups within cluster 4 can be identified at a lower similarity level and only one of these subclusters is represented at the first plateau. On the other hand, clusters 1 and 5 and the second subgroup of 4 are only populated with compounds having much higher tolerated cutoff values (see Supporting Information).

3.3. Peptide Backbone Conformation. The distance constraints and dihedral analysis for the acceptable conformers from simulated annealing are summarized in Tables 3 and 4. In agreement with the measured temperature coefficients (Table 2), no hydrogen bonds were found. The N-terminal peptide residues D-Lac and $L\text{-}Ala^1$ connected to the carbohydrate are well-defined in an extended conformation in accord with experimental data. The backbone dihedrals for Ala^1 are in the characteristic range for β-sheet conformations ($-163/127°$) with larger fluctuations in SA and MD observed for the ψ-angle. The aliphatic bridge between $D\text{-}iGln^2$ and Dap^3 is found to be very flexible in MD simulations, and the SA ensemble also reveals a high degree of disorder or flexibility in this region (cf. Table 4). The orientation of the terminal Cα-$CONH_2$ is known due to the long-range ROE between $CONH_2$ and $D\text{-}Ala^4$-Hα, counting for a backfolding of the C-terminal region.

For Dap^3, the backbone dihedrals adopt a left-handed α-helical conformation (61/64°) corresponding to one of the sterically allowed regions for L-α-substituted amino acids.[55] For $D\text{-}Ala^4$, the backbone angles are in the characteristic range for an extended conformation; larger fluctuations are observed in MD simulation for both dihedrals of the terminal $D\text{-}Ala^5$ (Table 4). It is seen from Table 2 that for each residue (except Ala^5) one of the allowed ϕ values estimated from the ^3J(NH,Hα) coupling constants are in agreement with the results from SA and MD calculations (Table 4). No evidence was, however, apparent

(51) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons: New York, 1986.

(52) Dyson, H. J.; Merutka, G.; Waltho, J. P.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1992**, *226*, 795−817.

(53) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5*, 14−24.

(54) Higashijama, T.; Kobayashi, Y.; Nagai, U.; Miyazawa, T. *Eur. J. Biochem.* **1979**, *97*, 43.

(55) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 174−175.

Table 4. Comparison of Selected Peptide Backbone, Side Chain, and Saccharide Dihedral Angles[a]

(a) Eleven Conformers from Simulated Annealing[b]

torsion	mean	std	high	low	torsion	mean	std	high	low
D-Ala^5-ψ	−62.97	36.21	−49.90	−172.10	D-iGln^2-ϕ	90.93	21.71	155.90	81.10
D-Ala^5-ϕ	−54.25	3.40	−51.60	−64.00	L-Ala^1-ψ	126.88	20.07	142.90	86.10
D-Ala^4-ψ	−61.95	6.39	−57.00	−74.40	L-Ala^1-ϕ	−163.71	2.88	−159.80	−168.30
D-Ala^4-ϕ	203.59	52.35	290.40	165.70	D-Lac-ψ	−150.45	1.58	−147.70	−152.50
m-Dap^3-ψ	64.41	18.16	86.40	34.50	D-Lac-ϕ	81.82	1.16	83.00	80.00
m-Dap^3-ϕ	61.21	1.99	63.50	56.90	MurNAc-ϕ'	−158.85	4.34	−151.10	−164.20
m-Dap^3-χ^1	−54.21	2.17	−52.10	−59.90	MurNAc-ψ'	65.8	0.2	66.1	65.3
m-Dap^3-χ^2	181.22	1.31	182.90	179.20	GlcNAc-ψ	117.68	6.61	134.00	114.20
m-Dap^3-χ^3	−136.83	53.07	−69.00	−176.40	GlcNAc-ϕ'	174.11	2.86	176.10	166.70
m-Dap^3-χ^4	−175.22	2.54	−172.40	−178.60	MurNAc-C56	58.50	1.60	59.80	54.50
D-iGln^2-ψ^1	74.70	8.95	88.70	64.80	GlcNAc-C56	48.60	33.50	59.40	−52.20
D-iGln^2-ψ^2	73.57	10.65	86.70	58.30	MurNAc-C3−C2−N2−C(=O)	−153.3	0.2	−153.0	−153.5
D-iGln^2-ψ^3	165.16	10.89	192.90	154.80	GlcNAc-C3−C2−N2−C(=O)	−154.0	2.8	−150.1	−157.6

(b) MD Simulation in Water [c]

torsion	mean	std	plateau	τ_2	τ_4	ID	torsion	mean	std	plateau	τ_2	τ_4	ID
D-Ala^5-ψ	91.66	67.08	0.68	5.10	3.52	1	D-iGln^2-ψ^3	177.76	8.05	0.15	16.95	18.15	13
D-Ala^5-ϕ	24.50	70.25	0.61	2.94	1.77	2	D-iGln^2-ϕ	123.85	52.41	0.71	3.82	2.89	14
D-Ala^4-ψ	−125.63	43.74	0.68	5.24	3.44	3	L-Ala^1-ψ	130.54	48.33	0.69	3.31	2.83	15
D-Ala^4-ϕ	150.11	19.02	0.28	7.91	10.66	4	L-Ala^1-ϕ	−160.46	9.80	0.08	12.20	8.67	16
m-Dap^3-ψ	67.28	14.75	0.29	8.77	13.48	5	D-Lac-ψ	−166.18	8.74	0.30	17.96	6.08	17
m-Dap^3-ϕ	64.18	10.80	0.12	4.87	11.22	6	D-Lac-ϕ	84.60	10.68	0.20	19.19	11.33	18
m-Dap^3-χ^1	−58.44	9.96	0.10	13.37	12.70	7	MurNAc-ϕ'	−161.56	15.37	0.42	10.13	6.55	19
m-Dap^3-χ^2	183.22	8.72	0.09	15.88	12.67	8	GlcNAc-ψ	118.04	11.48	0.28	16.92	14.29	20
m-Dap^3-χ^3	174.77	10.42	0.20	12.03	15.65	9	GlcNAc-ϕ'	164.87	7.36	0.14	16.40	25.57	21
m-Dap^3-χ^4	175.17	9.89	0.11	18.67	11.33	10	MurNAc-C56	60.02	10.80	0.13	19.77	15.77	22
D-iGln^2-ψ^1	76.99	12.23	0.32	16.58	13.14	11	GlcNAc-C56	−60.88	9.90	0.06	26.41	20.24	23
D-iGln^2-ψ^2	80.89	10.45	0.13	19.98	12.22	12							

[a] Definitions for saccharide torsion angles: MurNAc-ϕ', C4−C3−O3−Lac-α; GlcNAc-ϕ', C2−C1−O1−C4(MurNAc); GlcNAc-ψ, C1−O1−C4(MurNAc)−C3(MurNAc); MurNAc-ψ', N2−C2−C3−O3. [b] Definitions: mean, averaged torsion [deg] for 11 conformers obtained from SA calculations; std, standard deviation [deg]; high, highest torsion [deg]; low, lowest torsion [deg]. [c] Definitions: mean, averaged torsion [deg] from MD in water; std, standard deviation [deg]; plateau, plateau value for the dihedral autocorrelation function in the range between 3 and 10 ps; τ_2, dihedral fluctuation relaxation times [ps] within a period of 2 ps; τ_4, dihedral fluctuation relaxation times [ps] within a period of 4 ps; ID, dihedral angle ID from the DCCM graph (Figure 10).

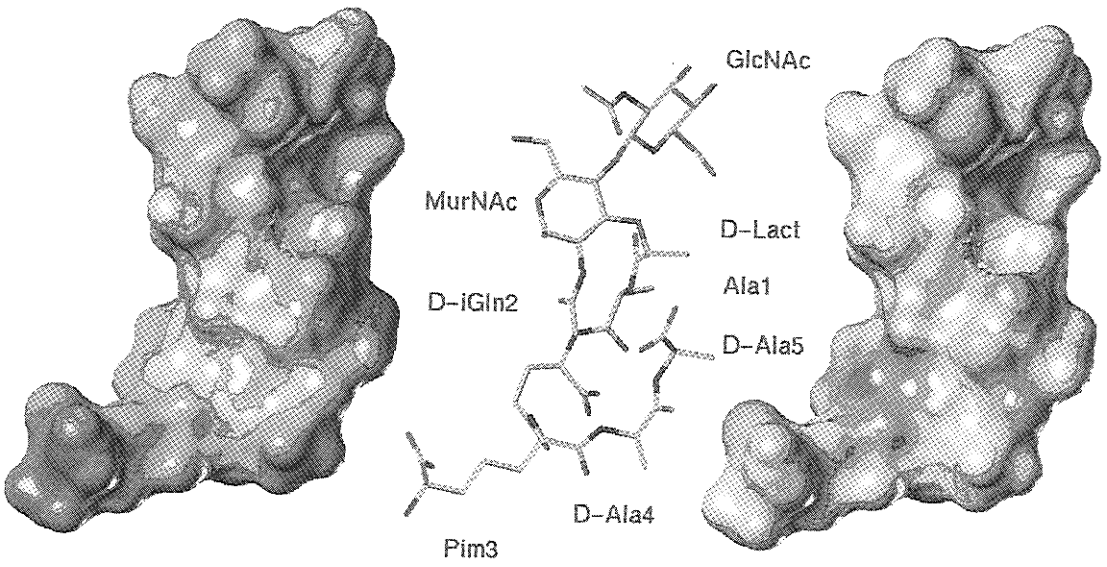


Figure 4. Representative low-energy conformer of PGM from simulated annealing calculations with lipophilicity potential (left) and electrostatic potential (right) mapped onto a Connolly molecular surface. The gray scale for the lipophilic potential ranges from dark (highest lipophilic area) to light (highest hydrophilic area), while for the electric potential light-medium gray regions indicate positive and dark regions indicate negative areas.

to support formation of β-turnlike conformations in the region from MurNAc acetamido group through D-iGln^2 as proposed earlier[5,8] for MDP, a smaller analog of PGM, in DMSO solution.

A representative low-energy conformer of PGM from SA calculations is displayed in Figure 4 to illustrate its molecular properties. The lipophilicity potential (LP) (Figure 4, left)[56] and electrostatic potential (EP) (Figure 4, right) of PGM were mapped onto a Connolly molecular surface[57] using the program MOLCAD.[29,58] The gray scale for the lipophilic potential ranges from dark (highest lipophilic area) to light (highest hydrophilic

(56) (a) Audry, E.; Colleter, J. C.; Dallet, P.; Dubost, J. P. *Eur. J. Med. Chem.* **1986**, *21*, 71−72. (b) Viswanadhan, V. N.; Ghose, A. K.; Revankar, G. R.; Robins, R. K. *J. Chem. Inf. Comput. Sci.* **1989**, *29*, 163−172.
(57) Connolly, M. L. *Science* **1983**, *221*, 709−713

area), while for the electric potential light-medium gray regions indicate positive areas and dark regions indicate negative areas. A lipophilic region at one side including the side chains of D-Lac, Ala[1], D-Ala[4], and D-Ala[5] (dark, backside) is in contrast to the hydrophilic side chain of Dap[3] (medium gray) the disaccharide hydroxyl groups (medium gray at the top) and the backbone (frontside) of the residues mentioned above, thus forming a molecule with two different lipophilicity sites, while the electrostatic potential extrema are scattered over the entire surface.

3.4. Disaccharide Conformation. The glycosidic dihedral angles are defined, according to Sundaralingam[59] and Bush et al.,[60] as follows: ϕ, O5−C1−O1−C4′; ψ, C1−O1−C4′−C3′. In addition, we report the glycosidic dihedral angle ϕ' C2−C1−O1−C4′ for comparison (e.g. Figure 4). The *Z-anti* orientation[61] of the MurNAc N2-acetamido group with respect to MurNAc-C3 is in agreement with the strong ROE observed between NH and H3 (Table 3). The corresponding C3−C2−N2−C′ dihedral angle is −126° for crystalline β-D-GlcNAc-N-Asn,[62] while −153° was observed here (Tables 2 and 4). In contrast, conformational averaging is observed for the N-acetyl group of GlcNAc, in agreement with literature observations.[61,63] A network of three distances from GlcNAc-NH to GlcNAc-H1/H2 and H3 was measured (2.37, 2.29, and 2.38 Å, respectively), while only the ROE involving GlcNAc-H3 is fulfilled in the SA ensemble (2.43 Å), where for C3−C2−N2−C′ a dihedral angle of −154° is found (Table 4). This set of conflicting ROEs counts for multiple rotameric states of the N2-acetamido group. Conformational equilibria involving fast-interconverting species[64] can be handled using numerous approaches.[65] A better description of NMR observables can be achieved by averaging of conformational ensembles.[66] After a grid search procedure around the C3−C2−N2−C′ torsion, three low-energy rotamers were identified which are separated by barriers up to ca. 7 kcal/mol: c1 60° (8.9 kcal/mol), c2 −160° (7.7 kcal/mol), and c3 −90° (8.9 kcal/mol). These conformers are in agreement with the value (9.6 Hz) measured for ^3J(Hα,-NH) while the ROE intensities could not be quantitatively explained assuming a single conformer. This is in accord with recent findings on the contributions of *Z-anti* as well as *Z-syn* conformers to the rotameric distribution around the C2−N2 bond of GlcNAc and derivatives.[61]

The PLS method,[67] using three experimental ROE intensities as dependent variables and the back-calculated normalized intensities as x-block, was used to estimate the rotameric populations around the C2−N2 bond. The distributions calculated for c1/c2/c3 (46/27/27% for r^{-3} dependence or 59/21/20% for r^{-6} dependence) are in support of conformational flexibility around the C2−N2 bond with a dominantly populated c1 (60°) rotamer.

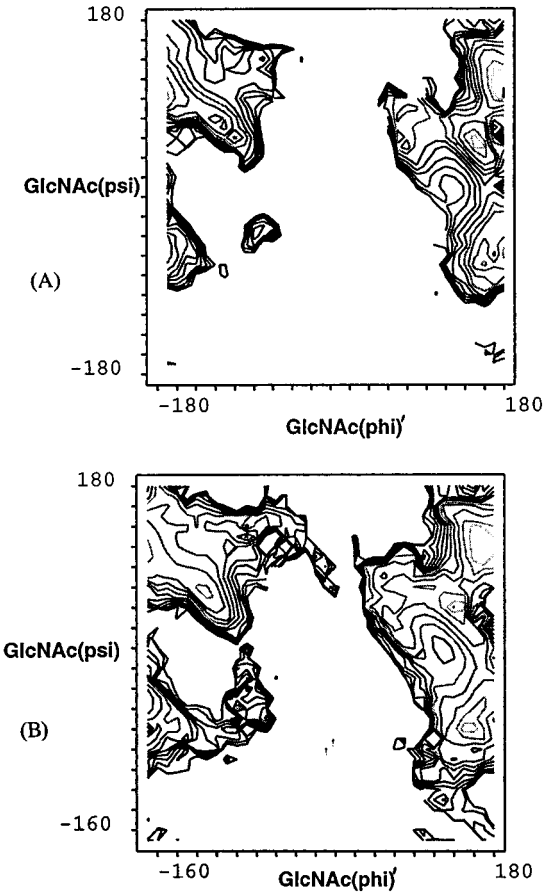


Figure 5. Relaxed conformational maps obtained using grid search procedures with subsequent energy minimization for the model disaccharide 3-*O*-Me-4-*O*-βGlcNAc-αMurNAc using (a) the standard TRIPOS 6.0 force field and (b) the extended set of carbohydrate force field parameters developed by Perez et al. (PIM parameters, see text). The dihedral angle ϕ' (C2−C1−O1−C4′) on the *x*-axis is plotted against ψ (C1−O1−C4′−C3′), and contour levels are drawn each 1 kcal/mol.

The glycopeptide linkage between MurNAc and D-Lac adopts a well-defined, rigid conformation (MurNAc ϕ'/ψ'-dihedral angles of −158/66°, Table 4) as confirmed by three characteristic peptide−carbohydrate ROEs (MurNAc-H3 to D-Lac-Hα/Hβ and MurNAc-H2 to Ala[1]-Hα).

The conformation of the β(1-4) glycosidic bond is well-defined and in agreement with experimental data, and there are no experimental indications for a significantly populated second conformation. The strong ROE effect between GlcNAc-H1 and MurNAc-H4 is compatible with ϕ'/ψ = 174/118° for GlcNAc. The relaxed conformational maps for the β(1-4) glycosidic bond of the model disaccharide 3-Me-4-βGlcNAc-α-MurNAc are plotted in Figure 5. The map obtained using the standard TRIPOS 6.0 force field is shown in Figure 5a. Three different energy minima at ϕ'/ψ-dihedral angles of 180/110° (7.99 kcal/mol), 150/60° (8.31 kcal/mol), and 180/−70° (8.73 kcal/mol) can be identified, in agreement with the stereoelectronic "exoanomeric" effect.[68] A qualitatively similar relaxed conformational map is obtained using the extended set of PIM

(58) Heiden, W.; Moeckel, G.; Brickmann, J. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 503.
(59) Sundaralingam, M. *Biopolymers* **1968**, *6*, 189−213.
(60) Bush, C. A.; Rao, B. N. N.; Yan, Z. Y. *J. Am. Chem. Soc.* **1986**, *108*, 6168−6173.
(61) Fowler, P.; Bernet, B.; Vasella, A. *Helv. Chim. Acta* **1996**, *79*, 269−287.
(62) Delbaere, L. T. *Biochem. J.* **1974**, *143*, 197−205.
(63) Kessler, H.; Matter, H.; Gemmecker, G.; Kling, A.; Kottenhahn, M. *J. Am. Chem. Soc.* **1991**, *113*, 7550−7563.
(64) (a) Rowan, R.; Warshel, A.; Sykes, B. D.; Karplus, M. J. *Biochemistry* **1974**, *13*, 970−981. (b) Scarsdale, J. N.; Yu, R. K.; Prestegard, J. H. *J. Am. Chem. Soc.* **1986**, *108*, 6778−6784. (c) Kessler, H.; Griesinger, C.; Lautz, J.; Muller, A.; van Gunsteren, W. F.; Berendsen, H. J. C. *J. Am. Chem. Soc.* **1988**, *110*, 3393−3396. (d) Kim, Y.; Ohlrogge, J. B.; Prestegard, J. H. *Biochem. Pharmacol.* **1991**, *40*, 7−13. (e) Kessler, H.; Geyer, A.; Matter, H.; Köck, M. *Int. J. Pept. Protein Res.* **1992**, *40*, 25−40. (f) Matter, H.; Kessler, H. *J. Am. Chem. Soc.* **1995**, *117*, 3347−3359.
(65) (a) Torda, A. E.; Scheek, R. M.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1989**, *157*, 289−294. (b) Brüschweiler, R.; Blackledge, M.; Ernst, R. R. *J. Biomol. NMR* **1991**, *1*, 3−11. (c) Landis, C.; Allured, V. S. *J. Am. Chem. Soc.* **1991**, *113*, 9493−9499. (d) Bonvin, A. M. J. J.; Rullmann, J. A. C.; Lamerichs, R. M. J. N.; Boelens, R.; Kaptein, R. *Proteins: Struct., Funct., and Genet.* **1993**, *15*, 385−400. (e) Mierke, D. F.; Kurz, M.; Kessler, H. *J. Am. Chem. Soc.* **1994**, *116*, 1042−1049. (f) Cicero, D. O.; Barbato, G.; Bazzo, R. *J. Am. Chem. Soc.* **1995**, *117*, 1027−1033.
(66) Jardetzky, O. *Biochim. Biophys. Acta* **1980**, *621*, 227−232.
(67) Wold, S.; Dunn, W. J.; Ruhe, A.; Wold, H. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735−743.

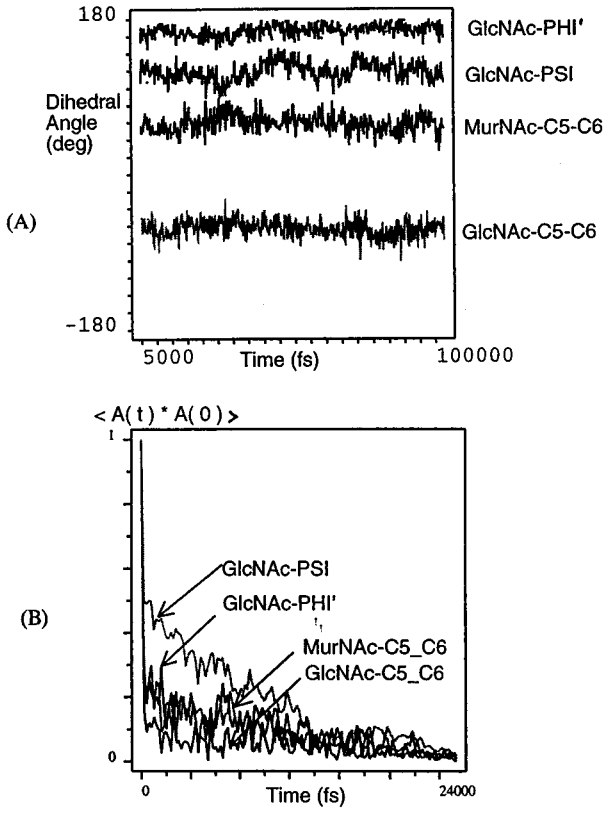


Figure 6. (a) Time development of characteristic exocyclic disaccharide dihedral angles during a 104 ps MD trajectory of PGM in aqueous solution. The dihedral angles ϕ' (C2−C1−O1−C4′), ψ (C1−O1−C4′−C3′), and ω (O5−C5−C6−O6) for GlnNAc and ω for MurNAc are plotted against the simulation time [fs]. (b) Autocorrelation functions for the same disaccharide dihedral angles.

carbohydrate parameters[32] (cf. section 2.3.1 and Figure 5b), with the following locations of the energy minima: 170/110° (−5.94 kcal/mol), 140/60° (−6.03 kcal/mol), and 160/−60° (−5.79 kcal/mol). The relaxed map computed using the MM3(94) force field shows only a single-broad minimum (150/80°, Supporting Information).

Only those energy minima corresponding to ψ-dihedrals between 60 and 120° are in accord with the strong ROE effect GlcNAc-H1 to MurNAc-H4 across the glycosidic linkage. The barriers between ψ-dihedral angle minima are low; hence, these conformers can easily exist in fast equilibrium. However, only the minimum at $\psi = 110°$ is experimentally identified and no conformational interconversion of this glycosidic bond was observed during the MD trajectory in water (Figure 6a). This energy minimum also is in agreement with literature ϕ'/ψ-values for α-lactose (146/95°), [69] β-cellobiose (167/106°),[70] or β-2-deoxylactose (163/110°).[71]

For each 6-membered ring, the standard $^4C_1$ conformation is identified and maintained during the time course of all MD simulations in water. The rms deviations of these internal ring dihedrals are nearly uniform with typical values between 5 and 10°, in agreement with published data.[72] Due to missing

(68) (a) Lemieux, R. U. *Pure Appl. Chem.* **1971**, *25*, 527−548. (b) Kirby, A. J. *The Anomeric Effect and Related Stereoelectronic Effects at Oxygen*; Springer: Berlin, 1983.

(69) Fries, D. C.; Rao, S. T.; Sundaralingam, M. *Acta Crystallogr., Sect. B* **1971**, *27*, 994−1005.

(70) (a) Jacobsen, R. A.; Wunderlich, J. A.; Lipscomb, W. N. *Acta Crystallogr.* **1961**, *14*, 598−607. (b) Chou, S.; Jeffrey, G. A. *Acta Crystallogr., Sect. B* **1968**, *24*, 830−838.

(71) Kessler, H.; Matter, H.; Gemmecker, G.; Kottenhahn, M.; Bats, J. W. *J. Am. Chem. Soc.* **1992**, *114*, 4805−4818.

(72) Brady, J. W. *J. Am. Chem. Soc.* **1989**, *111*, 5155−5165.

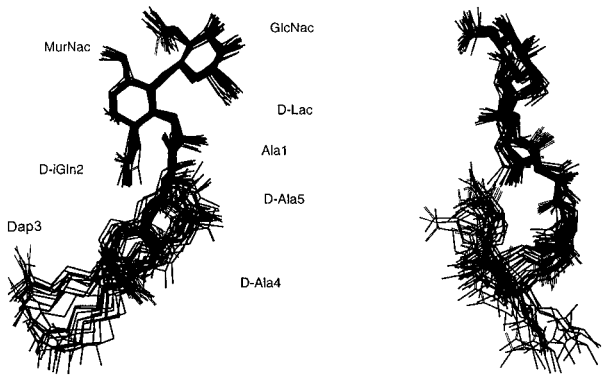


Figure 7. Orthogonal view of 26 PGM conformers sampled as snapshots of the entire 104 ps MD trajectory of PGM in water. The "rigid" and "flexible" part of the molecule are visible, and both are connected via the aliphatic "hinge" side chain of D-iGln[2].

experimental data, it is not possible to analyze conformational preferences for the exocyclic C5−C6 bond. These torsion angles were found to adopt only a single conformation during the MD simulation in aqueous solution (Figure 6a).

3.5. Molecular Dynamics Studies. To gain additional insight into the dynamic behavior of this glycopeptide on the picosecond time scale, MD simulations without experimental constraints were performed. The superposition of an ensemble of 26 snapshot structures for PGM shows the rigidity of the N-terminal peptide moiety and the connected carbohydrate, while the C-terminal peptide part undergoes conformational variations around the "hinge" amino acid D-iGln[2] (Figure 7). In Table 4b the averaged dihedral angles and their fluctuations are reported.

To analyze the trajectory in detail, a hierarchical cluster analysis based on molecular shape descriptors was performed using the "rigid" carbohydrate and N-terminal part of PGM as an alignment rule. Seven conformational families were identified on the basis of the molecular steric fields with the following populations: 1/85, 2/53, 3/59, 4/3, 5/20, 6/28, and 7/8. Clusters 1 and 2 can be grouped into a larger family, which is also true for clusters 3, 4, 5, and 6, 7, respectively. Representative conformers for the dominant clusters are shown in the Supporting Information.

The application of a PCA to the molecular steric fields[73] contracts the large number of correlated variables (steric field energies at discrete gridpoints) to a few, descriptive and orthogonal *principal conformational properties* (PCPs, cf. section 2.3.4). The first two orthogonal PCPs in the model (i.e., PCA scores) could explain ca. 80% of the variance in the data matrix, while the remaining components were less important. These PCPs show an interesting pattern, when time evolves along the MD trajectory (Figure 8). Starting in the upper left quadrant of the scores plot at ca. −1/1, corresponding to cluster 1, a second cluster (cluster ID 2) is formed at −0.5/0. This subgroup still is similar to the first one, and clusters 1 and 2 can therefore be grouped together to form a first cluster family. Then a second family, consisting of three clusters (cluster IDs 3, 4, 5) and formed after a conformational transition, is present in the lower half of the PCP plot (−1/−1). This second family differs in both PCPs from the first one, while the next, third cluster family (−1/2) shows conformational differences only in the second PCP. It consists of the remaining clusters with the IDs 6 and 7.

(73) van de Waterbeemd, H.; Clementi, S.; Costantino, G.; Carrupt, P.-A.; Testa, B. In *3D-QSAR in Drug Design. Theory, Methods and Applications*; Kubinyi, H., Ed.; ESCOM: Leiden, The Netherlands, 1993; pp 697−707.

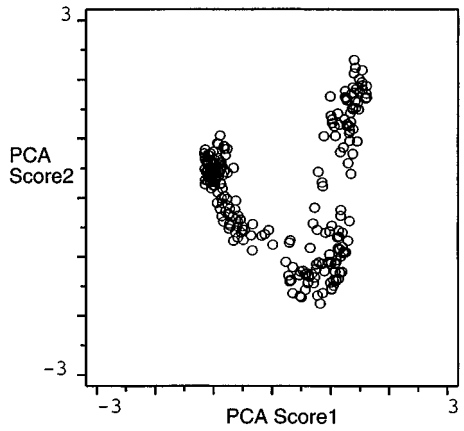


Figure 8. Plot of first vs second score from principal component analysis (PCA) of the molecular shape descriptor. Each score refers to a major principal conformational property (PCP).

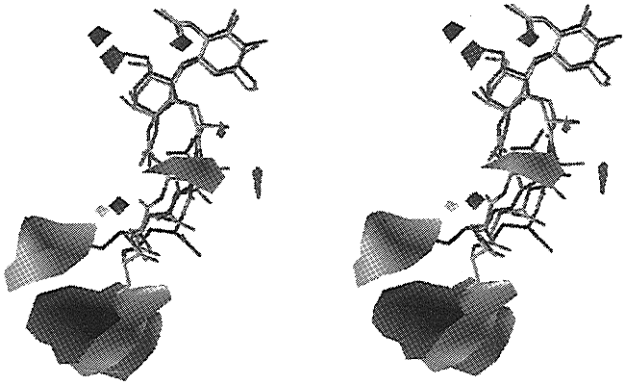


Figure 9. Stereoview of two representative conformers from MD simulations (cluster 1, white carbons; cluster 6, light gray carbons) together with contributions of the PCA loadings for the major principal conformational properties PCP1 and PCP2 obtained from principal component analysis based on CoMFA steric molecular fields. Of the four patches at the bottom, the first and third ones (counted downward, middle gray shading) correspond to 15 and 85%, respectively, and contributions of the loadings for the first PCP, while the second and fourth levels (darker shadings) correspond to 15 and 85%, respectively, for the second PCP (see the text for details).

To interpret the derived PCPs of the MD trajectory, the PCA loadings can be graphically displayed. The absolute value of the loading (p_{ak}) tells how much a field data point contributes to a PCP. The PCA-derived loadings were displayed by plotting their contributions within a Cartesian coordinate system. Figure 9 shows a stereoview of two representative conformers from clusters 1 and 6 together with the contributions for the loadings of both PCPs. Of the four patches at the bottom, the first and third ones (counted downward, middle gray shading) correspond to 15 and 85%, respectively, contributions of the loadings for the first PCP, while the second and fourth levels (darker shadings) correspond to 15 and 85%, respectively, for the second PCP. Thus, the main spatial difference expressed as first PCP is the movement of the Dap[3] side chain from the first to the third area during the MD trajectory, while the second PCP can be rationalized as an additional movement of this Dap[3] side chain in another direction (from the second to the fourth area) together with a conformational change at the C-terminal residues. More information on backbone conformational changes is given below.

The dynamic cross-correlation map[50] (DCCM) for dihedral angles (Figure 10) allows to extract correlated torsion fluctuations during the MD simulation in solvent. The type of information is complementary to the molecular shape analysis.

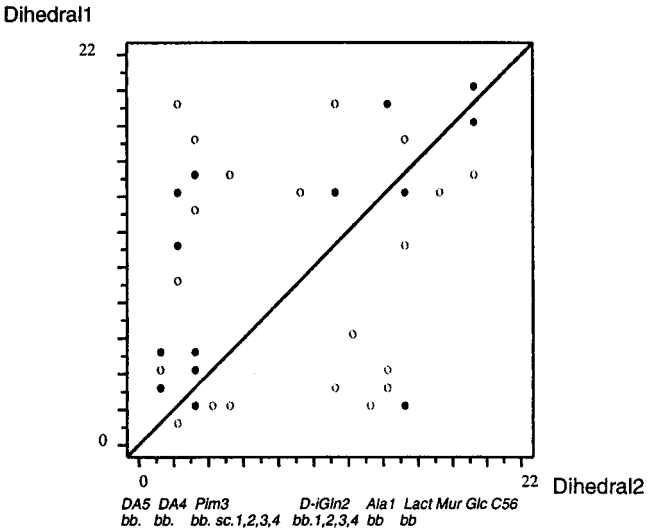


Figure 10. Dynamic cross-correlation map (DCCM) obtained from analyzing dihedral angle fluctuations during the MD trajectory of PGM in water. Crosspeaks above the diagonal correspond to correlated fluctuations, while anticorrelated fluctuations are displayed on the lower diagonal part. Filled circles represent extrema (correlation coefficient higher than 0.55 or lower than −0.71), while intermediate values are shown by open circles. The correlation between numbering of the dihedral angles and the name of the torsion is given in Table 4b. See the text for details.

The upper diagonal part in Figure 10 refers to correlated dihedral angle fluctuations, while anticorrelated phenomena are indicated in the lower part. The time courses of the D-Ala[5]-ϕ and D-Ala[4]-ψ dihedral angles are highly anticorrelated (maximum at 3/2 in the DCCM, cf. Table 4b for the dihedral angle numbering), due to a flexible D-Ala[4]-D-Ala[5] peptide bond, which affects other torsions within those two C-terminal residues (cf. Figure 10 and Table 4). However, this conformational change influences not only the local environment of D-Ala[4] but also the Ala[1]-D-iGln[2] peptide bond (DCCM maximum at 15/14; Ala[1]-ψ/D-iGln[2]-ϕ, cf. Table 4b and Supporting Information) and other torsions in the vicinity of this central residue.

The time evolution for selected anticorrelated dihedral angles is shown in Figure 11. Here, the conformational change involving D-Ala[5]-ϕ and D-Ala[4]-ψ can be seen. The other two dihedral angles displayed in Figure 11 are flanking another amide bond, Ala[1]-D-iGln[2], which is the partner for a hydrogen-bonding interaction during a short period. After ca. 20 ps simulation time, a conformational change corresponding to the opening of a weakly populated hydrogen bond between Ala[1]-C═O and D-Ala[4]-NH occurs. A more detailed analysis of the structural correspondence to the ROE-derived experimental data, the ensemble of SA structures, and the amide proton temperature coefficients shows that only the later part of the MD trajectory is in accord with the experiment. Free MD simulations in solvent also show which possible conformational alternatives are accessible to the C-terminal peptide while maintaining the rigidity of the N-terminal additionally stabilized by the carbohydrate moiety. Since the overall structure in this region is conserved, fluctuations and differences in neighboring ϕ/ψ-angle pairs compensate approximately.

Time−time correlation functions[74] were used to analyze the dynamic behavior of dihedral angle fluctuations for all torsion angles from Table 4b. Selected dihedral angle autocorrelation

(74) (a) Zwanzig, R. *Annu. Rev. Phys. Chem.* **1965**, *16*, 67−102. (b) Kushick, J.; Berne, B. J. *Mod. Theor. Chem.* **1977**, *6*, 41−63. (c) Lautz, J.; Kessler, H.; van Gunsteren, W. F.; Weber, H. P.; Wenger, R. M. *Biopolymers* **1990**, *29*, 1669−1687.

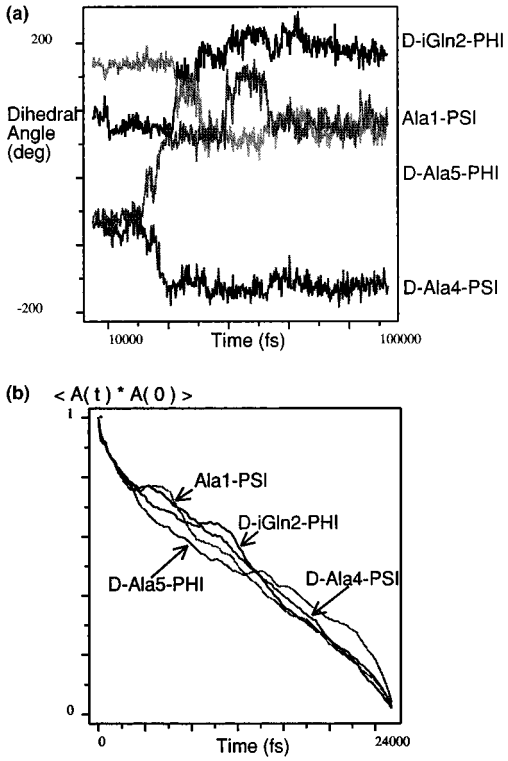


Figure 11. (a) Time evolution of selected dihedral angles (Ala[1]-ψ, D-iGln[2]-ϕ, D-Ala[4]-ψ, and D-Ala[5]-ϕ) from peptidic residues obtained from the MD trajectory of PGM in water. The conformational change between 20 and 30 ps corresponds to the opening of one intramolecular hydrogen bond and the formation of another bond. (b) Autocorrelation functions for the same dihedral angles showing high plateau values corresponding to a complicated dynamic phenomena on an intermediate time scale.

functions are shown in Figure 11b for the peptide torsions involved in this conformational change. The short-time dihedral angle fluctuation was estimated,[75] and computed values for the dihedral angle fluctuation time, τ, are reported in Table 4b (obtained for 2 and 4 ps). In addition to this short-time dynamic descriptor, the plateau value in the range of 3−10 ps for the autocorrelation vector was computed. Large plateau values indicate complicated dynamic processes on a medium-range time scale, as shown in Figure 11. These plateau values correspond to the fluctuations of the dihedral angles obtained from averaging over the MD trajectory. Moreover, they are in agreement with the C-terminal region undergoing conformational changes, such as the formation and opening of a hydrogen bond and the change of the ϕ angle of the "hinge" residue D-iGln[2] (cf. Figure 11).

Those dihedral angles involved in a torsion transition also show shorter relaxation times than conformationally stable dihedral angles: this is shown in Figure 6a with selected exocyclic disaccharide dihedrals, in Figure 6b with the corresponding dihedral angle autocorrelation functions, and in Table 4b for the plateau values and relaxation times of the dihedral angle fluctuations. The fluctuation of these torsions correspond to a short-time process compared to the accessible time scale for the MD simulation, which implies the conformational stability on the medium-range picosecond time scale of the carbohydrate.

Another anticorrelated dihedral angle pair in the N-terminal region can be identified (MurNAc-ϕ'/D-Lac-ϕ), which undergoes a correlated motion with a smaller amplitude than the dihedral angle pairs already mentioned. No major conformational change

(75) van Gunsteren, W. F.; Karplus, M. *Biochemistry* **1982**, *21*, 2259−2274.

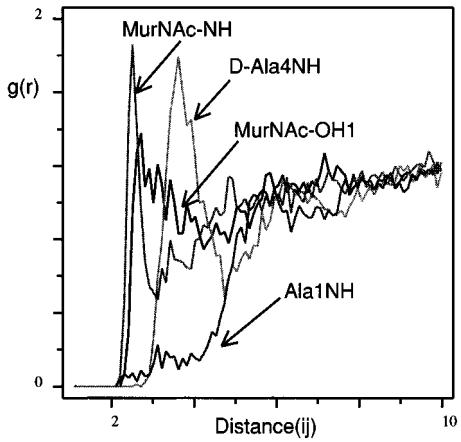


Figure 12. Radial distribution functions (RDFs) computed from MurNAc-NH, MurNAc-OH1, Ala[1]-NH, and D-Ala[4]-NH to solvent oxygen atoms.

can be detected except a dihedral angle change of MurNAc-ϕ' between −160 and −130° during a short period from 50 to 60 ps of the MD trajectory. This change is compensated by the D-Lac-ϕ dihedral angle. The autocorrelation functions of the MurNAc-ϕ' and D-Lac-ϕ dihedral angles reflect this mobility by higher plateau values and shorter relaxation times (cf. Table 4b and Supporting Information), especially for MurNAc-ϕ'.

Radial distribution functions (RDFs) show the variation of the distribution of solvent atoms from a random orientation; they can be used to develop potential parameters for liquids.[76] The RDFs for all polar solute hydrogens to solvent oxygen atoms were computed to obtain information regarding the solvation of PGM. RDFs for selected peptide and saccharide protons are shown in Figure 12, and the distance of the first maximum and its intensity gives some indication about the occurrence of intermolecular solute−solvent hydrogen bonds. From the profiles of the RDFs, the PGM polar hydrogens can be subdivided into four classes: (a) strong hydrogen bond to the solvent (e.g. MurNAc-NH in Figure 12), (b) intermediate hydrogen bond to the solvent (e.g. GlcNAc-OH3, MurNAc-OH1), (c) weak hydrogen bond partners (e.g. D-Ala[4]-NH), and (d) polar solute protons shielded from the solvent (e.g. Ala[1]-NH). Only the latter amide proton is shielded from the solvent, but only partially involved in a low-populated hydrogen bond (<10%), for which there is no experimental indication.

The reduced temperature coefficient for MurNAc-NH could suggest a solvent shielding, but it is not in the range to support the finding of an intramolecular hydrogen bond. However, a strong solute−solvent hydrogen bond is observed for this polar proton during MD simulations (Figure 12). One possible explanation could be that an intramolecular hydrogen bond mediated by a bound solvent molecule is formed as MurNAc-NH···water···MurNAc-O1. During the MD simulation only three weakly populated intramolecular hydrogen bonds were observed (<15% of the simulation time) in agreement with the experimental data.

In Figure 13, the time development of selected interatomic distances within the solute and between solvent and solute is displayed to illustrate the formation of a weakly populated intramolecular hydrogen bond of D-Ala[4]-NH to Ala[1]-CO in the inital phase of the MD simulation and its opening after 20 ps. At the same time, another hydrogen bond (D-iGln[2]-NH to MurNAc-CO) is formed and populated during the remaining simulation time. The formation of intermolecular hydrogen

(76) (a) Mierke, D. F.; Kessler, H. *J. Am. Chem. Soc.* **1991**, *113*, 9466−9470. (b) Rao, B. G.; Singh, U. C. *J. Am. Chem. Soc.* **1990**, *112*, 3803−3811.

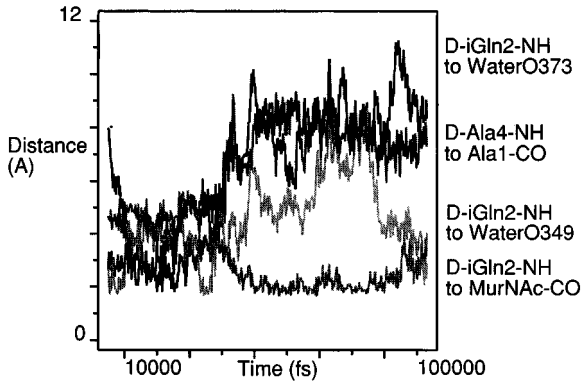


Figure 13. Time development of selected interatomic distances between hydrogen bond donors and possible intra- and intermolecular acceptor atoms during the MD simulation of PGM in water: (1) D-Ala^4-NH to Ala^1-CO, (2) D-iGln^2-NH to MurNAc-CO, (3) D-iGln^2-NH and O349, and (4) D-iGln^2-NH and O373.

bonds between D-iGln^2-NH and oxygen atoms from two different solvent atoms are also displayed in Figure 13 to illustrate the substitution of intermolecular vs intramolecular hydrogen bonds. This formation of solute−solvent hydrogen bonds is important for the stabilization of the glycopeptide structure. It has been noted that conformations of peptides are not stabilized by intramolecular hydrogen bonds in water but by stronger hydrogen bonds to the solvent[77] and those solute−solvent interactions also play an important role in protein conformation and stability.[78]

4. Conclusion

A combined NMR and restrained simulated annealing approach has been employed to elucidate conformational preferences for the biologically relevant peptidoglycan monomer (PGM). SA calculations yielded a set of 11 conformers showing a well-defined N-terminal peptide part connected to a rigid carbohydrate unit, while the C-terminal part exhibits more conformational freedom in accordance with experimental results. Conflicting experimental data for the GlcNAc-N-acetyl group could be explained using a conformer population analysis based on ROE intensities and J(NH,Hα) counting for a conformational equilibrium with one dominantly populated rotamer.

The disaccharide conformation from relaxed conformational maps, computed for the disaccharide model 3-*O*-Me-4-*O*-βGlcNAc-αMurNAc using various force fields, is in agreement with experimental data and literature results. Not only the interglycosidic bond but also the MurNAc to peptide linkage exists only in a single, well-defined conformation, for which no conformational changes on the medium-range time scale could be detected during the MD simulations in aqueous solution. The low-energy conformer exhibits two different lipophilicity sites with residues D-Lac, Ala^1, D-Ala^4, and D-Ala^5 in a lipophilic region and the Dap^3 side chain together with the carbohydrate moiety being in a hydrophilic one.

Unrestrained MD simulations in water gave some insight into the conformation and the dynamic behavior of this glycopeptide. While interpreting these results one should keep in mind that the entire conformational space cannot be thoroughly searched within the MD time frame explored here. As known for other linear peptides,[79] some parts of the structure calculated with experimental restraints proved to be unstable in water; major fluctuations and conformational changes occur during the analyzed MD simulation.

To monitor conformational changes along the MD trajectory, a novel conformer classification method was proposed. Hierarchical cluster analyses were performed using CoMFA molecular steric fields as global shape descriptors. When those steric fields using principal component analysis (PCA) are analyzed, principal conformational properties (PCPs) can be extracted and visualized. This allows one to project the conformational flexibility into a space with lower dimensionality which makes it possible to analyze fluctuations using a plot of the corresponding PCA scores. The present study provides evidence for the conformational stabilizing effects of carbohydrates to glycosylated amino acid sequences.

Acknowledgment. The authors thank Dr. Anne Imberty and Dr. Serge Perez (INRA, Nantes, France) for providing the set of PIM parameters developed for carbohydrates. H.M. thanks Dr. Peter Hecht (Tripos GmbH, München) for stimulating discussions. L.Sz. wishes to thank Dr. Péter Sándor, Dr. Lajos Radics (Varian GmbH, Darmstadt), and Dr. Eric Guittet (ICSN, Gif-sur-Yvette, France) for help with the NMR measurements and the Hungarian National Science Fund (OTKA) for financial support (grant no. T23814). The DRX-500 NMR spectrometer was acquired with the aid of equipment grants from OTKA (no. A084), the National Committee for Technology Development (no. OMFB Mec-930098), and PHARE-Accord (no. H-9112-0198). The 750 MHz ^{1}H NMR spectrum was recorded through the courtesy of Bruker GmbH (Rheinstetten, Germany).

Supporting Information Available: Additional figures and tables listing dihedral angle fluctuations, additional 43 simulated-annealing conformers, and details of the cluster analysis on the SA ensemble including representative conformers for each cluster, additional relaxed conformation maps for MM3(94), respresentative cluster members from the MD trajectory in water and selected dihedral angle autocorrelation functions, a 750 MHz ^{1}H NMR spectrum, and a series of 1D TOCSY spectra of PGM recorded for a solution in H_2O/D_2O (14 pages). See any current masthead page for ordering and Internet access instructions.

JA962776Z

(77) Snyder, J. P. *J. Am. Chem. Soc.* **1984**, *106*, 2393−2400.
(78) (a) Creighton, T. E. *Proteins. Structures and Molecular Properties*; W. H. Freeman and Co.: New York, NY, 1983. (b) Burley, S. K.; Petsko, G. A. *Adv. Protein Chem.* **1988**, *49*, 125−189.
(79) Guba, W.; Haessner, R.; Breipohl, G.; Henke, S.; Knolle, J.; Santagada, V.; Kessler, H. *J. Am. Chem. Soc.* **1994**, *116*, 7532−7540.
(80) Wishart, D. S.; Bigam, G. C.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D. *J. Biomol. NMR* **1995**, *6*, 135−140.
(81) Bystrov, V. *Prog. Nucl. Magn. Reson. Spectrosc.* **1976**, *10*, 41−81.
(82) Pardi, A.; Billeter, M.; Wüthrich, K. *J. Mol. Biol.* **1984**, *180*, 741−751.