ELSEVIER

Carbohydrate Research 256 (1994) 189–222

CARBOHYDRATE RESEARCH

The aqueous solution structure of the tetrasaccharide–ribitol repeat-unit from the lipoteichoic acid of *Streptococcus pneumoniae* strain R6 determined using a combination of NMR spectroscopy and computer calculations

Roger A. Klein [a,*], Rudolf Hartmann [a], Heinz Egge [a], Thomas Behr [b], Werner Fischer [b]

[a] *Institut für Physiologische Chemie der Universität Bonn, Nußallee 11, D-53115 Bonn 1, Germany,*
[b] *Institut für Biochemie der Medizinischen Fakultät der Universität Erlangen-Nürnberg, Fahrstraße 17, D-91054 Erlangen, Germany*

(Received February 9th, 1993, accepted in revised form September 9th, 1993)

Abstract

High-resolution 1D- and 2D-correlation ^{1}H NMR and ^{13}C NMR, at 500 and 125 MHz, respectively, permitted assignment of the majority of the resonances in the per-*N*-acetylated, phosphorylated tetrasaccharide–ribitol repeat-unit, and in the complete polymer ($n = 5–7$) containing between five and seven repeating units attached to the deacylated lipid anchor, for the lipoteichoic acid from *Streptococcus pneumoniae* strain R6; the ^{31}P resonances were also assigned. Comparison of the ^{31}P spectra obtained for the per-*N*-acetylated oligosaccharide and for the oligosaccharide having the AATG 4-NH_2 group still free, indicate a conformational difference brought about by interaction between the amino group and the neighboring phosphate group.

* Corresponding author.

Abbreviations used: AATG, 2,4-diamino-2,4,6-trideoxy-D-galactopyranose; Glc, D-glucopyranose; Gal, D-galactopyranose; GalNAc, 2-acetamido-2-deoxy-D-galactopyranose; PC, phosphorylcholine; RibP, ribitol 5-phosphate; FABMS, fast atom bombardment mass spectrometry; TSP, 4,4-dimethyl-4-silapentanoate-2,2,3,3-d_4; ABNR, adopted-basis Newton–Raphson; CJ, conjugate gradient; SD, steepest descents; TPPI, time proportional phase increment; TIP3, TIP3 water molecule; CHARMm, chemistry at Harvard molecular mechanics force field; MM2, Allinger's molecular mechanics force field; MMX, extended version of MM2 from Serena Software; MCS, Monte Carlo simulation; Φ, (H-1–C-1–O-1–C-X); ψ, (C-1–O-1–C-X–H-X); Ω, (C-1–O-1–C-6–C-5); rms, root mean square; LTA, lipoteichoic acid.

SSDI 0008-6215(93)E0333-V

Collectively these results prove the chain structure previously elucidated by chemical procedures and FAB-mass spectrometry, namely:

-[→6)-β-Glc-(1→3)-AATG-α-(1→4)-[cholineP-6]

-α-GalNAc-(1→3)-[cholineP-6]-β-GalNAc-(1→1)-Ribitol-5P]-

Proton–proton through-space connectivities were determined by NOESY experiments for both the oligomer ($n = 5–7$) and the isolated tetrasaccharide–ribitol repeat-unit. Quantitative estimation of the interproton distances from the NOE measurements allowed determination of the most probable 3D-conformations of the glycosidic linkages of the tetrasaccharide–ribitol repeat-unit in the teichoic acid polymer. Computer calculations by CHARMm molecular mechanics and dynamics methods, Monte Carlo simulation, and with experimental NOE-derived distances as constraints, indicate the most likely solution conformation for the tetrasaccharide–ribitol, taking account of charge, hydrogen-bonding, and solvent effects.

The torsional mobility of the glycosidic linkages between the hexopyranose units is shown to be rather low, whereas the mobility of the final glycosidic linkage in the repeat-unit between the second choline-substituted galactopyranose unit and the ribitol residue is much higher. Two adjacent polymer chains, lying next to one another, are capable of forming a relatively stable, pseudo-helical complex. The biological significance of these findings in terms of aggregation phenomena and the possible role of the side-chain phosphorylcholine moieties is discussed.

1. Introduction

Forssman antigenicity, a general feature of all known strains of *Streptococcus pneumoniae*, has its origin, as shown by serological methods, in the lipoteichoic acid localized on the cell surface of Gram-positive bacteria [1]. The relationship of the lipoteichoic acid, based on its constituent fatty acids, to the pneumococcal teichoic acid, or C-polysaccharide, is clearly established based on the presence in both structures of glucose, 2-acetamido-4-amino-2,4,6-trideoxygalactose, substituted *N*-acetylgalactosamines, ribitol, choline, and phosphate groups, subsequently being identified as a choline-containing lipoteichoic acid [2–6].

We have recently demonstrated, using a combination of chemical and mass-spectrometric techniques, that the structure of the lipoteichoic acid obtained from the cell surface of *Streptococcus pneumoniae* R6 consists of a tetrasaccharide–ribitol repeating unit:

β-Glc-(1→3)-α-AATG-(1→4)-α-GalNAc

-[6PC]-(1→3)-β-GalNAc-[6PC]-(1→1)-Ribitol-5-P

linked to a membrane lipid anchor β-Glc-(1→3)-βAATG-(1→3)-α-Glc-(1→3)-acyl$_2$-glycerol [7] where all sugars are in the pyranose form and belong to the D-series. The repeating units are interconnected by a phosphate diester bond between the 5-position of the ribitol and the primary 6-OH group of the nonreduc-

ing terminal glucose. The complete oligomer, as isolated, contains 5–7 tetrasaccharide–ribitol repeating units and has a molecular weight in the range 7000–8000.

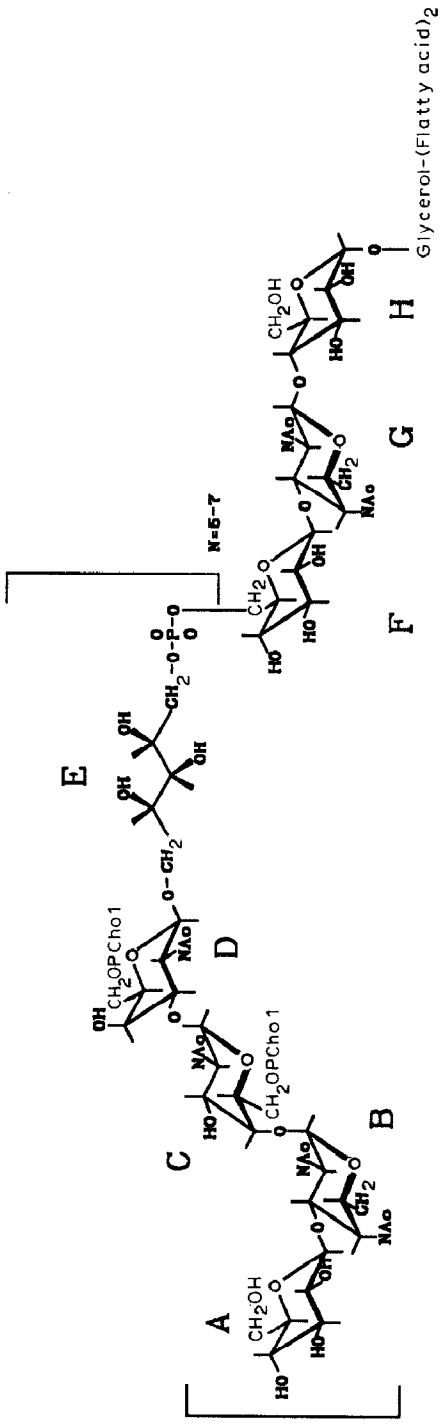
Forssman antigenicity and the C-polysaccharide are typical of the pneumococci, with lipoteichoic acid (LTA) being solely responsible for the expression of Forssman reactivity although it has been shown that lipoteichoic acid and teichoic acids do show serological cross-reactivity [1,8–10]. Forssman activity has been shown to be associated with the bacterial cytoplasmic membrane [10].

Behr et al. [7] have pointed out that the precise identity of the Forssman determinant remains unknown. It is, however, clear from the literature that Forssman haptens from a variety of sources have in common with one another the nonreducing disaccharide moiety α-GalNAc-(1 → 3)-β-GalNAc-(1 → (Refs. [11–15]). All 83 known strains of *Streptococci pneumoniae* express Forssman antigen and C-polysaccharide [9]. These authors have also shown that the pneumococcal teichoic acid cross-reacts with anti-(Streptococcal Group C)-antiserum in 152 out of 164 cases. As Behr et al. [7] point out, although the Forssman disaccharide [α-GalNAc-(1 → 3)-β-GalNAc-(1 →] is a constituent part of the lipoteichoic acid from *Streptococcus pneumoniae*, it is unlikely of itself to show Forssman antigenicity because the galactopyranose units are within the repeat-unit, that is not terminal, and are substituted with phosphorylcholine. This cryptic Forssman reactivity may, however, become apparent after partial degradation, residing in molecular species that contain incomplete phosphorylcholine-free terminal units. Alternatively Forssman antigenicity may be due to lipid-linked biosynthetic precursors, presumably co-extracted with LTA (Behr et al. [7]).

Other properties of the Forssman antigen, such as the effect on the activity [16,17] or cellular distribution [18] of pneumococcal autolytic enzymes, or the activation of the mammalian complement cascade system [19,20], require the presence of phosphorylcholine residues and may, in contrast to Forssman antigenicity, represent biological activities of the lipoteichoic acid itself.

This tetrasaccharide–ribitol repeating unit has a similar structure [7] to that reported for the pneumococcal C-polysaccharide by Jennings et al. [21], with the exceptions that the lipoteichoic acid has both galactosamine residues *N*-acetylated and a second 6-phosphate diester-linked choline substituent attached to the α-GalNAc next to the 2,4-diamino-2,4,6-trideoxygalactose residue. A number of NMR studies of other polysaccharide repeat-units have also been reported for the capsular polysaccharide from various strains of *Streptococcus pneumoniae*, ranging from a simple β-Glc-(1 → 4)-RibP to more complex structures [6,22,23].

The presence of two (1 → 3)-glycosidic linkages to a galactose ring in the repeating unit, one having the α- and the other the β-anomeric configuration, means that it is possible to obtain additional conformational information for these glycosidic linkages by NOE NMR-spectroscopy. Observed interproton through-space connectivities and thus distance information between adjacent residues are usually restricted for glycosidic linkages to that between the anomeric proton H-1 and the H-X proton on the substituted carbon of the aglycon. In the case of a glycosidic linkage to the 3-position of a galactopyranose unit, however, an additional NOE is often observed between the anomeric H-1 and the H-4 of the

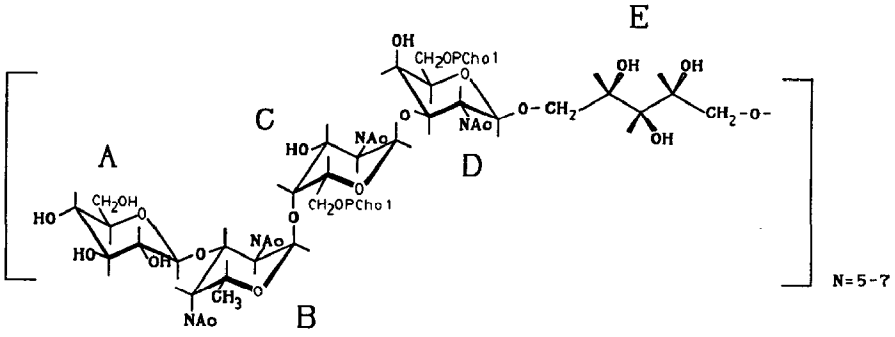

A
B
C
D
E
F
G
H
I
N=5-7
Glycerol-(Flatty acid)2

aglycon ring, as a consequence of the equatorial orientation of the galactopyranose H-4.

The potential availability of two NOE-derived distance constraints for a glycosidic bond greatly decrease the uncertainty in determining the most likely time- and distance-averaged combination of torsional angles [24]. Even if it were possible to derive interproton distances with a high degree of accuracy from measurements of NOE effects, a single quantitative NOE effect only determines which distance contour limits in Φ, ψ space, usually having the form of an elliptical annulus, are appropriate. The presence of a second NOE effect, however, reduces the uncertainty to two intersection points or regions. Under unfavourable geometric conditions it should be noted, in principle at least, that intersecting ellipses can have four real roots. Ambiguities between two possible conformational solutions may be resolved by calculating the energetically most favourable conformer populations and thus the location in conformational space.

In this paper we examine the conformation of the per-*N*-acetylated tetrasaccharide–ribitol repeat-unit obtained from the streptococcal lipoteichoic acid in aqueous solution, using a combination of H,H-distance information experimentally derived from NOE NMR-spectroscopy and computer calculations of the conformer energies. We used the per-*N*-acetylated tetrasaccharide–ribitol, either in the form of the isolated repeat-unit or as part of the complete oligomer containing 5–7 repeat-units, for the studies reported here because the spectral resolution was significantly better than for the oligosaccharide with a free 4-NH_2 group in the AATG moiety. This made the assignment of the resonances easier and more reliable. In a subsequent paper we compare the present results with those for the material containing the free 4-NH_2 group and show that both spectral resolution and overall conformation are affected by an interaction between the free amino group and a neighboring phosphate group.

A comparison of the NOE data for the isolated tetrasaccharide–ribitol repeat-unit and the natural oligomer chain ($n = 5$–7) shows that the conformation of the repeating units changes very little. We therefore extended our conformational analysis by "polymerizing" repeating units into oligomers of a size equivalent to that isolated from the natural source, as well as addressing the possibility of



complex formation between adjacent *n*-oligomers. As an essential part of the conformational study we have assigned all of the conformationally important ^{1}H, ^{13}C, and ^{31}P resonances unequivocally using a variety of homo- and hetero-correlation techniques and in doing so confirmed the structure proposed on the basis of chemical and mass spectrometric measurements for the lipoteichoic acid repeat-unit by Behr et al. [7].

The results and conclusions of the conformational analysis are also relevant for the peptidoglycan-linked teichoic acid or C-polysaccharide of *Streptococcus pneumoniae*. A recent reinvestigation of this polymer revealed a chain structure that is identical in all details with that of the lipoteichoic acid [25].

2. Materials and methods

The oligomer attached to the 2-*O*-deacylated lipid anchor and the tetrasaccharide–ribitol repeat-unit were prepared from *Streptococcus pneumoniae* strain R6 as described previously [7].

NMR spectroscopy.—NMR spectra were determined using a Bruker 500-MHz spectrometer and standard acquisition software. Samples were repeatedly exchanged against 99.96% D_2O by freeze drying (at least three times). For measurement the sample was dissolved in 0.5 mL 99.96% D_2O containing a small amount of the Na^+ salt of TSP as the shift reference. Spectra were recorded at 298 K. H,H-COSY spectra were recorded using standard Bruker software, with a 2048 × 256 matrix and 96 scans, followed by sine-squared multiplication without phase-shifting and finally matrix symmetrization. 2D-NOESY experiments were carried out using the pulse sequence contained in the standard software, with a mixing time satisfying the initial rate approximation (200 ms). A matrix size of 2048 × 512 was used with 256 scans in TPPI mode, followed by zero-filling to 2048 × 2048 and processing with a phase-shifted sine filter. After phase correction in T_1 and symmetrization, cross-peak volumes were determined using the Bruker standard software.

Macura and Ernst [26] have derived the relationship between NOE cross-peak intensities and cross-relaxation rates for a pair of spin $\frac{1}{2}$ nuclei showing through-space dipolar interaction. In the large molecule limit, the cross-peak intensity is an exponential function of the mixing time τ_m. Kumar et al. [27] have shown that this relationship, in the short mixing time limit, is a good approximation for multispin systems. For short mixing times (the initial rate approximation) cross-peak intensities in a NOESY experiment are directly proportional to cross-relaxation rates and thus to the inverse sixth power of the interproton distance. Assuming that identical motions modulate all interactions (this may not be justifiable in a molecule with regions of differing mobility as discussed below), then the ratio of an unknown interproton distance to that of a reference pair is simply the sixth root of the ratio of the cross-peak intensities

$$r_{ij}/r_o = (a_o/a_{ij})^{1/6}$$

In most oligosaccharides found naturally the pyranose rings are held rigidly in the chair conformation, as evidenced from coupling constant data [28,29]. Through-space interactions between ring protons are therefore capable of providing reference distance data for NOE calculations; based upon X-ray crystallographic data [30,32], the pyranose ring protons form suitable reference pairs. In β-D-glucopyranosides and β-D-galactopyranosides the H-1–H-3, H-3–H-5, and H-1–H-5 protons are fixed at distances of ~2.50 Å (Ref. [31]), whereas for α-D-glucopyranosides and α-D-galactopyranosides the H-1–H-2 and H-3–H-5 pairs are 2.50 Å apart [30]. Additionally for α-D-Neu5Ac residues the proton pairs $H\text{-}3_{ax}$–H-5, $H\text{-}3_{eq}$–H-4, and H-4–H-6 provide a distance reference of 2.50 Å (Ref. [33]).

Errors in the measurement of the cross-peak volumes may produce propagated errors in the derived distances; a 30% uncertainty in the intensity of the cross-peak volumes will produce, because of the r^{-6}-dependence, an error of ~5% that is not quite symmetrical. Volume errors are larger the weaker the cross-peak intensity because of the signal-to-noise ratio. For cross-peaks corresponding to distances greater than 3 Å, or with significant overlapping, the volume error may be as high as 40–50%. Estimates of the volume error, based on measurement reproducibility and peak overlap, gave distance uncertainties of typically ±10% at 3.0 Å, with nearer ±5–7% at smaller distances.

Computational methods.—The CHARMm force-field [34] was used for all molecular mechanics and molecular dynamics calculations, as an implementation in conjunction with the Quanta user-interface Version 3.2 (Polygen GmbH, Köln). We also used an extended version of Allinger's MM2 molecular mechanics program called MMX (Refs. [35–37]); as well as the GESA program for HSEA calculations [38,39], modified in this laboratory to improve the grid-search capability. An adapted parameter set suitable for carbohydrates was developed from values in the literature, based on the Jeffrey–Taylor acetal parameter approach [40,41]. Starting coordinates for the heavy atoms of the individual sugar rings were obtained, where available, from either X-ray or neutron diffraction data contained in the Cambridge Structural Database [32]. Additional functional groups and hydrogen atoms were added by means of the CHARMm/Quanta molecular editing facility and the position of the hydrogen atoms refined by energy minimizing the structure obtained.

All computer calculations were carried out on a Silicon Graphics 4D25TG Workstation. In order to obtain a reasonable starting geometry for the tetrasaccharide–ribitol repeat-unit, individual disaccharide units were minimized using a geometrical, near-neighbour distance algorithm followed by at least 100 cycles of a steepest-descent (SD) minimization procedure. Full minimization was then carried out using the adopted-basis Newton–Raphson (ABNR) method or the conjugate-gradient method [42], until convergence was obtained, as judged from the rms deviation in the force value. Atomic charges, determined using the Gasteiger and Marsili method [43] and hydrogen-bonding terms were included in the minimization. The unit charge carried by the phosphate and the quaternary ammonium groups of the choline residues corresponded to a pH value around neutrality.

The Φ angle for a normal glycosidic linkage is defined by the torsional set H-1–C-1–O-1–C-X, where C-1 is the anomeric carbon; the ψ angle is defined by the set H-X–C-X–O-1–C-1, in which C-X is the substituted carbon in the aglycon. In the case of a 6-linked residue an additional angle, Ω, is defined by the set O-1–C-6–C-5–O-5. For the oligomer, and indeed the tetrasaccharide–ribitol repeat-unit, the torsional angles are defined and abbreviated numerically in the following manner for the four glycosidic linkages within the repeat-unit:

$$\Phi(\text{H-1–C-1–O-1–C-X}) = \text{tor}\left[(8 \cdot (\text{R} - 1)) + (2 \cdot \text{S} - 1)\right], \text{ i.e., tor 3,}$$

$$\psi(\text{H-X–C-X–O-1–C-1}) = \text{tor}\left[(8 \cdot (\text{R} - 1)) + 2 \cdot \text{S}\right], \text{ i.e., tor 4,}$$

where S is the sugar or glycosidic bond starting from the nonreducing end of the repeat tetrasaccharide–ribitol, and R is the sequence number of the repeat-unit within the oligomer, again starting from the nonreducing β-Glc-(1 → 3) end.

An initial search of conformational space for each glycosidic linkage in either the disaccharide or trisaccharide containing the linkage, was carried out using a grid search of the Φ and ψ angles using a 10° search increment, in order to identify appropriate areas of conformational space. A variety of techniques are available for searching conformational space in addition to grid searching. Although a grid search, either as a rigid or relaxed procedure, may be used to systematically sample conformational space, a full search using an increment of 2.5°, representing a compromise between resolution and computational time, and full energy minimization at every point, is extremely time-consuming for all but disaccharides or possibly trisaccharides and becomes rapidly impractical as the degrees of freedom increase with larger oligosaccharides. In individual cases, a grid search was performed for glycosidic linkages in the complete tetrasaccharide–ribitol with a restricted angular window in order to confirm the results with the smaller disaccharides or trisaccharides. The minima observed with di- and tri-saccharides were used as starting geometries for minimizations involving the entire molecule. For larger structures, such as the oligomer studied in the present work, Monte Carlo simulation or molecular dynamics simulation are more efficient [52], in terms of computer time, in sampling multi-dimensional conformational space than grid searching. It has been our experience that the Boltzmann jump procedure with minimization (within Quanta/CHARMm) or MD simulation for the 5-oligomer may be achieved on approximately the same time-scale as a fully relaxed grid search for the tetrasaccharide–ribitol repeat-unit alone, i.e., a gain of approximately fivefold in CPU time spent making the calculation.

Solvent (water, D_2O) effects were included in the calculations either implicitly using the distance-dependent dielectric constant option for correcting for effects of solvation near to charged groups [34], or explicitly by carrying out the calculations in a TIP3 water box of 30 Å side [44]. The presence of one molecule of solute in such a $10 \times 10 \times 10$ cube of water molecules represents a not unrealistic concentration of 55 mM which, for an oligosaccharide with a molecular weight of ~ 1000, corresponds to a 5% w/v solution. NOE-derived interproton distance constraints, obtained experimentally by integrating the NOESY cross-peak volumes and appli-

Table 1
Chemical-shift assignments for the isolated tetrasaccharide–ribitol repeat-unit containing phosphorylcholine groups but without the 5-position of the ribitol phosphorylated

Atom	Residue					Choline
	A	B	C	D	E	
						(C)
H-1	4.61	4.99	5.21	4.67	3.93;4.00	$CH_3N = 3.30$
H-2	3.28	4.38	4.38	4.16	4.00	$OCH_2 = 4.35$
H-3	3.51	4.33	3.95	3.92	3.75	$NCH_2 = 3.74$
H-4	3.37	4.55	4.13	4.23	3.83	
						(D)
H-5	3.46	4.68	4.08	3.88	3.68;3.83	$CH_3N = 3.30$
H-6a	3.70	1.07	4.07	4.11		$OCH_2 = 4.39$
H-6b	3.92	(CH_2)				$NCH_2 = 3.75$
H-6c						
C-1	104.49	98.79	93.63	101.56	70.98	$CH_3N = 54.26$
C-2	73.31	49.38	49.72	50.89	70.87	$OCH_2 = 59.75$
C-3	76.06	75.55	67.22	74.90	72.31	$NCH_2 = 66.30$
C-4	69.67	53.63	76.71	63.44	72.24	
C-5	76.05	66.48	71.09	73.80	62.81	
C-6	61.25	15.88	64.00	64.68		
PO_4			1.35	1.83		

cation of the NOE r^{-6}-dependence, were used to limit the number of conformational possibilities for the glycosidic linkages during minimization.

By computationally polymerizing the energy-minimized tetrasaccharide–ribitol repeat-unit, we generated structures consisting of up to 5 repeat-units equivalent to the naturally occurring oligomer; this represents about the practical limit in terms of the subsequent requirement for CPU time during minimization or simulation. The 5-oligomer was energy minimized and the most stable conformation identified; this structure was then used to study the possibilities of complex formation between adjacent oligomers.

3. Results and discussion

The bis-*(phosphorylcholine)-tetrasaccharide–ribitol.*—The assignments for the ^1H, ^{13}C, and ^{31}P resonances are shown in Tables 1, 2, and 3. Based upon the 500-MHz H,H-COSY and ^1H-detected C,H-COSY (HMQC) spectra (not shown) it is possible to assign unequivocally all of the resonances for the four hexopyranose residues of the tetrasaccharide–ribitol repeat-unit. Signals for α-anomeric protons are observed at 4.99 ppm ($^3J_{H,H}$, 4.0 Hz) and 5.21 ppm ($^3J_{H,H}$, 3.9 Hz), and for the β-anomeric protons at 4.61 ppm ($^3J_{H,H}$, 7.8 Hz) and 4.67 ppm ($^3J_{H,H}$, 8.6 Hz). The 6-methyl group resonance at 1.07 ppm, belonging to the 2,4-acetamido-2,4,6-trideoxygalactose residue with its associated anomeric proton at 4.99 ppm, served to

Table 2
Chemical-shift assignments for the phosphorylated tetrasaccharide–ribitol repeat-unit contained within the oligomer $n = 5–7$. Chemical shifts shown in italics are for residues contained within the nonreducing end terminal repeat tetrasaccharide–ribitol

Atom	Residue					Choline
	A	B	C	D	E	
						(C)
H-1	4.63	5.01	5.22	4.68	3.92;4.01	
	4.57					$CH_3N = 3.30$
H-2	3.34	4.40	4.39	4.18	4.06	
	3.39					$OCH_2 = 4.35$
H-3	3.55	4.36	3.96	3.92	3.82	$NCH_2 = 3.74$
H-4	3.61	4.51	4.14	4.23	3.95	
						(D)
H-5	3.55	4.67	4.08	3.89	4.15	$CH_3N = 3.30$
H-6a	4.18	1.11 (CH_3)	4.04	4.12		$OCH_2 = 4.39$
H-6b						$NCH_2 = 3.75$
H-6c						$NCH_2 = 3.75$
C-1	*103.56*	98.56	93.55	101.45		$CH_3N = 54.20$
	103.45					
C-2	72.35	49.39	49.74	50.86	66.90	$OCH_2 = 59.80$
	71.03					
C-3	75.83	75.28	67.19	75.00	71.76	$NCH_2 = 66.30$
C-4	69.00	52.98	76.55	63.46	71.00	
C-5	75.10	66.77	71.00	73.74	66.90	
C-6	64.35	15.92	64.00	64.85		
PO_4			1.45	1.85	3.60	

Table 3
Chemical-shift assignments for the deacylated lipid anchor attached to the native oligomer

Atom	Residue			
	F	G	H	I
H-1	4.63	4.75	4.94	
H-2	3.34	4.05	3.67	
H-3	3.55	4.21	3.90	
H-4	3.61	4.42	3.53	
H-5	3.55	3.97		
H-6a	4.18			
H-6b		1.23		
H-6c				
C-1	103.56	102.30	99.05	
C-2	72.35		71.34	
C-3	75.83		82.39	
C-4	69.00	53.05	68.60	
C-5	75.10	70.15		
C-6	64.35	16.05		

differentiate the two α-galactopyranosyl spin systems. The β-sugars were distinguishable based upon the absence of the glycosylation shift for the glucopyranose ring. Assignment of the resonances for the ribitol unit (E) was achieved by comparison with the spectrum of free ribitol 5-phosphate. The ^{31}P resonances at 1.45, 1.85, and 3.60 ppm, measured at 200 MHz, were assigned using $^{31}P-^{1}H$ correlation spectroscopy as arising from the choline phosphate on residues C and D, and the ribitol phosphate respectively. The carbohydrate sequence followed directly from the cross-peaks observed in the 2D-NOESY spectrum. ^{13}C Chemical shifts were obtained and assigned using a C,H-correlation HMQC experiment with ^{1}H-detection.

*The [*bis-*(phosphorylcholine)-pentamer]-oligmer* (n = 5–7).—Because of the small differences in chemical environment, one would not expect to be able to resolve the proton resonances of individual repeat-units. The exception is for the terminal glucose residue which, because of the missing phosphate substituent at the 6-CH_2OH position, is observed as an additional sugar unit without difficulty. The resonances of both β-glucose residues, one in the repeat-unit and the other in the lipid anchor, coincide. The two other sugars in the lipid anchor, the β-2,4-diacetamido-2,4,6-trideoxygalactose and the α-glucose, would at any rate be expected to be of low intensity. Through a comparison of the 500-MHz H,H-COSY spectrum for the oligomer with that for the free pentamer, one obtains very good consistency between the chemical-shift values for sugar units B–D in the tetrasaccharide–ribitol repeat-unit. The remaining three sugar moieties, with reduced signal intensities, can be distinguished based on their anomeric proton resonances and the 6-methyl group. The ribitol 5-phosphate unit, as well as the assignment of ^{13}C chemical shifts, was achieved through HMQC experiments.

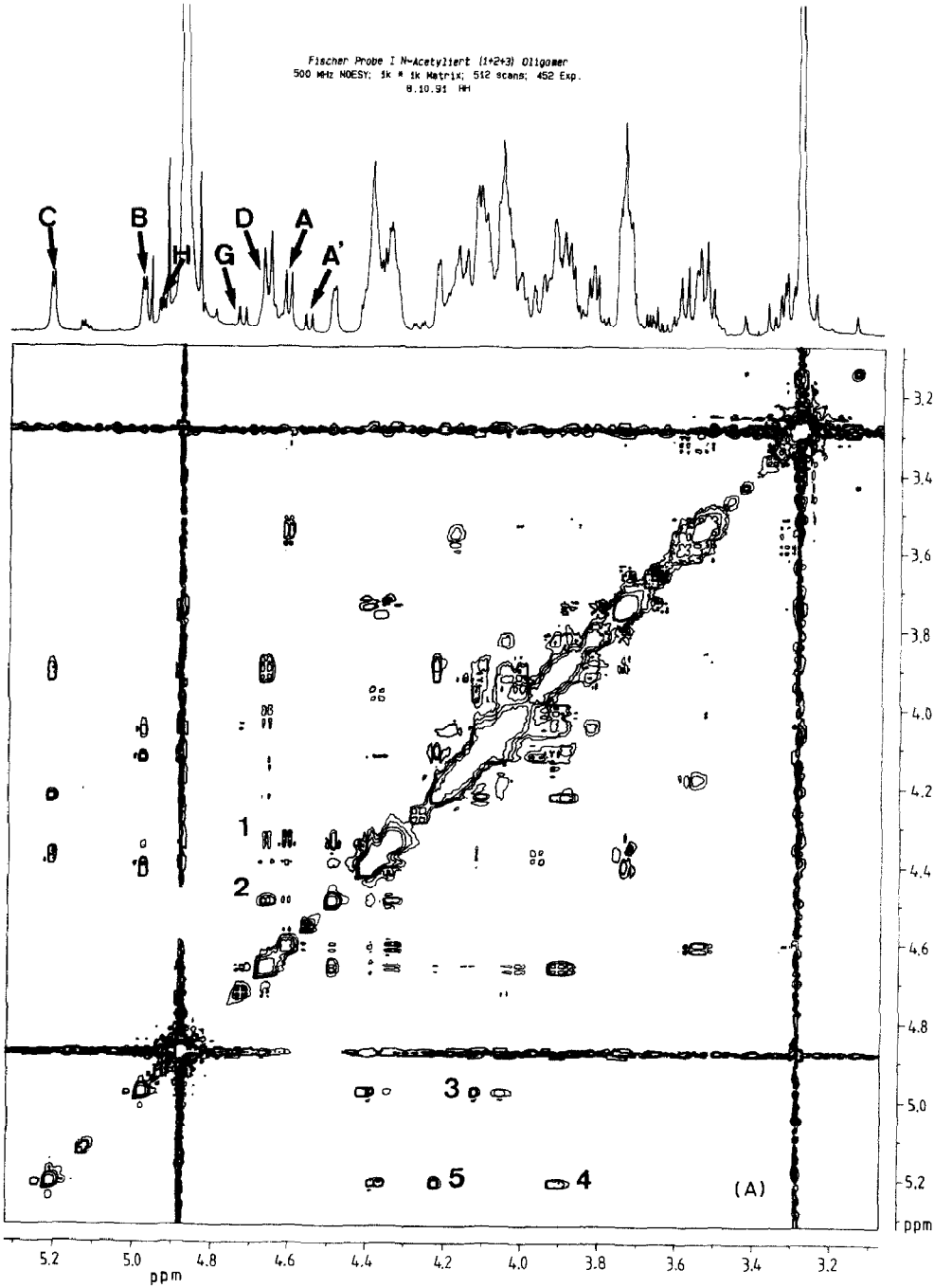
NOE through-space connectivities.—The 2D-NOESY spectra of (A) the oligomer and (B) the tetrasaccharide–ribitol repeat-unit cleaved at the phosphate diester linkage between the ribitol and glucose units are shown in Fig. 1. The interproton distances derived from integration of the 2D-NOESY cross-peaks for the tetrasaccharide–ribitol repeat-unit are given in Tables 4A and 4B using reference distances as indicated in the legends.

In the limit of the initial rate approximation the cross-peak intensity is proportional to the mixing time, τ_c:

$$I_{AB}(\tau_m) \propto \tau_c \cdot \tau_m / \mathbf{r}_{AB}^6.$$

The correlation time, τ_c, refers to the reorientation of the internuclear vector, r_{AB}. In systems with isotropic reorientation, such as globular proteins or otherwise rigid structures, τ_c is usually assumed to be equal for all AB proton pairs. This

Fig. 1. 2D-NOESY spectrum at 500 MHz for the isolated per-*N*-acetylated oligomer (A) and the tetrasaccharide–ribitol repeat-unit (B) in aqueous solution at 300 K. The individual NOESY cross-peaks labelled 1–5 are those shown with the same numbering in Table 4; these cross-peaks are the conformationally most important ones. The anomeric protons are indicated using the same alphabetical sugar unit labelling as in Table 4; the use of a prime, e.g., A′, indicates that the sugar residue belongs to the nonreducing end terminal pentamer.


Fischer Probe I N-Acetyliert (1+2+3) Oligomer
500 MHz NOESY; 1k * 1k Matrix; 512 scans; 452 Exp.
8.10.91 HH
C
B
H
G
D
A
A'
1
2
3
5
4
(A)
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
ppm

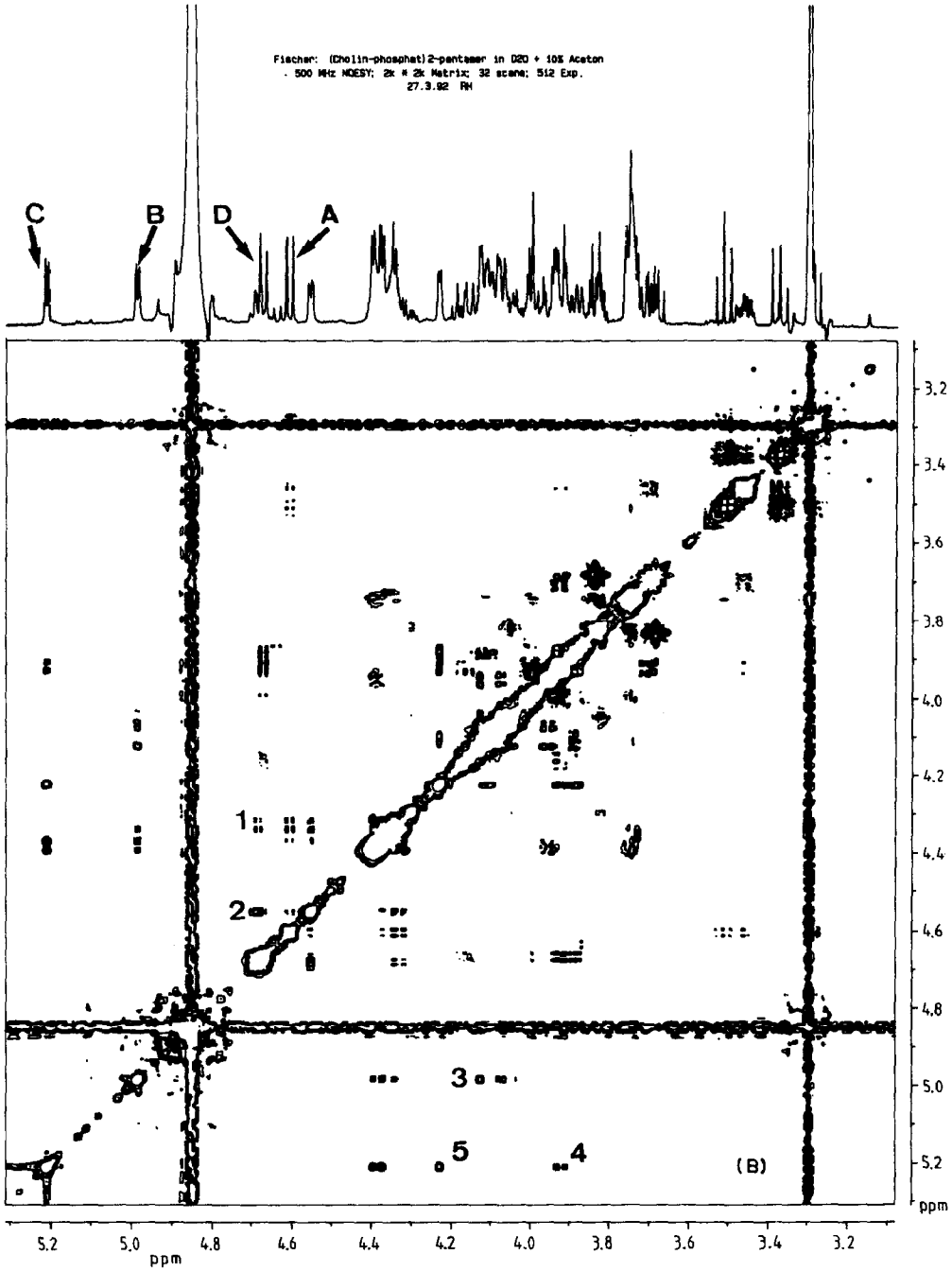


Fig. 1 (continued).

Table 4A
The effect of mixing times between 100–250 ms on the interproton distances calculated from the integration of important cross-peaks in the 2D-NOESY spectrum of the tetrasaccharide–ribitol repeat-unit

Cross-peak [a]	Mixing time (ms)				Average
	100	150	200	250	
C-1 → D-4	2.16	2.14	2.21	2.21	2.18 ± 0.03
C-1 → D-3	2.38	2.33	2.38	2.37	2.36 ± 0.02
B-1 → C-4	2.15	2.14	2.19	2.20	2.17 ± 0.03
B-1 → C-5	2.36	2.34	2.38	2.38	2.37 ± 0.02
	2.47	2.44	2.46	2.46	2.46 ± 0.01
A-1 → B-3	2.38	2.30	2.35	2.33	2.34 ± 0.03
C-1 → D-3	2.42	2.38	2.43	2.44	2.42 ± 0.03

[a] All distances in Å were calculated based on the H-1 → H-3 and H-1 → H-5 distance of 2.50 Å in β-glucopyranosides, using cross-peaks from residue A (glucose), H-1–H-3 and H-1–H-5. Distances were calculated for cross-peaks in nonsymmetrized spectra using reference signals from the same region of the 2D-spectrum. Relative NOE effects were normalized using the diagonal peak intensities for the anomeric protons in residues A, B, and C.

assumption certainly does not hold for mobile side-chains, which, incidentally, may not even fulfil the slow-motion approximation ($\tau_c \gg \omega_0^{-1}$), for example, in the present case, the 2-acetamido or choline phosphate groups as seen in MD simulations.

If the mixing time, τ_m, is increased beyond the range of the initial rate approximation, it becomes necessary to consider the higher terms of the expansion of $\exp(L\tau_m)$, rather than the use of the approximation

$$\exp(L\tau_m) \approx 1 + L\tau_m$$

For sufficiently short mixing times, however, approximate expressions for the amplitude can be derived by invoking the initial rate approximation. In the initial rate approximation, NOE cross-peaks increase in a monotonic fashion until they

Table 4B
NOE Cross-peaks observed in the 500 MHz 2D-NOESY spectrum of the repeat tetrasaccharide–ribitol, measured in D_2O at 303 K. H–H Distances were calculated using the H-1–H-3–H-5 distances in GalNAc and the H-1–H-5 distance in AATG as references; values shown are reliable to ~ ±5–10%

Peak	PPM(1)	PPM(2)	Distance	Assigned H(1) → H(2) [a]
1	4.625	4.363	2.48	A-1 → B-3
2	4.625	4.507	2.72	A-1 → B-4
3	5.012	4.142	2.26	A-1 → C-4
4	5.222	3.922	2.33	C-1 → D-3
5	5.222	4.230	2.35	C-1 → D-4

[a] Residue: A, β-Glc; B, 2,4-diamino-2,4-dideoxy-β-Gal2NAc; C, α-GalNAc-(6-PC); D, β-GalNAc-(6-PC); E, ribitol 5-phosphate.

become damped by external relaxation processes. Second-order Overhauser effects may produce a sigmoidal relationship. The detailed basis for the initial rate approximation is discussed in the monograph by Ernst et al. [45].

Important experimental consequences of the initial rate approximation, particularly if it applies to all the interproton vectors and thus NOE cross-peaks in the spectrum of interest, are that: (*i*) the normalized integrated cross-peak volumes should increase monotonically with mixing time; (*ii*) the calculated interproton distances, based on the r^{-6th}-dependence of the NOE effect and suitable reference distances, should not depend on the mixing time, within experimental error; (*iii*) the relative cross-peak volumes for reference proton pairs with the same interproton distance, for example 2.50 Å for the H-1–H-3 and H-1–H-5 protons in β-gluco- and galacto-pyranosides, provide evidence for any differences in correlation time, τ_c, within the molecule.

Distances derived experimentally from the integration of 2D-NOESY cross-peak volumes represent a weighted-average of all conformational states based on their probability, i.e., energy differences, and the sixth-power dependence of the interproton distances.

In Tables 4A and 4B we show NOE-derived distances for the repeat-unit containing the sugar residues A–D and the ribitol. Table 4A contains interproton distances determined using the H-1–H-3 and H-1–H-5 proton pairs of the nonreducing terminal β-D-glucose unit (residue A) as reference distances, at four different mixing times in the range 100–250 ms. Plots of the relative NOE effect for each proton pair, normalised against a set of diagonal peaks, versus mixing time showed, since this relationship was linear and monotonic without evidence of saturation for all the peaks integrated, including the reference peaks, that the measurements had been conducted under the conditions of the initial rate approximation. A consequence of this condition being met is that distances determined at the four different mixing times are seen to be indistinguishable from one another, within experimental uncertainty, as shown in Table 4A.

A comparison of the relative values for the H-1–H-3 and H-1–H-5 reference cross-peak volumes in the terminal β-glucose (A) and in the phosphorylcholine-substituted *N*-acetyl-β-galactosamine (D) linked to the ribitol, provided evidence that the correlation time, τ_c, for residue D was approximately twice that for residue A. Because residues B and C are both α-galactopyranose rings, the well-resolved H-1–H-2 reference cross-peaks could be used to determine the relative correlation times for these sugars. A reasonable assumption is that a gradient exists in the time for internuclear vector reorientation, i.e., the correlation time, τ_c, on passing from A → B → C → D and that, intuitively at least, the terminal glucose unit (A) is likely to be more mobile than the β-galactopyranose in the middle of the repeat-unit. Integration of the cross-peak volumes for the reference H–H pairs in residues A, B, C, and D, give the following results for the relative correlation times: $A = 1.0$; (1.0); $B = 2.1 \pm 0.3$ (1.13); $C = 2.3 \pm 0.4$ (1.15); $D = 2.1 \pm 0.1$ (1.13). The value in parentheses is, in each case, the factor for correcting the derived H–H distances. The relative correlation times for residues B, C, and D are the same within experimental error, suggesting that this section of

the repeat-unit is fairly stiff. The correlation time for terminal residue A, β-D-glucose, is approximately half that of the other residues.

Table 4B shows the same distances as in Table 4A but determined using different proton pairs as the reference cross-peak volumes. In this case the H-1–H-3 and H-3–H-5 pairs from the β-galactopyranose ring (residue D) linked to ribitol were used. As expected, because of the difference in correlation time, the distances are ~ 10–15% greater than those in Table 4A. In the discussion which follows below, the NOE-derived distances for the β-(1 → 3) linkage between residues A and B has been corrected for the shorter correlation time of the first residue in the repeat-unit.

The anomeric resonances for the individual sugars and the conformationally important cross-peaks are indicated in both Figs. 1A and 1B. On comparing Figs. 1A and 1B it is quite clear that, apart from minor differences in chemical-shift values, the spectra are very similar and that the same NOE cross-peaks are observed both for the oligomer and for the tetrasaccharide–ribitol. This important observation means that the conformation of the tetrasaccharide–ribitol moiety in aqueous solution does not change, within the sensitivity of our measurements, irrespective of whether it is present in the form of the oligomer ($n = 5–7$), or as the isolated oligosaccharide. The resolution in the spectrum of the tetrasaccharide–ribitol is considerably better than for the oligomer because of overlapping resonances with similar, but not actually identical, chemical shifts for the repeated sugar units in the latter.

Calculation of conformational preferences.—Our general approach to the calculation of conformational preferences has been to use a starting geometry determined using experimentally derived distance constraints, as shown in Table 4. An initial minimization was carried out using preliminary SD minimization to remove bad contacts, followed by the ABNR method, with a high degree of constraint applied to the specified H–H distances by means of the CHARMm SCALE command. A full ABNR minimization was then carried out under less constrained conditions, with a very much reduced SCALE factor, and an allowed distance deviation of ±0.3 Å (~ 10–15%) equivalent to the standard error. The conformation thus obtained is referred to as the "constrained" conformation. All distance constraints were then completely removed and the structure reoptimized using the ABNR method. Hydrogen-bonding terms were included and partial atomic charges derived according to the method of Gasteiger and Marsili [43]. The structure obtained in the absence of distance-constraints is subsequently referred to as having the "free" or "unrestrained" conformation. Restrained and unrestrained conformational calculations were also carried out in the presence of a 30 Å-sided cube of TIP3 water molecules in order to simulate solvent effects and particularly solute–solvent interactions. A single molecule of oligosaccharide, or for that matter any structure, "dissolved" in a 30 Å water box is equivalent to a solute concentration of ~ 55 mM.

The β-D-Glc-(1 → 3)-D-AATG linkage.—The experimentally determined NOE-derived interproton distances for the H-1A–H-3B and H-1A–H-4B proton pairs were found to be 2.18 ± 0.25 and 2.39 ± 0.25 Å, respectively, both corrected for

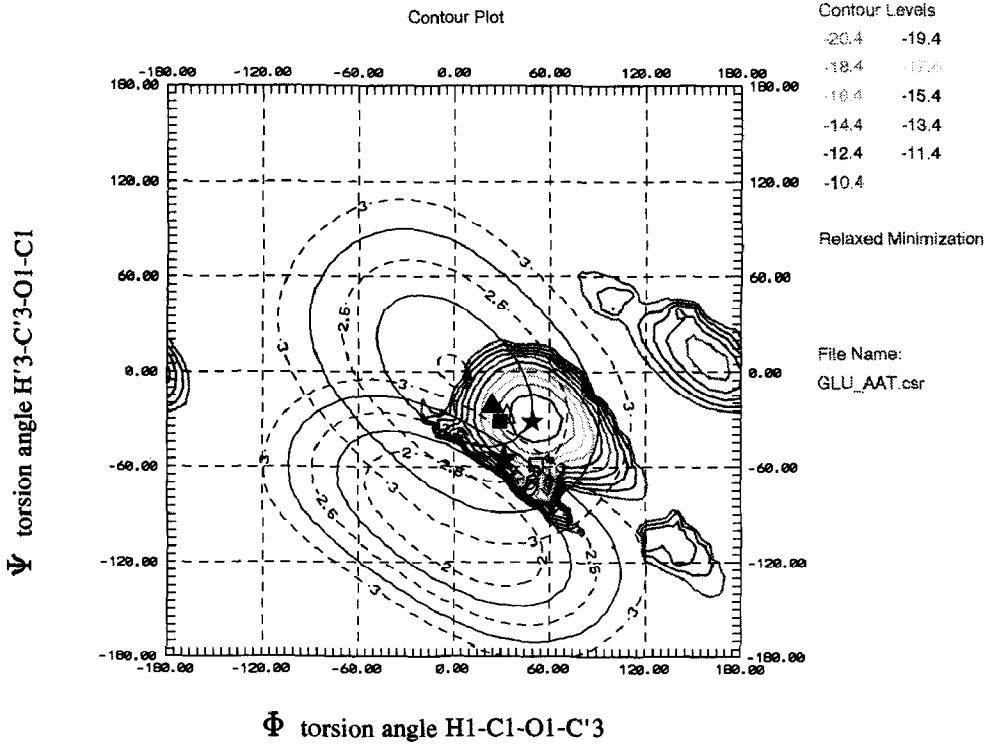


Fig. 2. Contour plot of the Φ, ψ torsional angles against energy for the β-Glc-(1 → 3)-AATG linkage determined using the CHARMm force-field with a 5° increment grid search as described in the text. Individual minimization results are indicated for CHARMm, unrestrained (△), with NOE distance constraints (□), unrestrained in the presence of a 30 Å TIP3 water box (▲), and with NOE distance constraints in a 30 Å TIP3 water box (■); and for the MMX force-field (★). Iso-energy contours are shown at 1 kcal/mol intervals. The distance contours for the β-(1 → 3) and β-(1 → 4) interproton spacings are also shown, for distances between 2 and 3 Å at 0.25 Å intervals.

correlation time (Table 4B – see above). Estimation of the H-1A–H-3B distance at four different mixing times using reference peaks within the β-glucose unit, gave values of 2.27 and 2.34 Å. The differences between the values for the distance H-1A–H-3B of 2.18, 2.27, and 2.34 Å lie within probable experimental error. These results imply that the torsional angles Φ and ψ for the β-(1 → 3)-glycosidic linkage must satisfy these two distance constraints simultaneously, as illustrated in the contour map [53] shown in Fig. 2, and must therefore lie within an elliptical area between approximately +25°, −50° and −10°, −40°. Calculation of the minimum-energy conformations of this linkage, using the CHARMm force field and a grid search with a 10° increment, gave the contour plot shown in Fig. 2. Minima are present at +50°, −30° (0 kcal/mol relative); +165°/+10° (+5 kcal/mol), and +140°, −105° (+8 kcal/mol). The broad minimum centered at +50°, −30° appears to have a subsidiary and rather sharp minimum at +60, −70° within ±0.2 kcal/mol; although this additional minimum was preserved on reduc-

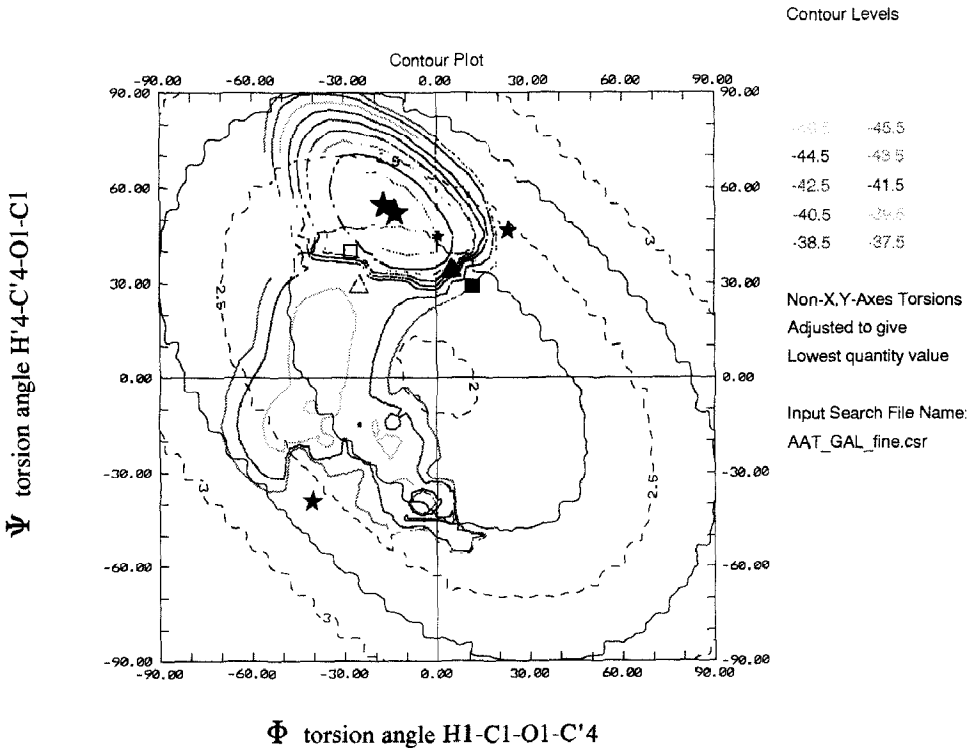
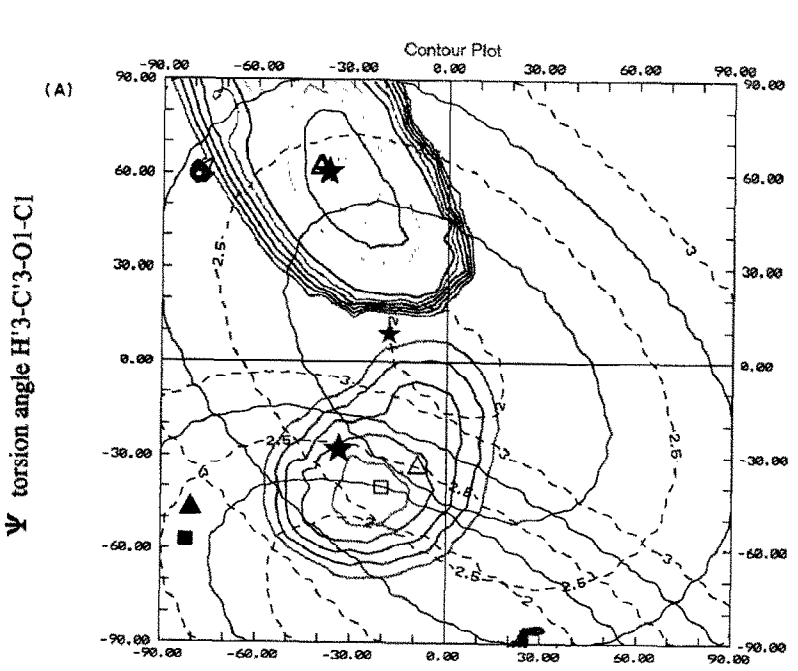
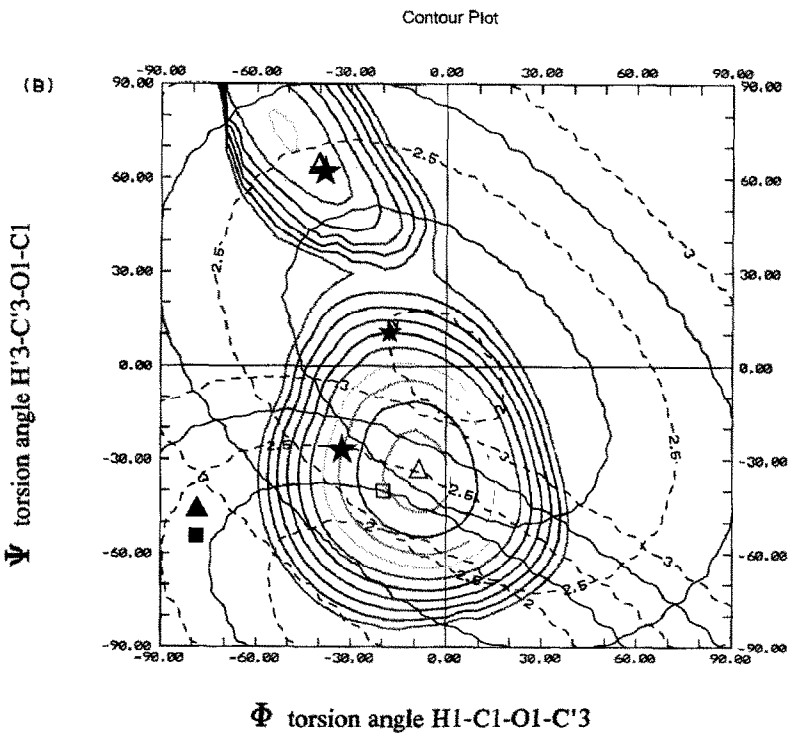


Fig. 3. Contour plot of the Φ, ψ torsional angles against energy for the α-AATG-(1 → 4)-GalNAc[6PC] linkage determined using the CHARMm force-field with a 5° increment grid search as described in the text. Individual minimization results are indicated for CHARMm, unrestrained (△), with NOE distance constraints (□), unrestrained in the presence of a 30 Å TIP3 water box (▲), and with NOE distance constraints in a 30 Å TIP3 water box (■); and for the MMX force-field (★). Iso-energy contours are shown at 1 kcal/mol intervals. The distance contours for the α-(1 → 4) interproton spacing is also shown for distances between 2 and 3 Å at 0.25 Å intervals. The wavy nature of the iso-energy contour lines is an artefact brought about by expansion of the grid.

Fig. 4. (A and B) Contour plot of the Φ, ψ torsional angles against energy for the α-GalNAc[6PC]-(1 → 3)-GalNAc[6PC] linkage determined using the CHARMm force-field with a 5° increment grid search as described in the text. Individual minimization results are indicated for CHARMm, unrestrained (△), with NOE distance constraints (□), unrestrained in the presence of a 30 Å TIP3 water box (▲), and with NOE distance constraints in a 30 Å TIP3 water box (■); and for the MMX force-field (★). Iso-energy contours are shown at 1 kcal/mol intervals. The distance contours for the α-(1 → 3) and α-(1 → 4) interproton spacings are also shown, for distances between 2 and 3 Å at 0.25 Å intervals. (A) is the contour map determined for the disaccharide and (B) the contour map for the entire tetrasaccharide–ribitol repeat-unit. (C and D) Contour plot of the Φ, ψ torsional angles against energy for the α-GalA-(1 → 3)-(2-acetamido-4-amino)-2,4,6-trideoxy-Gal linkage determined using the HSEA force-field with a 1° increment grid search. An explicit exoanomeric energy term was used in C and the Burckert torsional energy term in D. The position of the single minimum as determined by Bock et al. [7] is shown by the asterisk (★).


(A)
Contour Plot
Contour Levels
-50.5 -49.5
-48.5 -47.5
-46.5 -45.5
-44.5 -43.5
-42.5 -41.5
-40.5
Non-X,Y-Axes Torsions
Adjusted to give
Lowest quantity value
Input Search File Name:
GAL_GAL_fine.csr
Ψ torsion angle H'3-C'3-O1-C1
Φ torsion angle H1-C1-O1-C'3


(B)
Contour Plot
Contour Levels
-129.5 -128.5
-127.5 -126.5
-125.5 -124.5
-123.5 -122.5
-121.5 -120.5
Non-X,Y-Axes Torsions
Adjusted to give
Lowest quantity value
Input Search File Name:
GAL_GAL_PENTA_fine
Ψ torsion angle H'3-C'3-O1-C1
Φ torsion angle H1-C1-O1-C'3

ing the grid search increment to 2.5°, it is almost certainly an artefact based on its sharpness alone.

Calculation of the conformational energy using the MMX force field, which is an extended version of Allinger's MM2 programs [35,36], clearly demonstrated two energy minima, one at +33, −55° (0 kcal/mol) and the other at +49, −32° (+0.53 kcal/mol). The global energy minima calculated using the CHARMm force field in the presence and absence of experimentally derived distance constraints, as well as without or with a 30 Å TIP3 water box, gave Φ, ψ values for this linkage of (+51, −60°), (+33, −31°), (+30, −33°), and (+26, −23°), respectively, for the intact tetrasaccharide–ribitol.

The results obtained from CHARMm and MMX force-field calculations are shown superimposed on the contour map. Extensive MMX calculations using various conformer starting geometries for the di-, tri-, and tetra-saccharide containing the β-Glc-(1 → 3)-AATG linkage, gave two sets of Φ and ψ angles of (+28 ± 2°, −52 ± 2°) and (+34 ± 2°, −55 ± 2°) corresponding to relative con-

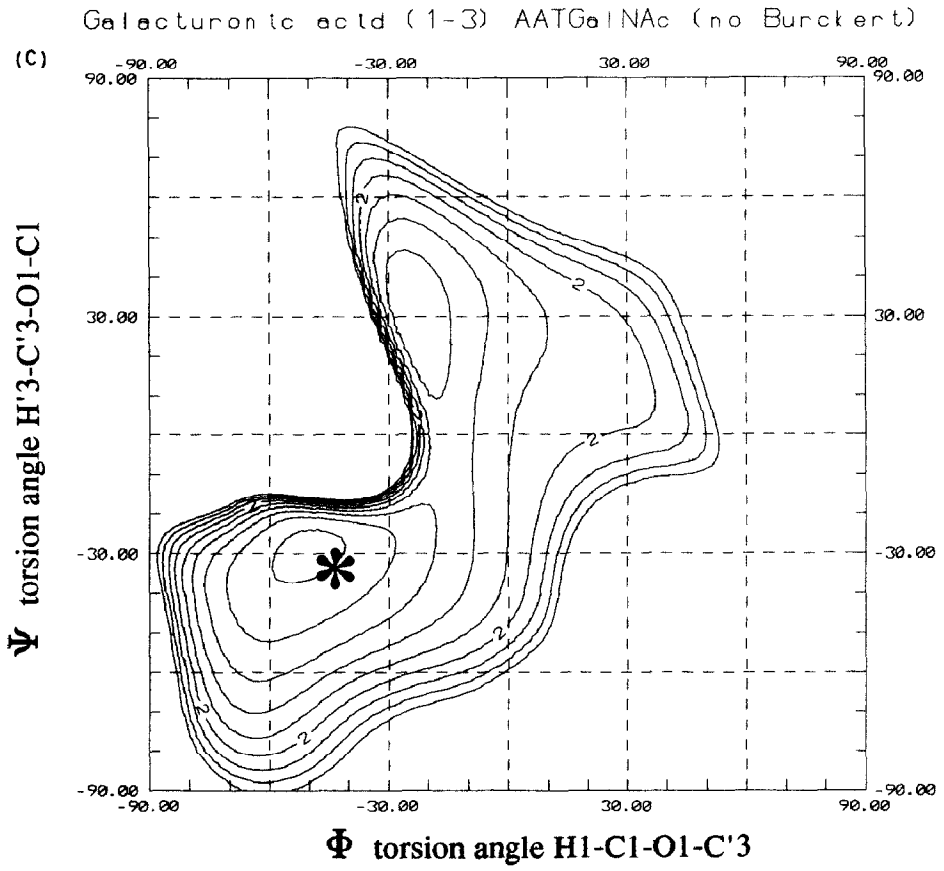


Fig. 4 (continued).

former populations of ~30% and 70%, respectively. Because of the proximity of these two minima it is not possible to detect conformational averaging on the basis of the NOE data. The very sharp additional minimum observed in the CHARMm grid search may be a consequence of a highly restricted conformer brought about through the presence of the two neighboring acetamido groups either side of the glycosidic linkage, although an artefact cannot be excluded.

The α-AATG-(1 → 4)-D-GalNAc[6PC] linkage.—It is much more difficult to assign an unequivocal conformation to this linkage because only one NOE-derived distance is available, that between H-1B and H-4C equal to 2.26 ± 0.25 Å. Exploration of Φ, ψ conformational space for this linkage using the MMX force-field identified five conformer populations with the following estimated percentage occupancies based on their calculated energies: −13°, +52° (58%); −40°, −39° (25%); +1°, +46° (12%); −50°, −32° (3%); −10°, +46° (2%). The weighted average for the distance H-1B → H-4C gives an estimated value of 2.33 ± 0.1 Å in excellent agreement with the experimentally determined value. A CHARMm grid search, performed over ±180° with a 10° increment, indicated the

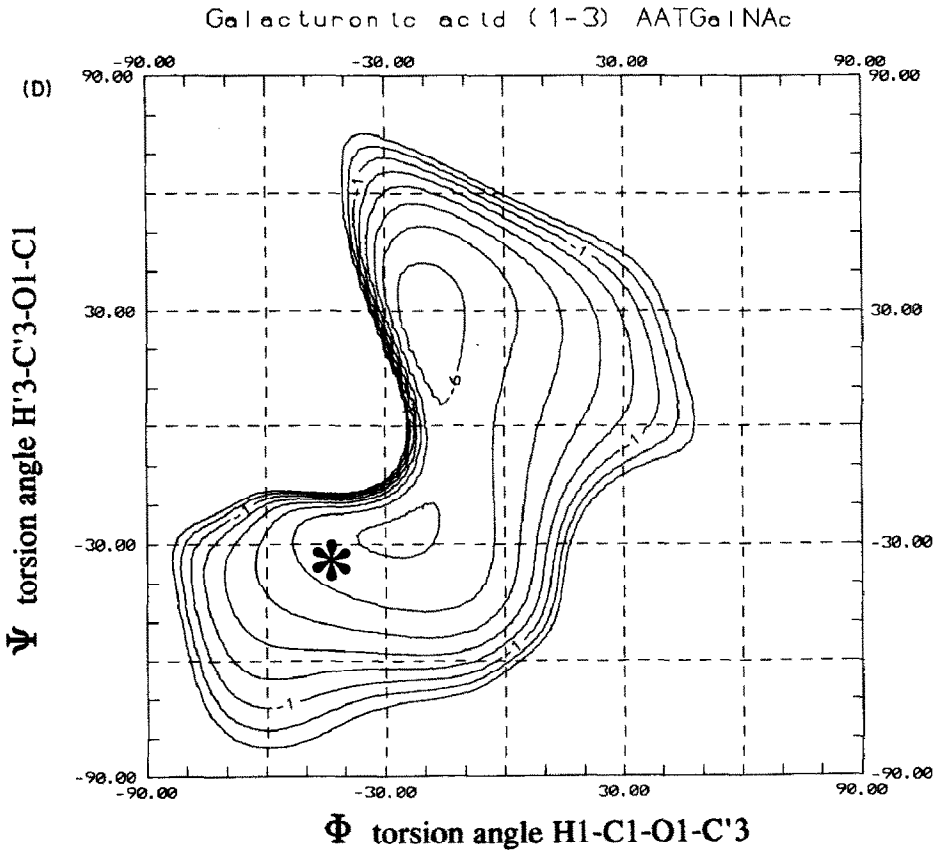


Fig. 4 (continued).

presence of two minima centred at −(20–40°), +(20–40°), and at −(5–15°), −(10–20°). Relaxed CHARMm calculations gave Φ and ψ values of −(25–30°), +(30–40°), and +5°, +32° broadly in line with the MMX calculations and showing similar energy differences between the conformer populations. The various Φ and ψ angles are shown superimposed for comparison on the contour map in Fig. 3.

The α-D-GalNAc-[6PC]-(1 → 3)-D-GalNAc-[6PC] linkage.—Sampling of the Φ, ψ-dependent energy surface, using the MMX force-field, for the α-(1 → 3) linkage between the two GalNAc residues carrying 6-linked phosphorylcholine groups, indicated a highly complex surface with multiple minima. We have identified six significant conformer populations for this linkage with occupancies estimated from the differences in total energy as follows: −38°, +59° (36%); −32°, −29° (32%); −18°, −11° (22%); −37°, −30° (9%); −16°, −33° (1%). The energy surface is undoubtedly made more complex than for a simple α-Gal-(1 → 3)-Gal linkage because each galactopyranosyl ring carries, in addition to the 2-acetamido group, a highly mobile, doubly charged zwitterionic phosphorylcholine group phosphodiester linked at the 6-position, thus increasing greatly the possibilities for stable conformers. Boltzmann averaging of these conformer populations derived from the MMX calculations, gives an estimated distance of 2.25 ± 0.1 and 2.71 ± 0.1 Å for the H-1C → H-3D and H-1C → H-4D proton pairs, respectively. Comparison of these values with those derived experimentally from the NOE measurements (2.33 and 2.35 Å) suggests that the contribution of the conformer (−, +) is overestimated.

Grid searching of conformational space, in 5° increments using the CHARMm force-field, gave iso-energy contour maps for the disaccharide unit and the intact repeat tetrasaccharide–ribitol as shown in Figs. 4A and 4B. Two broad minima are apparent which will be referred to as (−, +) and (−, −) according to the sign of their Φ and ψ angles. The grid search for the disaccharide in isolation shows that the (−, +) conformer is energetically the more favourable by some kcal/mol, whereas for the tetrasaccharide–ribitol the situation is reversed with the (−, −) minimum becoming the more likely. These results for the intact repeat-unit, with > 95% corresponding to the (−, −) conformer, suggest that this contour surface is in reasonable agreement with the experimentally determined NOE-derived distances, in contrast to the MMX calculations which overestimated the importance of the (−, +) conformer.

The energy surface for this glycosidic linkage involving a galactopyranosyl residue α-(1 → 3)-linked to a 2,4-diacetamido-2,4,6-trideoxygalactopyranosyl residue, is characterized by two minima (−, −) and (−, +), and is very similar to that found for α-galacturonic acid-(1 → 3)-[2-a cetamido-4-amino]-trideoxygalactose reported as the repeat-unit in another *S. pneumoniae* oligosaccharide by Behr et al. [7]. These authors report a single minimum as calculated using the HSEA method, at −42°, −32°, apparently consistent with their experimentally determined NOE distances. We have found, however, that minimization of the α-GalA-(1 → 3)-AATG linkage, also using an HSEA method [38,39], gave two minima as shown for comparison with the data of Bock et al. [46] in Figs. 4C and 4D. The

relative energy of the two minima depends on whether an explicit exoanomeric effect energy term is included (C), or whether the Burckert torsional term is used (D). The global minimum obtained using the explicit exoanomeric energy term (C) appears at first sight to correspond exactly with the experimental NOE-derived interproton distances of 2.45 Å for both H-1 → H-3′ and H-1 → H-4′ (Ref. [46]). It should be pointed out, however, that since the energy surface is rather flat, depending to some extent on the energy terms included, the distance-averaged NOE effects must also include substantial contributions from other parts of the energy surface. If an NOE calculation is performed [47,48] by sampling the entire energy surface, values for the H-1 → H-3′ and H-1 → H-4′ distances are obtained, 2.59 and 2.40 Å, respectively, which are still in reasonable agreement with the experimental NOE-derived distances, at least within experimental error. Since glycosidic linkages often exhibit rather flat minima, it seems preferable to use matrix sampling of the energy surface rather than relying on the position of the minima themselves, in order to estimate accurately r^{-6}-averaged interproton distances.

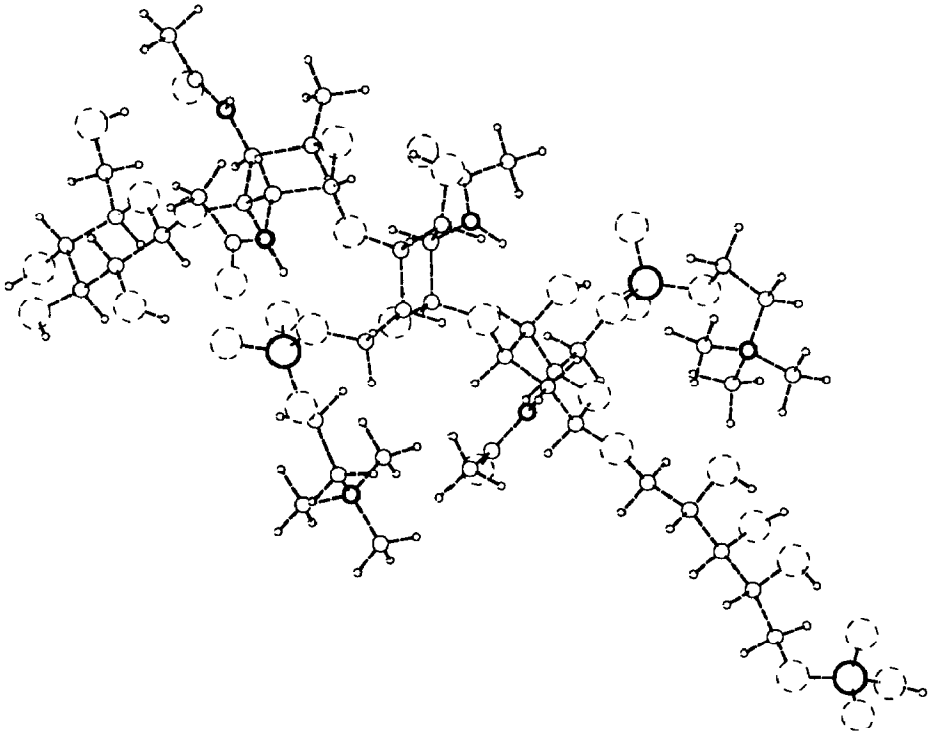


Fig. 5. The three-dimensional structure of the tetrasaccharide–ribitol repeat-unit with NOE-derived distance constraints applied during minimization. The NOE effects were measured in the repeat-unit and in the intact oligomer consisting of 5–7 repeat units (average mol wt 7000–7500). Oxygen atoms are shown in the form of dotted circles and phosphorus atoms as large circles in heavy type.

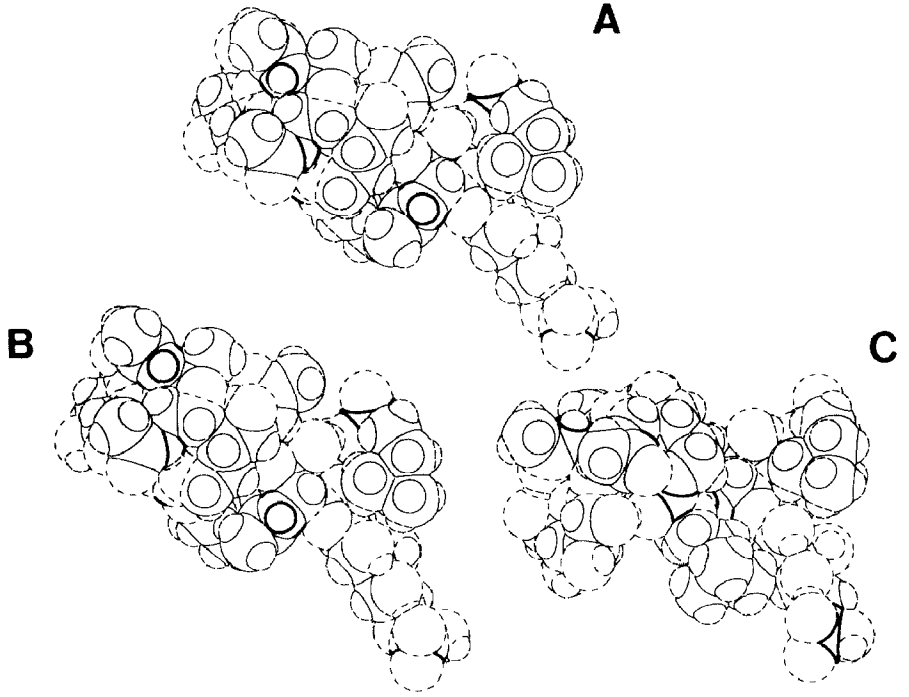


Fig. 6. The three-dimensional structures are shown of (A) the tetrasaccharide–ribitol repeat-unit with NOE-derived distance constraints applied, as well as (B) and (C) the two conformers corresponding to the major minima in the contour map for the α-GalNAc[6PC]-(1 → 3)-GalNAc[6PC] linkage shown in Fig. 4. Ball-and-stick representations are used. The torsional angles for the structures shown are (A) +35°, −48°; −21°, +34°; −15°, −39°; +62°, +60°, −57°: (B) +33°, −31°; −22°, +31°; −8°, −34°; +62°, +59°, −57°: (C) +33°, −31°; −23°, +31°; −37°, +61°; +63°, +62°, −56°. The Φ and ψ angles are given for linkages in the order of the units A → B, B → C, C → D, and D → E.

It is clear from the contour maps shown in Figs. 4A–4D that all three methods, namely CHARMm, MMX, and HSEA calculations, give very satisfactory agreement as to the shape of the energy surface and that this linkage is not especially environmentally sensitive to substituents on the two galactopyranosyl-derived rings – see later discussion of the variability of the Φ and ψ angles during Monte Carlo simulation.

The β-D-GalNAc-[6PC]-(1 → 1)-Ribitol-5-P linkage.—The beta linkage between the terminal GalNaAc-[6PC] residue and C-1 of the ribitol phosphate shows a single, well-defined minimum at +50°, +60°(−60°) corresponding to a g^+ and g^- orientation of the H-1a and H-1b protons of the ribitol chain, respectively, in relationship to the anomeric C-1–O-1 bond of the galactosamine ring. The ribitol chain shows, as expected, considerable conformational mobility.

Conformation of the tetrasaccharide–ribitol repeat-unit.—The average conformation predicted on the basis of the NOE measurements is shown in Fig. 5. In Figs. 6B and 6C the two most probable conformer populations, identified on the basis of

searching conformational space for all the linkages, are shown with the NOE-derived structure for comparison in Fig. 6A.

The conformation of the 5-oligomer of the repeat tetrasaccharide–ribitol.—We have sampled conformational space, using the Boltzmann jump procedure, for the 5-oligomer of the tetrasaccharide–ribitol consisting of repeat-units 5,6′-phosphate diester linked through their terminal D-glucopyranose to the ribitol of the previous repeat-unit, utilizing the NOE-predicted average structure as a reasonable starting geometry. This procedure is an implementation of the Metropolis Monte Carlo algorithm [49] to explore conformational space for energy minima. Random perturbations are carried out in torsional space whereas quenching is achieved in Cartesian space. If the perturbed conformation, occurring within a predefined torsion angle window, is of higher energy then a stochastic decision is made to retain or reject the new conformation, based upon the Boltzmann probability factor, $\exp(-E/kT)$. This method is capable of making an upward jump, or moving uphill in an energetic sense, and has the added advantage that it is less computationally intensive than a standard grid search with many degrees of freedom.

A total of 242 conformations were examined using the Boltzmann jump procedure, taking a total of 3 days CPU time. All nonring torsions were defined as valid for perturbation. We used a torsion angle window of 5° and an RMS torsion space of 2° as the cutoff for thermal perturbations; each minimization was performed using the CJ method with 300 steps. A high value of 5000 K for the apparent "temperature" was used in the formula which determines whether a conformation was accepted or rejected, so as to increase the chances of making an energetically uphill jump. We have found that the CJ method of minimization is more stable than the ABNR method as regards its ability to cope with bad starting geometries encountered during searches of conformational space, especially for molecules containing permanently charged groups such as phosphates or carboxylates, as in the present case.

The results of the Boltzmann jump procedure are shown in Fig. 7a. This figure shows the conformation number, which is arbitrary but sequential, plotted against the CJ-minimized conformation energy. The most stable conformers which were identified, have energies between −57 and −59 kcal/mol (conformers 175–205); another stable conformer occurs at −53 kcal/mol. Ball-and-stick representations of these two conformer populations, A and B, respectively, are shown in Fig. 6. In Table 5 the torsional angles are shown for the glycosidic linkages in structures A and B, as well as for a conformer from the higher energy plateau region C. The tetrasaccharide–ribitol repeat-units, five in all, are indicated by the Greek letters; individual glycosidic linkages are defined in the legend to the table. The torsional angles are sequentially numbered in pairs starting from the first intersugar linkage of the nonreducing end repeat-unit; they are, therefore, in pairs in sets of four for each repeat-unit, making forty glycosidic torsional angles in all.

Many of the torsional angles show only random variation with conformation number, and thus overall energy, amounting to ±5° and often as little as ±2°; an example of this is shown for the Φ and ψ angles of the first glycosidic linkage

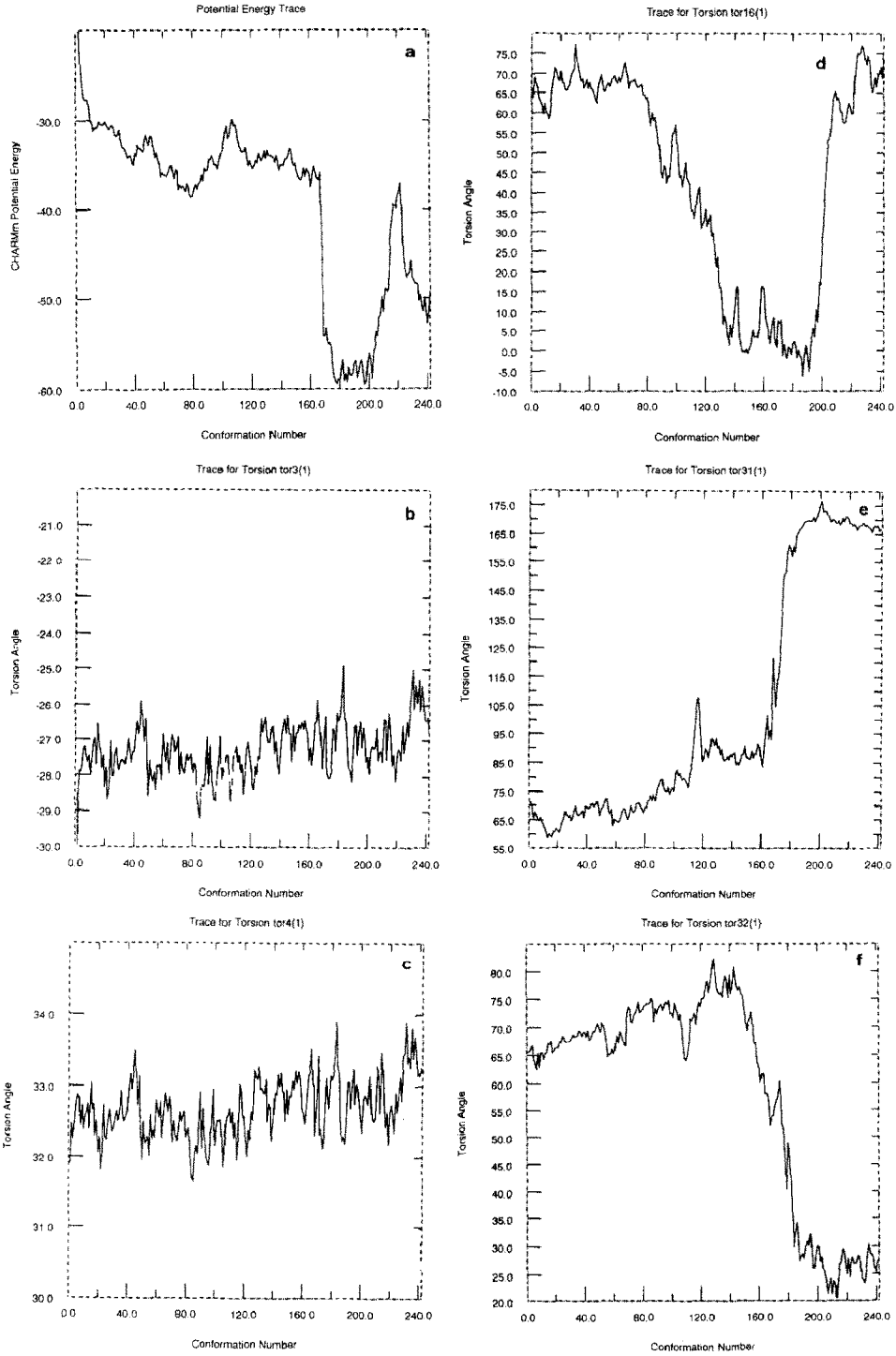

Potential Energy Trace
a
CHARMm Potential Energy
-30.0
-40.0
-50.0
-60.0
0.0
40.0
80.0
120.0
160.0
200.0
240.0
Conformation Number
Trace for Torsion tor16(1)
d
Torsion Angle
75.0
70.0
65.0
60.0
55.0
50.0
45.0
40.0
35.0
30.0
25.0
20.0
15.0
10.0
5.0
0.0
-5.0
-10.0
Conformation Number
Trace for Torsion tor3(1)
b
Torsion Angle
-21.0
-22.0
-23.0
-24.0
-25.0
-26.0
-27.0
-28.0
-29.0
-30.0
Conformation Number
Trace for Torsion tor31(1)
e
Torsion Angle
175.0
165.0
155.0
145.0
135.0
125.0
115.0
105.0
95.0
85.0
75.0
65.0
55.0
Conformation Number
Trace for Torsion tor4(1)
c
Torsion Angle
34.0
33.0
32.0
31.0
30.0
Conformation Number
Trace for Torsion tor32(1)
f
Torsion Angle
80.0
75.0
70.0
65.0
60.0
55.0
50.0
45.0
40.0
35.0
30.0
25.0
20.0
Conformation Number

Table 5
Torsion angles for the glycosidic linkages in the 5-oligomer. Values for the structures identified as (A), (B), and (C) on the energy trace of the Boltzmann jump procedure (Fig. 7) are listed, with energies of −59.8, −52.8, and −38.8 kcal/mol, respectively. The linkages shown are as follows: Link 1 = Ω (O′-1–C-6–C-5–O-5) for the D-glucose unit; Link 2, Φ and ψ angles for β-Glc-(1 → 3)-AATG; Link 3, for α-AATG-(1 → 4)-GalNAc[6PC]; Link 4, for α-GalNAc[6PC]-(1 → 3)-GalNAc[6PC]; and Link 5, for β-GalNAc[6PC]-(1 → 1)-RibP, respectively. The angles Φ and ψ are defined in the text

Unit		Link 1	Link 2	Link 3	Link 4	Link 5
Alpha	A	−50°	+35°, −30°	−26°, +33°	−31°, −35°	+67°, +63°
	B	−50°	+35°, −30°	−26°, +33°	−34°, −35°	+59°, +62°
	C	−50°	+36°, −30°	−28°, +32°	−34°, −36°	+61°, +61°
Beta	A	−70°	+27°, −22°	−27°, +32°	−30°, −35°	+70°, +7°
	B	−75°	+32°, −18°	−29°, +31°	−30°, −36°	+69°, +72°
	C	−78°	+27°, −37°	−25°, +34°	−32°, −36°	+85°, +63°
Gamma	A	−75°	+19°, −44°	−20°, +37°	−31°, −35°	+48°, +41°
	B	−61°	+19°, −34°	−20°, +38°	−30°, −33°	+68°, +77°
	C	−76°	+23°, −39°	−24°, +35°	−33°, −37°	+66°, +58°
Delta	A	−68°	+14°, −13°	−27°, +32°	−17°, −33°	**+171°, +26°**
	B	−68°	+5°, −30°	**+13°**, +29°	−23°, −32°	**+166°, +25°**
	C	−75°	+15°, −39°	−18°, +38°	−32°, −33°	+69°, +73°
Theta	A	−59°	+42°, **+14°**	−28°, +32°	−31°, −36°	+65°, +61°
	B	−58°	+47°, **+15°**	−29°, +31°	−32°, −35°	+63°, +61°
	C	−51°	+30°, −21°	−28°, +32°	−34°, −35°	+64°, +60°

between the AATG and GalNAc[6PC] residues (tor3 and tor4) in Fig. 7. Some torsional angles, however, especially those between the terminal GalNAc[6PC] and the ribitol residues in the penultimate repeat-unit (delta), show much greater variation associated clearly with the energy minimum in Fig. 7a, as seen in Figs. 7d–7f – tor16, tor31, tor32. The torsional angles within the delta repeat-unit show, in general, much greater variation with conformation number than those of the other units. On the other hand, Link 4 between the two α-(1 → 3)-linked GalNAc[6PC] residues is rather more conformationally stable than the others.

A similar conclusion can be drawn from plotting the results for the Monte Carlo simulation (Boltzmann jump) as a scatter plot of the Φ and ψ angles. In Figs. 8a and 8b scatter plots for tor1/tor2 and for tor7/tor8 are shown; these torsional angles refer to the glycosidic linkages between the terminal glucose and AATG, and between the choline-substituted galactosamine linked to the ribitol (see description of torsional angle numbering in Materials and methods). The conclu-

Fig. 7. Results of the Boltzmann jump procedure (Metropolis Monte Carlo simulation) for the 5-oligomer, showing (a), the total CHARMm energy versus conformation number equivalent to time of simulation; (b) and (c), the Φ and ψ torsion angles (tor3 and tor4) for the first α-AATG-(1 → 4)-GalNAc[6PC] linkage; (d), the ψ angle (tor16) for the second β-GalNAc[6PC]-(1 → 1)-RibP linkage; (e) and (f), the Φ and ψ angles (tor31 and tor32) for the fourth β-GalNAc[6PC]-(1 → 1)-RibP linkage; all corresponding to the same conformation numbers.

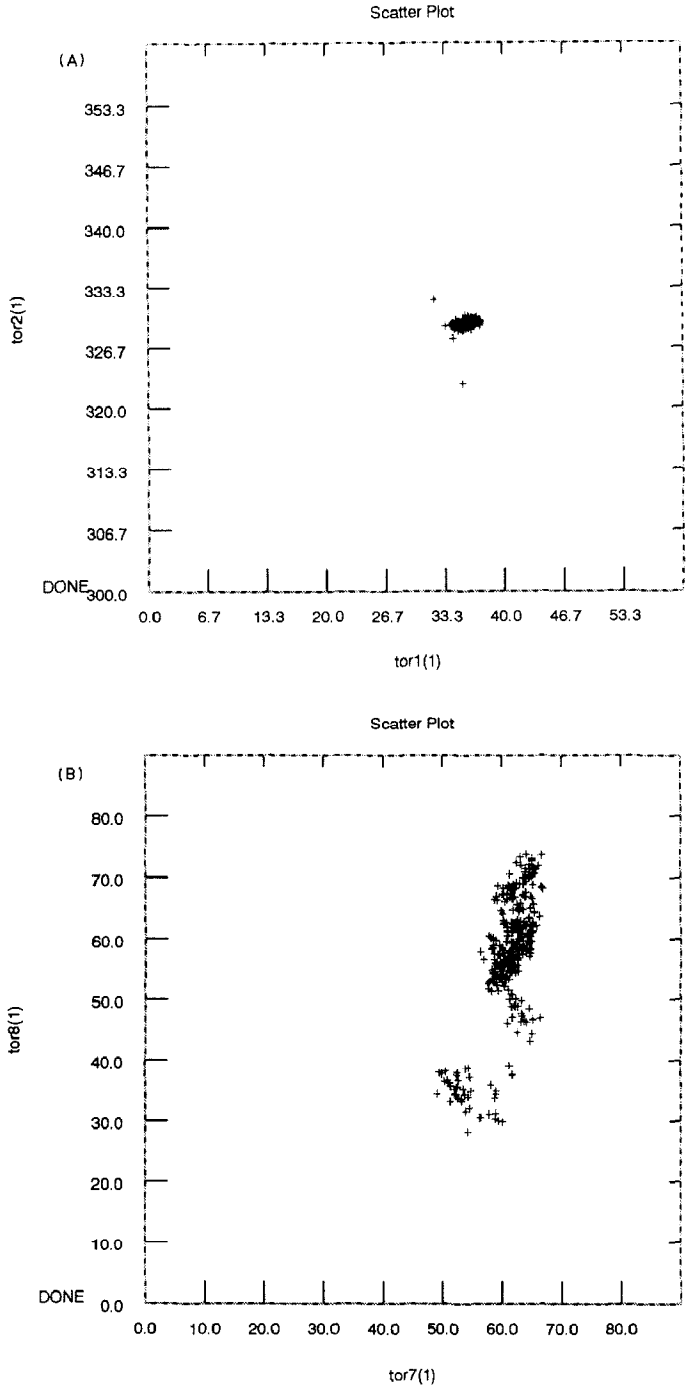


Fig. 8. Scatter plots of the torsional angles tor1/tor2 (A) and tor7/tor8 (B) generated by Monte Carlo simulation for the 5-oligomer, as described in the text.

sion to be drawn from both the MD and MC simulations is that the glycosidic linkages between the sugar units are relatively stiff, probably due to substituent steric effects, but that the linkage β-GalNAc-(1 → 1)-Ribitol is rather mobile in comparison. Additional calculations have shown, surprisingly, that the ribitol chain itself is not very mobile.

All of the above conformational results were obtained with the per-*N*-acetylated material. This was originally done so as to improve spectral resolution. A positive

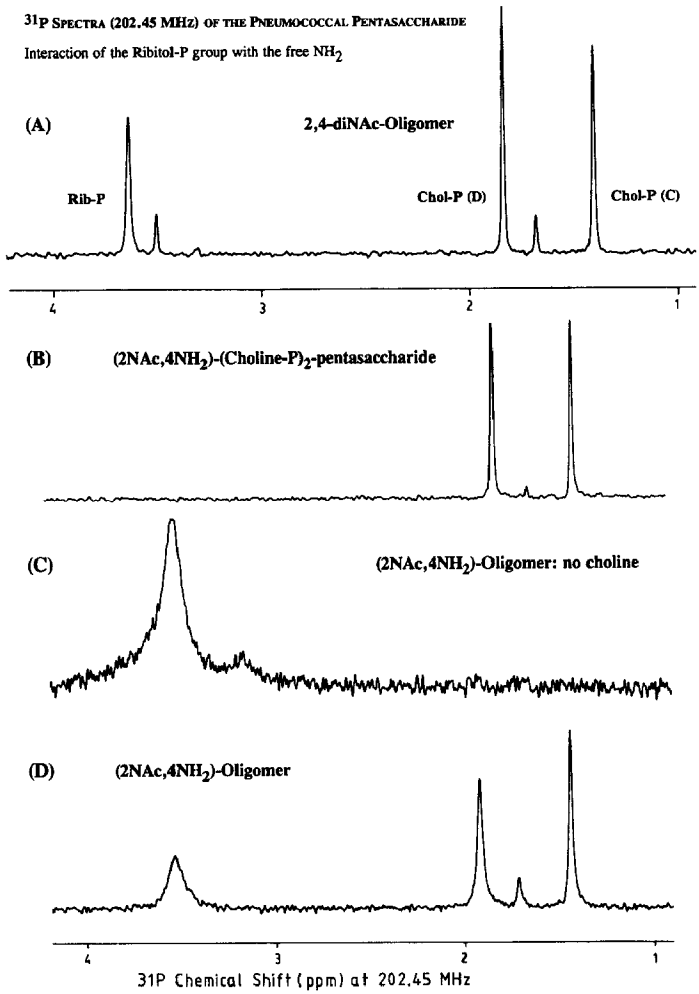
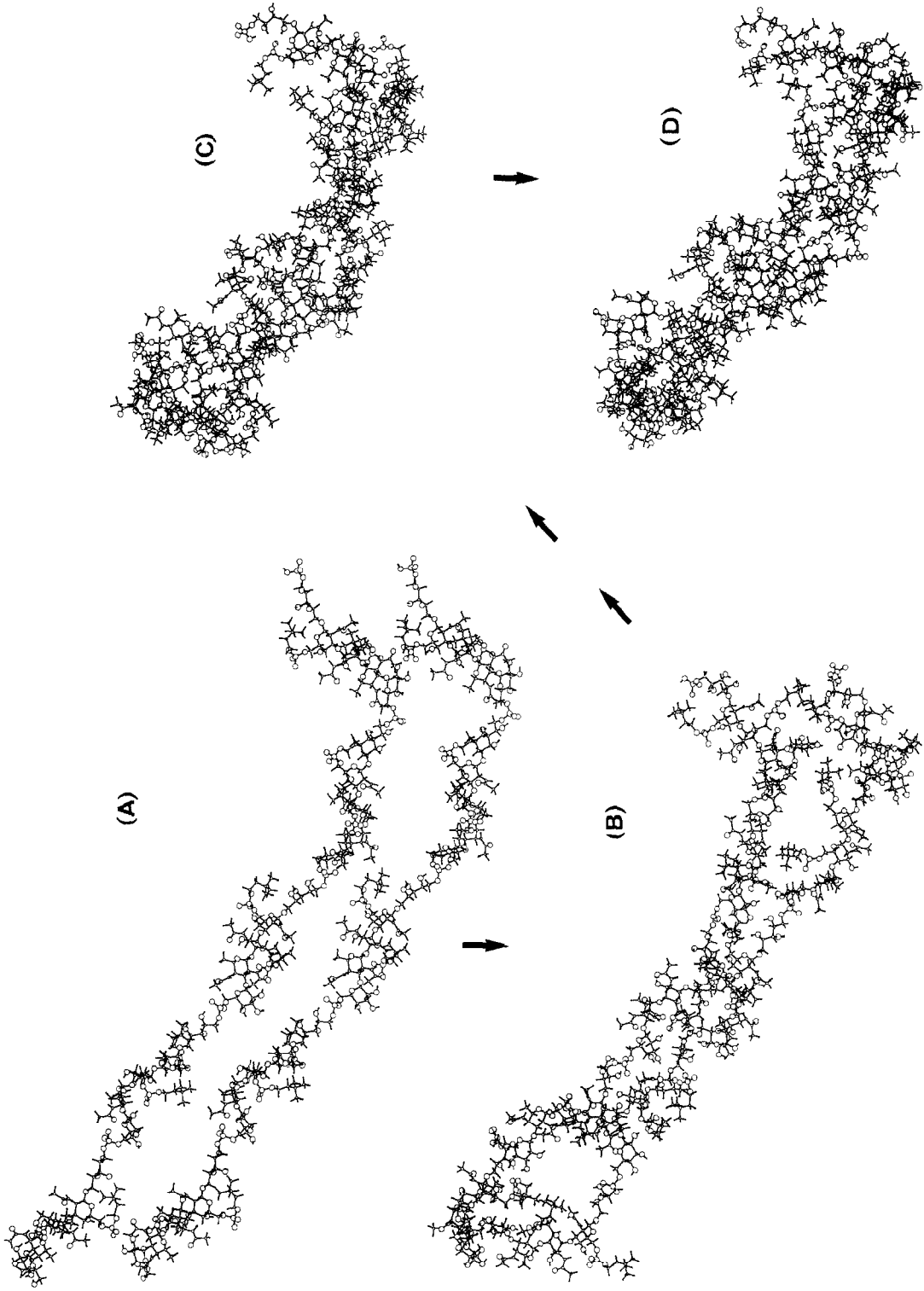


Fig. 9. ^{31}P spectra for various derivatives of the native oligosaccharide: (A) the per-*N*-acetylated oligomer; (B) the tetrasaccharide–ribitol repeat-unit with a free 4-NH_2 group on the AATG residue; (C) the native oligomer without choline phosphate side chains; (D) the native oligomer with free 4-NH_2 groups on the AATG residues. The small peaks marked with an asterisk are those originating from the terminal repeat units.


(C)
(D)
(A)
(B)

side-effect, however, was to uncover a very interesting effect on the ^{31}P resonances.

In the oligosaccharide containing a free 4-NH_2 group on the AATG residue, the ^{31}P resonances are broadened, not only for the ribitol-P but also for the phosphorylcholine-P, *only* in those preparations where there is both a ribitol phosphate group and a free amino group present (Figs. 9C and 9D). The free NH_2 group alone is insufficient to cause broadening of the ^{31}P signals (Fig. 9B).

The electrostatic interaction between the positively charged AATG residues in the nonper-*N*-acetylated material, and the intrachain phosphorylcholine groups, which is presumably the basis of the line broadening (Fig. 9), was also evident in the rate of periodate oxidation of the chain which required 120 h for completion rather than 4 h as observed for the per-*N*-acetylated derivative [7].

Apart from the ^{31}P resonance broadening seen in the spectrum of the native material with free 4-NH_2 groups, we have observed other NMR evidence which suggests that there are conformational differences compared to the per-*N*-acetylated derivative, particularly in the region of the glycosidic linkage between the glucopyranose residue and the 2-acetamido-4-amino-2,4,6-trideoxy-galactopyranose residue, as might be expected. It is particularly interesting that we have observed evidence for a change in the conformation of the galactopyranose ring itself, presumably brought about by the electrostatic interaction of the free amino group at the 4-position of the galactopyranose with the phosphate group on the 6-position of the glucopyranose. These conformational studies of the native material will be reported in a subsequent paper. Because of the evidence for differences in conformation between the native and *N*-acetylated material, care should be taken in extrapolating from the data presented in this paper to the biological situation.

Complex formation between adjacent oligosaccharide chains is driven by physical forces.—Given the location of oligomers of the tetrasaccharide–ribitol repeat-unit on the streptococcal cell surface, we have modelled the situation to investigate whether complex formation between adjacent oligomers is possible. Two energy-minimized 5-oligomers were aligned so that the chains were within ~5 Å of one another and then subjected to a short CJ minimization (300 steps) in order to remove bad contacts, before performing a molecular dynamics calculation: time step 0.001 ps; heating, 0 to 600 K, 1.08 ps; equilibration at 600 K, 2 ps; simulation at 600 K, 60 ps. This was followed by an annealing phase from 600 K to room temperature.

Fig. 10 illustrates the moving together of the two 5-oligomers as the MD simulation and annealing progressed. Examination of the final copolymer by rotating the structure on the screen, showed that it was stabilized by hydrogen

Fig. 10. The copolymerized complex of two 5-oligomer units is shown after different stages in the molecular dynamics simulation: (A) starting conformation; (B) after 10 ps simulation at 600 K; (C) after a further 50 ps simulation; and (D) after cooling (simulated annealing) from 600 K to absolute zero. The time sequence is indicated by the arrows. All four molecular structures are drawn to the same scale (originally 2 mm per Å, subsequently scaled down).

bonds, 157 in total, mainly between saccharide ring hydroxyl groups, with the beginning of a helical structure becoming visible. Unfortunately it was not possible for us to detect complex formation experimentally using NMR spectroscopy because the solutions were too dilute and we did not have sufficient additional material available.

Although the formation of complexes through Van der Waals and electrostatic interactions at low dielectric constant close to in vacuo values is a physical truism and the movement of chains together would be expected on physical grounds alone, it may, nonetheless, have considerable importance biologically. The cell surface close to the interfacial region between the hydrophobic core and the highly polarized, structured water layer forming part of the Stern double-layer, has a relatively low dielectric constant, at any rate much less than the value for bulk water ($\epsilon = 80$). Physical driving forces for complex formation in this region and the consequent association between oligosaccharide chains on the cell surface could be

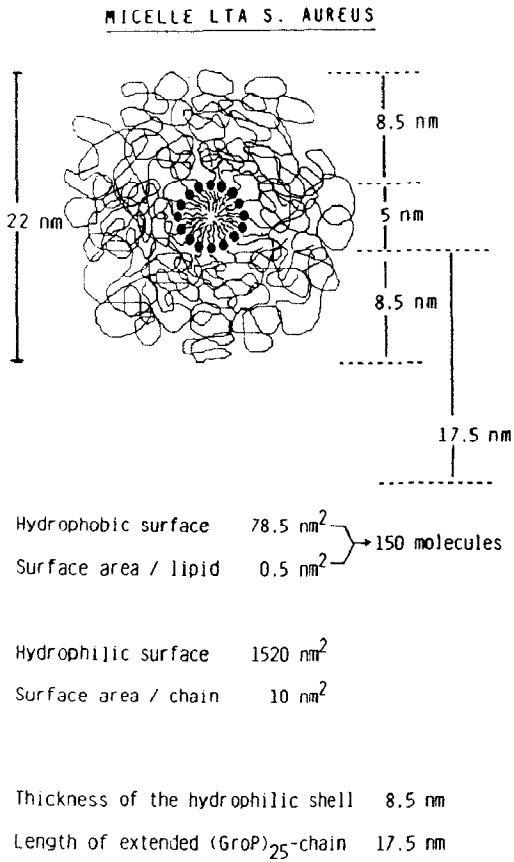


Fig. 11. Representation of the micelles formed from the lipoteichoic acid (LTA) of *Staphylococcus aureus* in aqueous dispersions (from Ref. [50]).

involved in protective or receptor functions; association in free solution, if it occurred, would have significance for aggregation phenomena in immunology.

Complex formation might be detected experimentally using micelles of the native lipoteichoic acid. These form spontaneously in aqueous dispersions and contain polymer chains at an appropriate distance from one another [50], lying parallel rather than anti-parallel and in a medium of presumed low dielectric constant and limited water molecule availability – i.e., not in free aqueous solution. Unfortunately this supramolecular structure formed by the micelles is unlikely to be directly amenable to NMR analysis because of anisotropy effects and line broadening. A representation of the micelles formed from *S. aureus* lipoteichoic acid is shown in Fig. 11 taken from Labischinski et al. [50].

Pneumococcal teichoic acid has the capacity to activate the indirect pathway of the complement cascade whereas pneumococcal lipoteichoic acid has not [51]. It can, however, acquire this capacity through binding to the surface of erythrocytes [20]; this binding occurs through insertion of the acyl residues into the lipid bilayer and predictably results in an oriented presentation of single chains. The inactivity in aqueous dispersions may therefore be caused by the micellar organization of lipoteichoic acid [50] in which required structural elements become inaccessible and complex formation between adjacent polymer chains may play a role.

4. Acknowledgments

We should like to acknowledge generous financial support from the Deutsche Forschungsgemeinschaft (DFG) in the form of a programme grant (SFB 284 A1) and (Fi 218/4-8).

5. References

[1] W.F. Goebel and M.H. Adams, *J. Exp. Med.*, 77 (1943) 435–449.
[2] D.E. Brundish and J. Baddiley, *Biochem. J.*, 110 (1968) 573–581.
[3] E.B. Briles and A. Tomasz, *J. Biol. Chem.*, 248 (1973) 6394–6397.
[4] M.J. Watson and J. Baddiley, *Biochem. J.*, 137 (1974) 399–404.
[5] I.R. Poxton, E. Tarelli, and J. Baddiley, *Biochem. J.*, 175 (1978) 1033–1042.
[6] B. Lindberg, B. Lindqvist, J. Lönngren, and D.A. Powell, *Carbohydr. Res.*, 78 (1980) 111–117.
[7] T. Behr, W. Fischer, J. Peter-Katalinic, and H. Egge, *Eur. J. Biochem.*, 207 (1992) 1063–1075.
[8] W.F. Goebel, T. Shedlovsky, G.I. Lavin, and M.H. Adams, *J. Biol. Chem.*, 148 (1943) 1–15.
[9] U.B.S. Sörensen and J. Henrichsen, J. Clin. Microbiol., 25 (1987) 1854–1859.
[10] U.B.S. Sörensen, J. Blom, A. Birch-Andersen, and J. Henrichsen, *Infect. Immun.*, 56 (1988) 1890–1896.
[11] B.A. Fraser and N.F. Mallette, *Immunochemistry*, 11 (1974) 581–593.
[12] B. Siddiqui and S. Hakomori, *J. Biol. Chem.*, 246 (1971) 5766–5769.
[13] S.J. Sung, W.J. Esselman, and C.C. Sweeley, *J. Biol. Chem.*, 248 (1973) 6528–6533.
[14] J.E. Coligan, B.A. Fraser, and T.J. Kindt, *J. Immunol.*, 118 (1977) 6–11.
[15] A. Slomiany and B.L. Slomiany, *Eur. J. Biochem.*, 76 (1977) 491–498.
[16] (a) J.V. Höltje and A. Tomasz, *Proc. Natl. Acad. Sci. USA*, 72 (1975) 1690–1694; (b) J.V. Höltje and A. Tomasz, *J. Bacteriol.*, 124 (1975) 1023–1024; (c) J.V. Höltje and A. Tomasz, *J. Biol. Chem.*, 250 (1975) 6072–6076.

[17] A. Tomasz, M. Westphal, E.B. Briles, and B. Fletcher, *J. Supramol. Struct.*, 3 (1975) 1–16.
[18] T. Briese and R. Hakenbeck, *Eur. J. Biochem.*, 146 (1985) 417–427.
[19] D.S. Hummell, A.J. Swift, A. Tomasz, and J.A. Winkelstein, *Infect. Immun.*, 47 (1985) 384–387.
[20] D.S. Hummell and J.A. Winkelstein, J. Clin. Invest., 77 (1986) 1533–1538.
[21] H.J. Jennings, C. Lugowski, and N.M. Young, *Biochemistry*, 19 (1980) 4712–4719.
[22] R.A. Byrd, W. Egan, M.F. Summers, and A. Bax, *Carbohydr. Res.*, 166 (1987) 47–58.
[23] C. Jones, F. Currie, and M.J. Forster, *Carbohydr. Res.*, 221 (1991) 95–121.
[24] E.W. Wooten, C.J. Edge, R. Bazza, R.A. Dwek, and T.W. Rademacher, *Carbohydr. Res.*, 203 (1990) 13–17.
[25] W. Fischer, T. Behr, R. Hartmann, J. Peter-Katalinic, and H. Egge, *Eur. J. Biochem.*, 215 (1993) 851–857.
[26] S. Macura and R.R. Ernst, *Mol. Phys.*, 41 (1980) 95–117.
[27] A. Kumar, G. Wagner, R.R. Ernst, and K. Wütrich, *J. Am. Chem. Soc.*, 103 (1981) 3654–3658.
[28] P.L. Durette and D. Horton, *Adv. Carbohydr. Biochem.*, 26 (1971) 49–125.
[29] C. Altona and C.A.G. Hasnoot, *Org. Magn. Reson.*, 13 (1980) 417–429.
[30] B. Sheldrick, *Acta Crystallogr. B*, 32 (1976) 1016–1020.
[31] F. Longchambon, D.A. Ohannessian, and A. Neuman, *Acta Crystallogr. B*, 31 (1975) 2623–2627.
[32] F.H. Allen, S.A. Bellard, M.D. Brice, B.A. Cartwright, A. Doubleday, H. Higgs, T. Hummelink, B.G. Hummelink-Peters, O. Kennard, W.D.S. Motherwell, J.R. Rodgers, and D.G. Watson, Acta Crystallogr. Sect. B, 35 (1979) 2331. Cambridge Structural Database (CSD) and subsequent versions.
[33] J.N. Scarsdale, J.H. Prestegard, and R.K. Yu, *Biochemistry*, 29 (1990) 9843–9855; J.N. Scarsdale, R.K. Yu, and J.H. Prestegard, J. Am. Chem. Soc., 108 (1986) 6778–6784.
[34] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus, *J. Comput. Chem.*, 4 (1983) 187–217; P. Cieplak and P. Kollman, J. Comput. Chem., 12 (1991) 1232–1236.
[35] N.L. Allinger, *J. Am. Chem. Soc.*, 99 (1977) 8127–8134.
[36] N.L. Allinger, R.A. Kok, and M.R. Iman, J. Comput. Chem., 9 (1988) 591–595.
[37] Serena Software (1990) Bloomington, Indiana. MMX - PC-Model implementation of an extended MM2 program for PCs and Silicon Graphics Workstations.
[38] H. Thøgersen, R.U. Lemieux, K. Bock, and B. Meyer, *Can. J. Chem.*, 60 (1981) 44–57; K. Bock, Pure Appl. Chem., 55 (1983) 605–622.
[39] H. Paulsen, T. Peters, V. Sinnwell, M. Hemme, and B. Meyer, *Carbohydr. Res.*, 156 (1986) 87–106.
[40] G.A. Jeffrey and R. Taylor, *J. Comput. Chem.*, 1 (1980) 99–109.
[41] G.A. Jeffrey, J.A. Pople, J.S. Binkley, and S. Vishveshwara, *J. Am. Chem. Soc.*, 100 (1978) 373–379.
[42] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes – the Art of Scientific Computing*, Cambridge University Press, Cambridge, UK, 1986.
[43] J. Gasteiger and M. Marsili, *Tetrahedron*, 36 (1980) 3219–3228.
[44] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, and M. Klein, *J. Chem. Phys.*, 79 (1983) 926–935.
[45] R.R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Clarendon Press, Oxford, UK, 1987.
[46] K. Bock, H. Lohn, and T. Peters, *Carbohydr. Res.*, 198 (1990) 375–380.
[47] J.H. Noggle and R.E. Schirmer, *The Nuclear Overhauser Effect*, Academic Press, New York, 1971.
[48] M.J. Forster, *J. Comput. Chem.*, 72 (1991) 292–300.
[49] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Telle, *J. Chem. Phys.*, 21 (1953) 1087.
[50] H. Labischinski, D. Naumann, and W. Fischer, *Eur. J. Biochem.*, 202 (1991) 1269–1274.
[51] J.A. Winkelstein and A. Tomasz, *J. Immunol.*, 120 (1978) 174–178.
[52] A. Nayeem, J. Vila, and H.A. Scheraga, *J. Comput. Chem.*, 12 (1991) 594–605.
[53] G.N. Ramachandran, C. Ramakrishnan, and V. Sasisekharan, *J. Mol. Biol.*, 7 (1963) 95–99.