ELSEVIER

Carbohydrate Research 265 (1994) 79–96

CARBOHYDRATE RESEARCH

NMR assignment and conformational analysis of the antigenic capsular polysaccharide from *Streptococcus pneumoniae* type 9N in aqueous solution

Trevor J. Rutherford [a,1], Christopher Jones [a,*], David B. Davies [b], A. Clare Elliott [c]

[a] *Laboratory for Molecular Structure, National Institute for Biological Standards and Control, South Mimms, Potters Bar, Hertfordshire, EN6 3QG, United Kingdom*

[b] *Chemistry Department, Birkbeck College, University of London, 29 Gordon Square, London, WC1H 0PP, United Kingdom*

[c] *ZENECA Agrochemicals, Jealott's Hill Research Station, Bracknell, Berkshire, RG12 6EY, United Kingdom*

Received 17 May 1994; accepted 11 June 1994

Abstract

Complete ^{1}H and ^{13}C NMR assignments, determined by one- and two-dimensional homo- and hetero-nuclear experiments, are reported for the antigenic capsular polysaccharide (CPS) from *Streptococcus pneumoniae* serotype 9N (S9 in American nomenclature). Distance constraints derived from 1D NOE difference experiments were combined with energy minimisation (simulated annealing) and molecular dynamics (MD) calculations to determine the most favoured conformation of S9 in aqueous solution at 70°C. NOE values simulated for several static conformational models using the NOEMOL program did not correlate well with experimental values, whereas time averaged interproton distances calculated from 500 ps of restrained MD (using the Tropp formalism to account for rapid internal mobility) were in close agreement with experimentally derived distance estimates.

Keywords: *Streptococcus pneumoniae*; Conformation; NMR spectroscopy; Molecular modelling; Capsular polysaccharide

* Corresponding author.

[1] Also at Birkbeck College, University of London. This work formed part of the Ph.D. thesis of T.J.R. (Birkbeck College, University of London, 1991). Current address: Department of Biochemistry, University of Dundee, Dundee, DD1 4HN, United Kingdom.

SSDI 0008-6215(94)00211-W

1. Introduction

The antigenic capsular polysaccharide (CPS) from *S. pneumoniae* serotype 9N (S9) is a regular repeating polymer with a pentasaccharide repeating unit [1] and is a component of all current pneumococcal vaccine formulations.

$$[\rightarrow 4)\text{-}\alpha\text{-Glc}p\text{NAc-}(1\rightarrow 4)\text{-}\alpha\text{-Glc}p\text{A-}(1\rightarrow 3)\text{-}\alpha\text{-Glc}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-Man}p\text{NAc-}(1\rightarrow 4)\text{-}\beta\text{-Glc}p\text{-}(1\rightarrow]_n$$

A B C D E

Residues are labelled A–E for convenient specification of atom numbering.

Serum antibodies raised against S9 are protective against other serogroup 9 strains (Danish types 9L, Ref. [2]; 9A, Ref. [3]; and 9V, Refs. [4,5]) with structurally related capsules [6], and antigenic cross-reactivity has also been observed between S9 and other polysaccharides possessing structural similarities (e.g., a common disaccharide in the repeating unit) [7–9]. An instructive approach is to compare the three-dimensional structures of the cross-reacting polysaccharides, and to map conserved and variable regions on to the three-dimensional structure.

The use of NMR spin-coupling constants and ^1H–^1H nuclear Overhauser enhancement (NOE) data to generate geometric constraints limiting the computed space in molecular modelling studies is well established for proteins [10] and oligosaccharides [11]. Little of this work has yet been extended to polysaccharide analysis and there are only a few heteropolysaccharides for which a solution conformation based upon experimental data has been reported [12–15]. Here we report the most energetically favourable conformations of S9 compatible with ^1H–^1H NOE and conformation-dependent chemical shifts at 70°C.

2. Experimental

A sample (30 mg) of S9 used for vaccine production, a gift from Merck, Sharpe, and Dohme (Rahway, NJ), was used without further purification. The molecular mass of the CPS was estimated to be $> 10^6$ Da using high-performance gel permeation chromatography [16]. Prior to NMR studies, the sample was partially depolymerised (dissolved in 1 mL 10 mM aq CF_3CO_2H, heated at 100°C for 10 min, neutralised with dil aq ammonia, dialysed, and lyophilised), to reduce the viscocity of the solution. The peak molecular weight of the partially hydrolysed sample was ~400 kDa (~2300 residues). Hydroxyl and amide protons were deuterium exchanged by repeated lyophilisation from D_2O, and the sample dissolved in D_2O (0.5 mL, nominally 100%) for analysis.

NMR spectroscopy.—500-MHz ^1H NMR spectra were acquired either on a Bruker AM-500 (COSY, relayed COSY, double-relayed COSY, TQF-COSY), a Jeol GSX-500 (NOESY, 1D-NOE, ^1H T_1), or a Varian Unity 500 spectrometer (HMBC). Spectral windows were chosen to include all of the signals of interest, generally 3.0–5.6 ppm. ^{13}C NMR spectra (67.5 MHz) were recorded on Jeol GSX-270 (^1H–^{13}C correlation) and Bruker AM400 (selective INEPT) spectrometers. All spectra were acquired at a nominal probe temperature of 70°C unless otherwise indicated.

NOESY spectra were acquired in the phase-sensitive mode using the complex transform procedure of States et al. [17]. One-dimensional truncated driven NOE experiments were performed using the Jeol microprogram NOEPUL, alternately acquiring on-resonance and off-resonance transients with a 180° transmitter phase shift between them. A recycle time of 5 s (ca. $5 \times T_1$) was allowed for relaxation between each pulse. The optimum preirradiation time (300 ms) under the conditions used was determined from a preliminary time-course experiment. Further time course experiments were carried out subsequently to estimate the time required to saturate the irradiated resonance under the conditions used (see text). The double-relayed COSY spectrum was optimised for interproton coupling constants of 10 Hz. The COLOC experiment was carried out according to Kessler et al. [18]. The FLOCK experiment was carried out as described by Reynolds et al. [19]. The selective INEPT spectrum [20] was obtained as described by Altman et al. [21] with magnetisation transfer optimised for J_{CH} 4 Hz, and omitting the refocussing pulse.

The HMBC spectrum [22] was obtained at 70°C on a Varian Unity 500 spectrometer. 1024 data points were collected in f_2 over a (^{1}H) spectral width of 3000 Hz, and 186 fids collected in f_1 covering a (^{13}C) spectral width of 8000 Hz. 128 transients were averaged in each fid. A phase-sensitive spectrum was obtained using the States method [17]. Chemical shifts were measured relative to internal 3-(trimethylsilyl)propionate-2,2,3,3-d_4 (sodium salt) at 0 ppm (^{1}H) or −1.8 ppm (^{13}C). Signals arising from amide protons were observed in 80:20 H_2O–D_2O using the $1\bar{3}3\bar{1}$ pulse sequence of Hore [23,24].

Molecular modelling.—Computer models were built using the Chem−X molecular graphics package (January 1988 version), developed and distributed by Chemical Design Ltd., Oxford, UK, running on a MicroVax II computer. Initial energy calculations for disaccharides used the Chem-X van der Waals force field and provided low energy starting structures for optimisation [25] with MM2 (QCPE, Indiana University), with Jeffrey–Taylor [26] parametrisation for use with carbohydrates (MM2CARB) [27]. The dielectric constant (ε) was set to 80 to part simulate bulk water. The initial monosaccharide model was constructed from MM2CARB optimised monosaccharides which have hydroxyl hydrogens in random, sterically unhindered orientations. Hydroxymethyl groups were set to the *gauche–gauche* rotamer ($\tau_{(O\text{-}6\text{–}C\text{-}6\text{–}C\text{-}5\text{–}H\text{-}5)} = 180°$).

Simulated annealing with NOE constraints was performed on a nonasaccharide model in vacuo (with a dielectric constant $\varepsilon = 80.0$ to part simulate bulk water) using the Discover molecular modelling package (version 2.8, Biosym Technologies, San Diego, CA), running on an IRIS INDIGO computer. The force field was AMBER [28] with the parameters described by Homans [29]. Initial velocities were assigned to a Maxwell–Boltzmann distribution at 300 K and were rescaled by coupling to a temperature bath with a time constant of 0.6 ps. Initially all torsion terms were scaled by a factor of 5 and 1 ps of dynamics was run at 500, 450, 400, and 350 K. The torsion terms were rescaled and the temperature was reduced in 10 K steps from 300 to 0 K, running 1 ps of dynamics at each temperature. The annealed structure was used as input to an MD simulation.

510 ps of MD was simulated at 300 K using the AMBER force field, with results recorded after every 250 steps of 1 fs. NOE constraints were applied throughout. The first 10 ps of the simulation was discarded before back-calculating NMR data. Tropp averaging of the MD simulations was carried out according to Homans and Forster [30]. NOE simulations were performed on rigid nonasaccharide models using the NEOMOL program [31] running

on a SUN 3/160 computer. A single correlation time of 5 ns was estimated from ^{13}C T_1 values, using an isotropic tumbling approximation.

3. Results and discussion

Sample homogeneity.—If S9 adopts multiple conformations in slow exchange, separate signals would be expected for each distinct conformation. Since only one multiplet was observed empirically for each resonance in the NMR spectra, no conformational exchange occurs that is slow on the chemical-shift-averaging timescale (ca. <1000 Hz). However, if part of the sample had greatly restricted mobility (e.g., due to partial aggregation) the resultant line-broadening may render that portion of the material NMR invisible. A D_2O solution was prepared with a weighed sample of S9 and two internal standards [3-(trimethylsilyl)propionate-2,2,3,3-d_4 (sodium salt) and NMe_4Cl] of known concentrations. Integrations of well-resolved multiplets in the ^1H NMR spectrum of S9 and those of the standards established that the S9 is 100% NMR visible. All conformational NMR data (NOE, chemical shifts, etc.), therefore, are representative of the entire sample, rather than a mobile subpopulation.

Spectral assignment.—The ^1H NMR spectrum of S9 (Fig. 1) has not previously been assigned, although that of a pentasaccharide derivative was reported [1]. The intact poly-

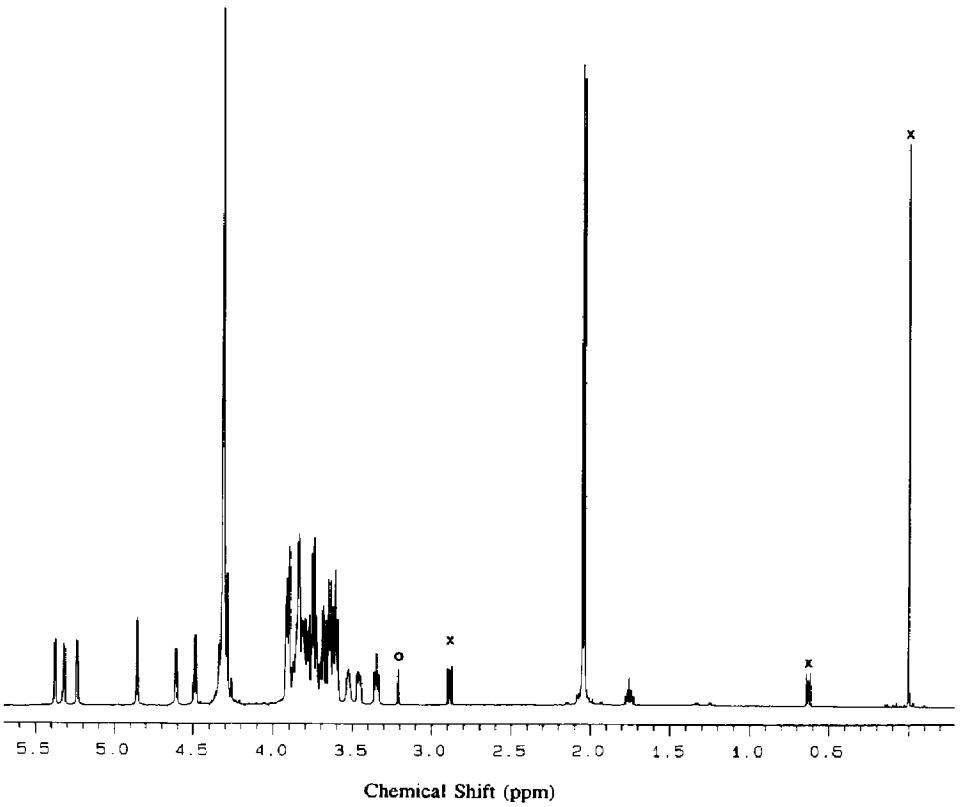

Fig. 1. ^1H NMR spectrum of the 9N PS, 600 MHz, 70°C.

Table 1
^{1}H chemical shift assignment for S9 at 70°C. Numbers in brackets are differences in chemical shift from model systems [a], where a downfield shift is positive. Linkage positions are underlined

Residue	H-1	H-2	H-3	H-4	H-5	H-6	H-6′
α-GlcNAc	5.38	3.91	3.83	3.68	3.91	3.81 [b]	3.76 [c]
	(+0.17)	(+0.03)	(+0.06)	(+0.17)	(+0.03)	(−0.04)	(−0.01)
α-GlcA	5.33	3.61	3.92	3.75	4.32		
	(+0.07)	(+0.02)	(+0.15)	(+0.20)	(+0.21)		
α-Glc	5.24	3.61	3.75	3.61	3.79	3.91 [c]	3.82 [c]
	(+0.01)	(+0.05)	(+0.01)	(+0.17)	(−0.07)	(+0.07)	(+0.06)
β-ManNAc	4.86	4.61	3.91	3.76	3.46	3.91	3.79
	(−0.17)	(+0.14)	(+0.06)	(+0.22)	(−0.01)	(−0.01)	(−0.04)
β-Glc	4.49	3.35	3.64	3.64	3.53	3.84	3.65
	(−0.17)	(+0.08)	(+0.12)	(+0.20)	(+0.05)	(−0.06)	(−0.07)

[a] P.-E. Jansson, L. Kenne, and G. Widmalm, *Carbohydr. Res.*, 188 (1989) 169–191. [b] Tentative assignment: resonance observed in NOE difference experiments when β-Glc H-1 is saturated. [c] Tentative assignment: correlation peak to C-5 observed in HMBC spectrum.

saccharide has 23 unresolved multiplets within an overlapping envelope between 3.5 and 4.0 ppm, 7 resolved signals at lower field, and 5 signals at higher field. The majority of ^{1}H resonances were assigned using analogous reasoning to that described previously for the types 9A and 9V CPSs [5], from COSY, and relayed COSY experiments, and the assignment is given in Table 1.

Resonances from hydroxymethyl protons of the α-GlcNAc and α-Glc residues (residues A and C) showed strong spin-couplings with H-5 and could not be assigned from homonuclear correlation experiments. Tentative assignment of one α-GlcNAc H-6 was made from an NOE difference experiment (see below) and for the remaining resonances from an HMBC experiment [22].

Assignment of ^{13}C NMR spectrum from the ^{13}C–^{1}H correlation spectrum (Fig. 2) was generally straightforward, but ^{13}C chemical shift arguments and comparison with model

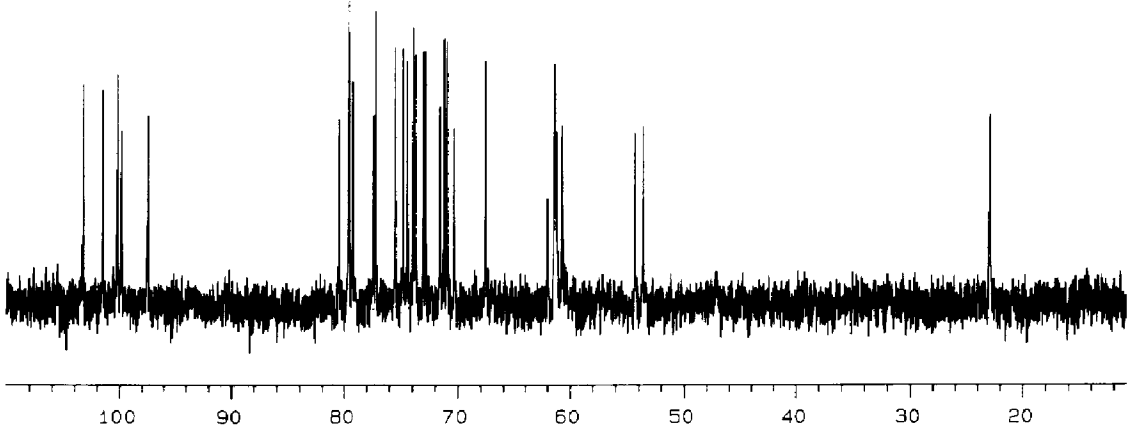

Fig. 2. ^{13}C NMR spectrum of the 9N PS, 67.5 MHz, 70°C.

Table 2
^{13}C NMR chemical shifts (glycosylation shifts) for S9 at 70°C. Linkage positions are underlined

Residue	C-1	C-2	C-3	C-4	C-5
α-GlcNAc	97.5 (+5.7)	54.3 (−0.7)	70.4 (−1.3)	79.7 (+8.4)	71.6 (−0.9)
α-GlcA	99.9 (+6.9)	71.3 (−1.0)	74.4 (+0.9)	77.4 (+4.5)	73.8 (+1.3)
α-Glc	101.6 (+8.6)	72.9 (+0.4)	80.0 (+7.0)	71.0 (+0.2)	73.1 (+0.7)
β-ManNAc	100.2 (+6.3)	53.7 (−1.2)	79.3 (+6.3)	67.6 (0.0)	77.2 (0.0)
β-Glc	103.3 (+6.5)	74.0 (−1.2)	74.9 (−1.9)	79.6 (+8.9)	75.5 (−1.2)

systems were used to distinguish within several sets of ^{13}C resonances where the ^{1}H resonances overlapped. These assignments were confirmed from an HMBC spectrum. The glycosylated positions of residues A and E resonate at identical ^{1}H and ^{13}C chemical shifts at 70°C, but the ^{13}C resonances could be resolved at lower temperature as only one was temperature-sensitive. Different temperature dependence of the glycosylation shifts may indicate different conformational behaviour at the D–E and E–A linkages, therefore it was necessary to assign the overlapping ^{13}C resonances. Two-dimensional long-range ^{13}C-detected C–H correlation experiments (COLOC [18] FLOCK [19]), were used, but were not sufficiently sensitive, probably because of fast T_2 relaxation. A selective INEPT spectrum [20,21] was acquired at 40°C with selective irradiation of H-1(D) and tuned for optimum magnetisation transfer when $J_{CH}=4$ Hz (typical of $^{3}J_{CH}$ across the glycosidic bond [32,33]), and gave an unambiguous assignment of C-4(E). The complete assignment of ring carbons is given in Table 2. ^{13}C resonances from methyl groups (23 ppm), hydroxymethyl groups (60.7, 61.1, 61.3 and 61.4 ppm), and carbonyl groups (174.3–175.7) were not individually assigned.

Nuclear Overhauser enhancement experiments.—The 500-MHz phase-sensitive NOESY of S9 at 70°C showed, as is typical for linear polysaccharides, few interresidue NOEs other than from H-1 to the protons attached to the glycosylated position of the aglycon (H-α). These interresidue NOEs confirmed the published sequence of residues. No 'long-range' enhancements were observed between residues not adjacent in the sequence. We chose to quantify these interresidue enhancements from one-dimensional truncated-driven difference experiments, using a 300-ms preirradiation time rather than measurement of cross-peak volumes in the NOESY experiment. Enhancement of peaks was measured as a percentage of the change in intensity of the selectively irradiated signal, and although the absolute integrals varied between 3 independent experiments, relative NOE values (i.e., $\sigma_{interresidue}/\sigma_{reference}$) (where σ, the cross relaxation rate, is proportional to the NOE) were reproducible to within ±5%. Each of the well resolved signals (including the *N*-acetyl singlets) was irradiated separately, but interresidue enhancements were only observed when irradiating H-1 resonances or H-2(D). The observed NOE, relative-NOE (average of three separate experiments) and interproton distances calculated from them, are given in Table 3.

Signal overlap was a hindrance to accurate quantification, eg. quantification of the H-1(A)···H-3(B) and H-1(E)···H-4(A) distances was hindered by overlap of H-3(B) with H-2(A) and H-3(E) with H-4(E) and H-4(A), respectively. Irradiation of H-2(D) enhanced the H-5(C) signal, which appeared as a broad multiplet and was difficult to

Table 3
^1H–^1H NOE values and calculated distance for S9 at 70°C

NOE from	Enhanced Peak		% NOE	Relative NOE	Interresidue ^1H–^1H Distance (nm)
	(ppm)	Assignment			
H-1(A)	3.91	H-2(A),H-5(A),H-3(B)	−16 [a]		
	3.83	H-3(A)	−1		
	3.75	H-4(B)	−5	2.98 ± .07	0.29
H-1(B)	3.61	H-2(B),H-2(C),H-4(C)	−15 [a]		
	3.92	H-3(B)	−2		
	3.75	H-3(C),H-4(B)	−8	1.94 ± .08	0.28
H-1(C)	3.61	H-2(C),H-4(C)	−11 [a]		
	3.75	H-3(C),H-4(D)	−2		
	3.91	H-3(D),H-6(D)	−12	1.06 ± .06	0.25
H-1(D)	4.86	H-2(D)	−5		
	3.91	H-3(D),H-6(D)	−10		
	3.46	H-5(D)	−7 [a]		
	3.64	H-3(E),H-4(E)	−17	0.37 ± .02	0.20
H-1(E)	3.35	H-2(E)	−1		
	3.64	H-3(E),H-4(E),H-4(A)	−12	0.91	0.25
	3.53	H-5(E)	−11 [a]		
	3.82	H-6(A),H-3(A)	−5	2.05 ± .39	0.31

[a] Reference NOE.

quantify. The intensity of the H-5(C) peak in the difference spectrum was approximately that of the H-3(D) and it was assumed that these protons were roughly equidistant from H-2(D).

Interresidue ^1H–^1H distances were estimated by the simple reference distance method, using the fixed distance between the H-1 and its closest intraresidue contact (for α-sugars H-1···H-2 = 0.248 nm, for β-ManNAc H-1···H-5 = 0.250 nm, and for β-Glc H-1···H-5 = 0.237 nm) as the reference. Wide error limits ($\pm$20%) were allowed on estimated distances, as this method does not allow for anisotropic reorientation, spin diffusion, integration errors or effects caused by rapid internal motion. Due to the r^{-6} dependence of NOE on the ^1H–^1H distance the limits chosen correspond to allowing an error of a factor of 3 in the observed NOE.

H-3(E) and H-4(A) resonated at the same chemical shift and each cross-relaxed with H-1(E). The H-1(E)···H-4(A) interproton separation was deduced to be 0.25 nm after correcting for the H-1(E)···H-3(E) NOE using an estimate based on the H-1(E)···H-5(E) NOE

Conformation-dependent chemical shifts.—Nonbonded interactions usually involve only substituents that are α and β to the linkage positions [34]. Negative glycosylation shifts have been measured for β-carbons in a range of disaccharides [34,35]; those not involved in steric interactions are generally ~0.5–1.5 ppm for carbons attached to an axial proton, and ~3 ppm when an equatorial proton is in close proximity to H-1′. Hence positive or large negative ^{13}C glycosylation shifts are indicative of interresidue associations [35].

Because many factors determine chemical shifts, glycosylation cannot be interpreted *a priori* in terms of specific geometric relationships, however, models were assessed by their

ability to explain conformation dependent shifts (Tables 1 and 2). The following glycosylation shifts (^{13}C more positive than -0.5 ppm or more negative than 1.5 ppm; $^{1}H > 0.1$ ppm in absolute value) of nonlinkage nuclei needed to be explained by satisfactory conformation models:

^{13}C, C-3(B) C-5(B) C-2(C) C-4(C) C-4(C) C-5(C) C-4(D) C-3(E)

^{1}H, H-3(B) H-5(B) H-4(C) H-2(D) H-4(D) H-3(E)

We also sought an explanation of the resistance to mild periodate oxidation of the 4-linked β-Glc residue (E) that was observed previously [1]. Glycosylation shifts of the E3 resonances indicate that periodate resistance may be due to steric hindrance, but it was recognised that α-GlcA (residue B) *is* periodate labile and its C-3 and H-3 resonances also had significant glycosylation shifts.

Coupling constant information: conformation of hydroxymethyl groups.—Large (9–10 Hz) vicinal ^{1}H spin-coupling constants for each proton pair H-2 through H-5 (except ManNAc H-2/H-3, 4 Hz), indicative of a diaxial configuration, confirmed that each ring adopts predominantly the expected $^{4}C_1$ chair conformation. The H-5/H-6 and H-5/H-6′ coupling constants for the two β-linked sugars were determined from the resolved H-5 multiplets in a 2D *J*-resolved spectrum (not shown). Hydroxymethyl rotamer populations, estimated from the H-5/H-6 coupling constants, using the approach of Nishida et al. [36], were as follows:

β-ManNAc 54% *gg* 39% *gt* 6% *tg*

β-Glc 61% *gg* 32% *gt* 7% *tg*

These values are in close agreement with previous calculations for β-Glc in aqueous solution, which show a preponderance of the *gg* rotamer (i.e., the rotamer with O-6 *gauche* to both O-5 and C-4) [36]. In residues A and C, H-5 resonances were not sufficiently resolved for spin-coupling constants to be estimated.

Hydroxymethyl rotamer populations in the molecular dynamics study were not consistent with those derived from H-5/H-6 spin-coupling constants, since a longer timescale must be simulated to study events as slow as interconversion of C-6 rotamers.

Coupling constant information: conformation of N*-acetyl groups.*—Two amide proton resonances were observed for the polysaccharide in 80:20 H_2O–D_2O at a range of temperatures using the $1\bar{3}3\bar{1}$ pulse sequence for water signal suppression. By comparison with a Karplus curve derived from peptide studies [37] the measured $^{3}J_{HCNH}$ coupling constants (9.8 and 10.2 Hz) [23,24] indicate H-2 and N–H are overwhelmingly anti-periplanar, consistent with X-ray crystal structures of acetamido sugars in the Cambridge Structural Database [38]. There was no NMR evidence to indicate the orientation of N–H with respect to C–O, but on the basis of crystal structures amide protons were set anti-periplanar to the carbonyl oxygen in all computer models.

Rapid spin–spin (T_2) relaxation in high-molecular-weight polymers decreases the sensitivity of experiments with extended pulse schemes, and we were not able to estimate glycosidic dihedral angles by comparing $^{3}J_{CH}$ with standard calibration curves [32,33].

Conformational energy calculations.—Methods for the computer modelling of carbohydrate conformations are well established and have recently been reviewed [27,39]. Theoretical analyses of di- and tri-saccharides suggest that several conformations about the glycosidic bonds are energetically accessible, and the NMR data represents an average of multiple conformations within the entire sample [40] (perhaps a ''virtual conformation''

Table 4
Constraints used to determine the glycosidic angles at each linkage, grouped according to type

Linkage	NOE constraints	^{13}C Chemical shift data ϕ ; ψ	Other constraints
GlcNAc-(α1→4)-GlcA	α-GlcNAc H-1···αGlcA H-4 (0.23–0.35 nm)	; −30	α-GlcNAc H-1···α-GlcA O-3 close
		δ_C C-1	δ_H H-1 is shifted downfield
	No other significant interresidue NOE observed		δ_C C-3 is temperature sensitive
GlcA-(α1→3)-Glc	α-GlcA H-1···α-Glc H-3 (0.22–0.34 nm)	; −20	α-GlcA H-1···α-Glc O-4 close
		δ_C C-1	δ_H H-1 is shifted downfield
	No other significant interresidue NOE observed		δ_C C-4 is temperature sensitive
			α-GlcA H-5···α-Glc O-2 close
			δ_H H-5 is shifted downfield
			δ_C C-2 is temperature sensitive
			δ_C C-5 is temperature sensitive
Glc-(α1→3)-ManNAc	α-Glc H-1···β-ManNAc H-3 (0.20–0.30 nm)	; −20	α-Glc H-1···β-ManNAc O-4
	β-ManNAc H-2···β-Glc H-5 (0.21–0.31 nm)	δ_C C-1	δ_c C-4 is temperature sensitive
	No other significant interresidue NOE observed		
ManNAc-(β1→4)-Glc	β-ManNAc H-1···β-Glc H-4 (0.18–0.24 nm)		β-ManNAc O-5···β-Glc OH-3 hydrogen bond as seen for cellobiose is postulated to explain periodate resistance of β-Glc residue
	β-ManNAc H-2···β-Glc H-4		
	No other significant interresidue NOE observed		
Glc-(β1→4)-GlcNAc	β-Glc H-1···α-GlcNAc H-4 (0.18–0.30 nm)		β-Glc O-5···α-GlcNAc OH-3 hydrogen bond postulated
	β-Glc H-1···α-GlcNAc H-6 (0.25–0.36 nm)		
	No other significant interresidue NOE observed		

as defined by Jardetzky [41]). Experimental evidence for multiple conformations is, however, difficult to obtain.

To assess the validity of energy-minimised conformational models NOE values were simulated, using the NOEMOL program [31] (*vide infra*), for comparison with experimental data. Refinement of the model against this experimental data reduces the dependence of the energy calculations on assumptions about the force field, the effects of hydration, and the contribution of factors such as the *exo*-anomeric effect. For the NOE simulations it was necessary to construct nonasaccharide models, with the CPS repeating unit as the central portion, and two residues on either end to account for 'end-effects'. Initially the simplest approach for exploring the potential energy surfaces at each link in the nonasaccharide was adopted, i.e., a rigid-residue investigation calculating only the van der Waals (VDW) energy. VDW energy surfaces were calculated for disaccharide models.

Potential surfaces (with 5°) increments calculated for disaccharide models of each element of the S9 repeating unit are shown in Fig. 3. In the models for these maps, pendant groups were set to experimentally determined preferred orientation; *N*-acetyl groups with N–H anti-periplanar to H-2 and C–O, and hydroxymethyl groups in the favoured *gg* orientation. Hydroxyl and amide protons were removed to avoid their involvement in bad contacts. The specific interresidue contacts that define the shape of the potential surface are labelled on the maps in Fig. 3.

Sequential minimisation of the glycosidic linkages in the nonasaccharide model was carried out by the same method, and the energy surfaces produced were almost identical, showing that there is little interaction between nonadjacent residues, consistent with the NOE results.

Disaccharide models (which included hydroxyl and amide protons) were placed successively in each of the local energy minima found on the energy surfaces and minimised using MM2CARB. NOEs were simulated for the nonasaccharide model with linkage geometries set to each of the MM2CARB-derived local minima, and compared with experimental data. The conformations in closest agreement with the NMR data are shown in Table 5.

The optimum MM2CARB model was used as the input to a simulated annealing study, using the AMBER force-field with the carbohydrate parameters of Homans [29] and the glycosidic dihedral angles are listed in Table 5. This structure is shown in Fig. 4. Simulated annealing is a convenient technique for optimising about all degrees of freedom simultaneously to produce low energy conformations. Using several starting structures explores a wide range of conformational space and the consensus conformation to which most structures converge is assumed to be the global minimum energy model. The high initial temperature and scaling of torsion terms is used to increase the range of the conformational search.

Comparison of simulated and experimental NOE data.—The NOEMOL program [31] simulates multi-spin NOE effects from the relaxation matrix for a given conformational (for estimating internuclear distances) and motional model (for estimating the spectral density function). In anisotropically tumbling molecules, such as stiff rods [31], the direct relationship between r^{-6} and the NOE is not valid, and terms involving the angle between the interproton vectors and the principal axes of rotation become important. ^{13}C T_1 values were measured to develop a motional model for S9; for an anisotropically tumbling molecule ^{13}C T_1 values are modulated by the angle σ between the C–H vector and the axes of rotation.

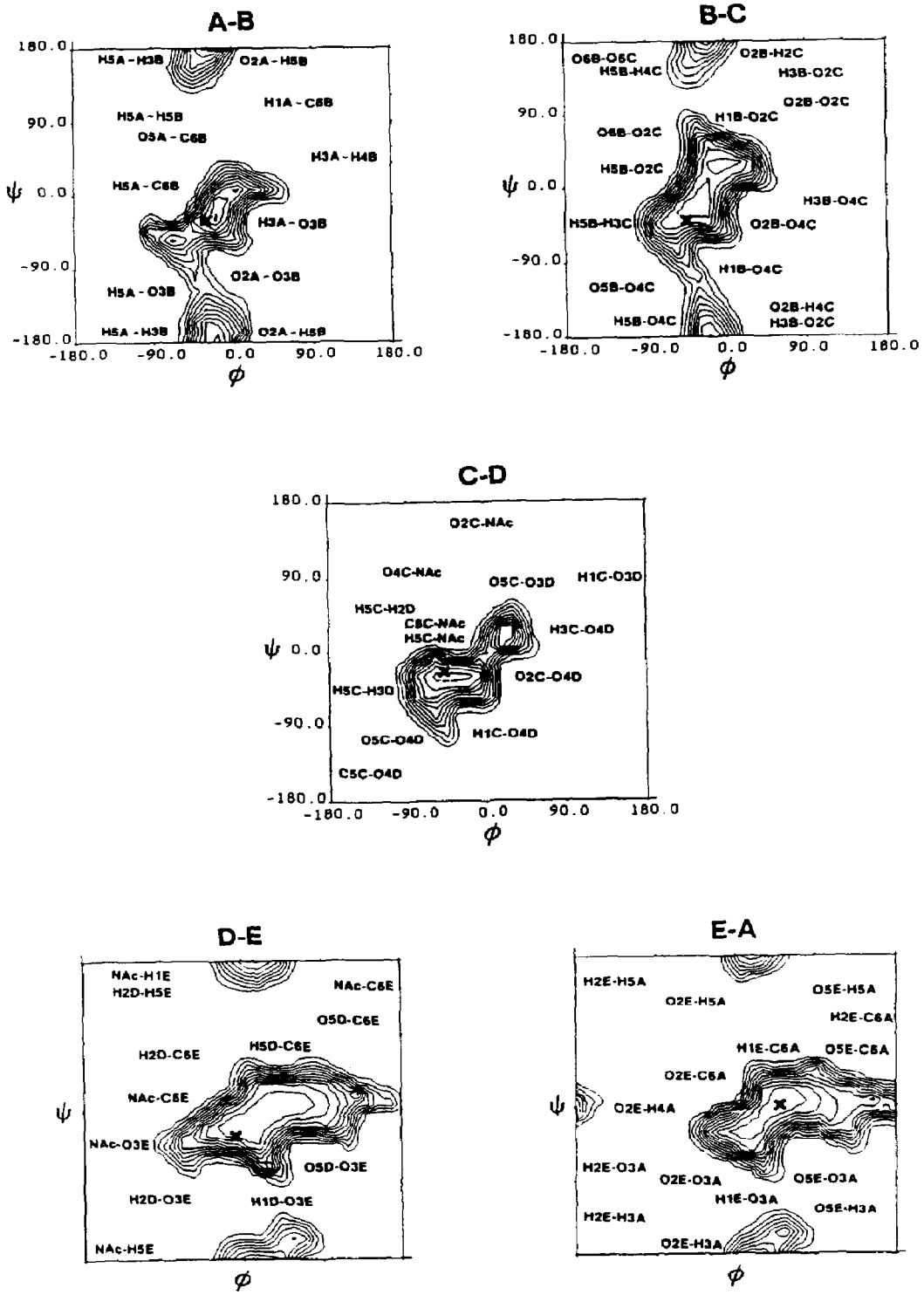

Fig. 3. Rigid-residue VDW potential surfaces of disaccharide model fragments of the 9N PS, 5° increments in ϕ and ψ, iso-energy contours are at equal (nominally 5 kcal/mol) intervals.

Table 5
Glycosidic torsion angles (ϕ; ψ) determined from molecular modelling calculations

Linkage	VDW [a]	MM2 [a]	AMBER annealed	MD [b] time-averaged
α-GlcNAc-(1→4)-α-GlcA	−34 ; −33	−33 ; −31	−42 ; −16	−44 ; −25
A–B	(−70 ; −56)	(−70 ; −66)		(10.8 ; 14.6)
α-GlcA-(1→3)-α-Glc	−18 ; −43	−44 ; −41	−46 ; −9	−50 ; −16
B–C	(−42 ; −43)	(−16 ; −43)		(11.2 ; 19.9)
α-Glc-(1→3)-β-ManNAc	−43 ; −19	−44 ; −20	−57 ; −12	−51 ; −3
C–D				(7.2 ; 17.9)
β-ManNAc-(1→4)-β-Glc	65 ; −10	54 ; 14	48 ; 3	45 ; −3
				(13.4 ; 15.1)
D–E	(31 ; −31)	(60 ; −3)		
	(52 ; 11)	(38 ; −41)		
β-Glc-(1→4)-α-GlcNAc	7 ; −43	47 ; −9	50 ; 2 [c]	409 ; 0
E–A	(69 ; 5)	(5 ; −47)	60 ; 0	(13.3 ; 12.2)
	(45 ; −7)	(63 ; 5)		

[a] Secondary minima are given in parenthesis. [b] The rms deviations of ϕ and ψ are given in parenthesis. [c] The two examples of this glycosidic linkage in the nonasaccharide optimised to different values.

The energy-minimised model of S9 was an extended helix, and the major axis of rotation was fitted as the helix axis, and T_1 values plotted against σ. As no simple function was found to relate ^{13}C T_1 and σ, the simplest model, isotropically tumbling with a single effective rotational correlation time (τ_c) was used. This correlation time was estimated as 5 ns from the mean ^{13}C T_1 value for methine carbons (256 ± 47 ms, at 67.5 MHz), assuming dipole–dipole relaxation through the attached proton ($r_{CH} = 0.116$ nm) as the sole relaxation mechanism.

NOE time courses were simulated for the NOE time course experiments performed. Under the conditions used for the NOE difference experiments, the low power irradiation of resonances of interest does not cause the instantaneous saturation assumed by the simulation routines. Extrapolation of the linear region of experimental NOE time course to zero

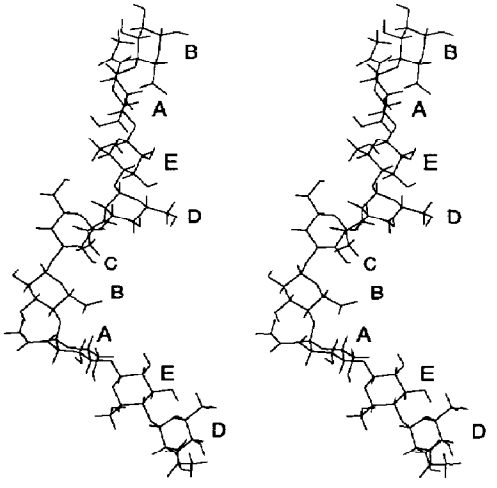

Fig. 4. Stereo diagram of the annealed 9N CPS model.

Table 6
Comparison of experimental and simulated NOEs. NOE simulations were carried out with NOEMOL using nonasaccharide models in which the glycosidic angles were set to those found in the lowest energy MM2CARB conformation and annealed AMBER conformation

Irradiated resonance	Observed resonance	Experimental NOE, 3 experiments	MM2 minimum	Simulated NOEs	AMBER annealed	Simulated NOEs
			ϕ ; ψ	σ	ϕ ; ψ	σ
α-GlcNAc H-1	α-GlcNAc H-2	−21; −18; −16	−33 ; −31	−19	−42 ; −16	−21
	α-GlcNAc H-3	−4; −3; −1		−2		−3
	α-GlcA H-4	−7; −6; −5		−27		−23
α-GlcA H-1	α-GlcA H-2	−19; −17; −15	−44 ; −41	−19	−46 ; −9	−22
	α-GlcA H-3	−1; −1; −2		−2		−3
	α-Glc H-3	−9; −9; −8		−24		−19
α-Glc H-1	α-Glc H-2	−15; −12; −11	−44 ; −20	−18	−57 ; −12	−20
	α-Glc H-3	−5; −2; −2		−2		−3
	β-ManNAc H-3	−14; −11; −12		−17		−12
β-ManNAc H-1	β-ManNAc H-2	?; −5; −5	54 ; 14	−17	48 ; 3	−16
	β-ManNAc H-3	−13; −9; −10		−14		−16
	β-ManNAc H-5	−7; −7; −7		−21		−20
	β-Glc H-3/H-4	−21; −18; −17		−23		−25
β-ManNAc H-2	β-ManNAc H-3	?= [a]	54 ; 14	−26	48 ; 3	−18
	α-Glc H-5	?=		−14		−26
β-Glc H-1	β-Glc H-2	−; −1.5; −1.5	47; −9	−6	50 ; 2	−6
	β-Glc H-3/H-4} α-GlcNAc H-4}	−; −13; −12		−41		−16, −2, −19 } −37
	β-Glc H-5	−; −11; −11		−20		−21
	α-GlcNAc H-6	−; −7; −5		−11		−9

[a] Not quantified: NOEs from β-ManNAc H-2 to β-ManNAc H-3 and β-Glc H-5 approximately equal in intensity.

NOE indicated a delay of 80–100 ms, hence the simulated values given in Table 6 (at 200 ms) are equivalent to those of the experimental values at 300 ms preirradiation. The absolute value of simulated enhancements were higher than experimental values (by up to 4×), and some enhancements that presumably arose via spin diffusion [e.g., H-1(A)···H-1(B) 0.5%]. Although most NOEs for the annealed structure in Table 5 have reasonable agreement with experiment, NOEMOL predicted many other large NOEs that were not observed in practice, e.g., H-1(D)···H-6R(E) no NOE observed, simulated NOE = −12%, no single rigid conformation could be found that satisfied all of the experimental measurements. It is likely that experimentally measured interresidue NOE values are attenuated by internal flexibility about the glycosidic dihedral angles.

The simulated distances to the H-6's were sensitive to the orientation of the hydroxymethyl groups, and, given the slow interconversion between the hydroxymethyl rotamers, the conformational space will not be sampled sufficiently on the timescale of the MD simulations. Each of the hydroxymethyl groups sampled all three rotamers during the MD timecourse, with between three and thirteen transitions observed. Since this was not a

statistically significant sampling of the transitions, distances measured to H-6's must be considered approximate.

Part of the discrepancy between simulations and experimental data is the neglect of internal motions and conformational averaging. The relatively narrow linewidths and short correlation times for such a high-molecular-weight polymer (10^6 Da) clearly indicate rapid internal motions within the polymer chains. NOE simulations of S9 fragments should allow for the influence of rapid internal motions.

The flexibility of S9 was explored by restrained molecular dynamics (AMBER force-field, in vacuo) on the annealed structure. NOE constraints were applied throughout (Table 3). The last 500 ps of a 510 ps simulation were used for back-calculating NMR data. ϕ/ψ trajectories for each linkage are shown in Fig. 5, and the time-averaged linkage geometries are given in Table 5. Effective interproton distances were evaluated from the MD data (Table 7) using the Tropp formalism to account for rapid interconversion of the conformational states. Tropp-averaging has previously been shown to apply to oligosaccharides [30] where the rates of internal motion are rapid compared with molecular rotation.

Tropp-averaged interproton distances were generally in close agreement with experimentally derived distances. The flexibility attenuates the magnitude of effective interresidue ^{1}H–^{1}H distances and therefore the NOE data is more consistent with the dynamic model than the static models, e.g., the H-1(D)···H-6R(E) NOE, was over estimated in the static model (-12%) but has an effective distance $r_{\text{Tropp}} = 0.357$ nm, which would not give rise to an observable NOE. S9 may approximate to a single conformation with glycosidic linkages oscillating rapidly within a single energy well, similar to the resticted random coil described for the *Shigella flexneri* type Y polysaccharide [13], which exhibits random ϕ,ψ oscillation of $\sim \pm 10°$.

Rationalisation of glycosylation shifts.—The most suitable models from restrained molecular dynamics calculations were examined for consistency with conformational-dependent chemical shifts. ^{13}C Glycosylation shifts of C-1 or C-α resonances have been correlated [41–44] with the glycosidic dihedral angle, ψ, and applying correlations obtained from glycosides of α-Glc and α-Gal gave estimates of ψ for the three α-linkages in good agreement with results from the modelling studies (Table 4). The resonances of a number of protons not associated directly with the glycosidic linkage experience downfield shifts (compared to model systems) and which are usually considered due to proximity to an oxygen atom in an adjacent residue. ^{1}H Glycosylation shifts are, however, affected both by through-space and through-bond effects, which complicates their interpretation and renders them useful only as a rough guideline of steric interactions.

The relatively large downfield ^{1}H glycosylation shift of the anomeric proton H-1(A) indicated proximity with an oxygen atom, consistent with the H-1(A)···O-3(B) time-averaged distance of 0.261 nm in the model. This interaction also accounts for the large ^{13}C glycosylation shift on C-3(B). Similarly, the downfield chemical shift of H-5(B) can be explained by its proximity to O-2(C) (0.268 nm), and both C-5(B) and C-2(C) experience glycosylation shift.

The time-averaged interoxygen distance O-5(D)···O-3(E) is 0.302 nm, convenient for an interresidue hydrogen bond as postulated in the crystal structure of cellobiose [45,46]. This would provide an explanation for the C-3(E) glycosylation shift and the resistance of the β-Glc residue to mild periodate oxidation [1], and likewise, the O-5(E)···O-3(A)

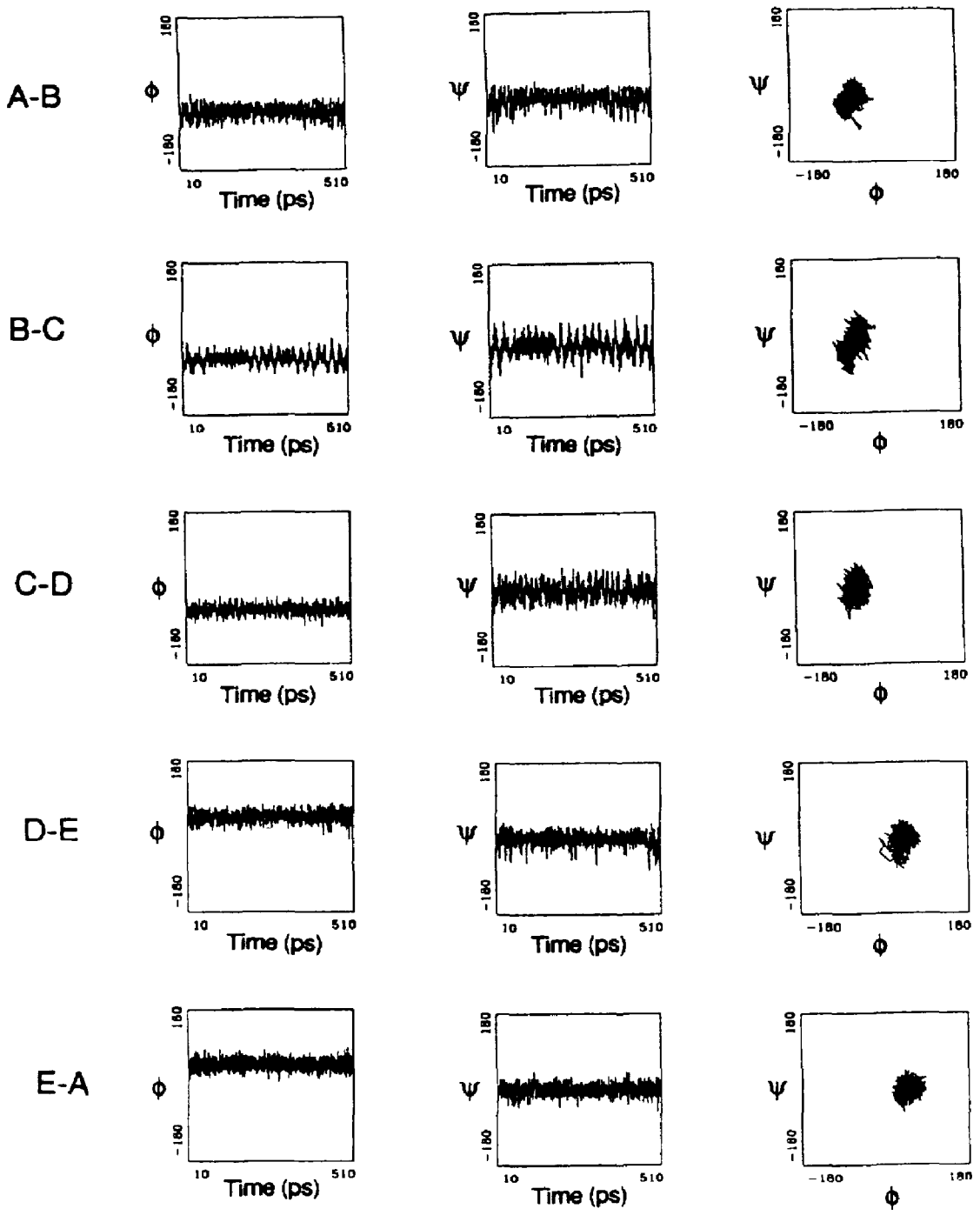

Fig. 5. Molecular dynamics ϕ/ψ trajectories.

distance (0.303 nm) would also facilitate H-bond formation. However, this hydrogen bond has not been observed in protic solvents [47].

The other significant ^{13}C glycosylation shifts were also accounted for by short (<3 Å) time-averaged interatomic distances; δ_C C-4(C) is affected by H-1(B)···O-4(C) proxim-

Table 7
Comparison of experimentally-derived ^1H–^1H distances with Tropp-averaged distances (r_{Tropp}) calculated from MD data

	Experiment [a]	r_{Tropp} (nm)
H-1(A)···H-4(B)	0.29	0.269
H-1(B)···H-3(C)	0.28	0.265
H-1(C)···H-3(D)	0.25	0.256
H-5(C)···H-2(D)	0.26	0.261
H-1(D)···H-4(E)	0.20	0.248
H-1(E)···H-4(A)	0.25	0.269
H-1(E)···H-6R(A′)	0.31	0.329

[a] Calculated using an isolated spin-pair approximation.

ity, δ_C C-5(C) by H-5(C)···H-2(D) proximity, and δ_C C-4(D) by H-1(C)···O-4(D) proximity.

Temperature dependence of conformation.—For practical reasons, NMR experiments were carried out at elevated temperature, and we were concerned with how the conformation might differ from that at physiological temperatures. Rheological studies were used to investigate the temperature dependence of the S9 conformation, as stress–strain rigidity moduli of polymer solutions are sensitive to changes in molecular organisation of the polymer. The shear storage modulus (G') and the shear loss modulus (G'') of S9 were measured at a range of temperatures (25–70°C) using a cone-and-plate viscoelastic rheometer. There was no significant change in either G' or G'' as a function of temperature, consistent with an overall molecular organisation of S9 that is essentially the same at 70°C as at physiological temperatures.

NOE values at 30°C were compared with those at 70°C, and indicated that at high temperatures H-1′···H-α distances are larger for links A–B and B–C but smaller for C–D and D–E. Due to the uncertain influence of spin diffusion at 30°C, interpretation of this data should be regarded with caution, but as a rough guide the calculated H-1′···H-α distances differed by ≤ 0.03 nm from those at 70°C (corresponding to $\leq 15°$ difference in ψ).

The correlations of carbon chemical shift with glycosidic angle proposed by Bock et al. [42] indicated that a change of 10° in ψ causes a 1–2 ppm change in the aglycon C-α chemical shift, and can be used to judge the magnitude of changes in glycosidic angle with temperature. Between 25 and 70°C, the chemical shifts of C-α resonances in S9 changed by 0–0.9 ppm [48] corresponding to, at most, a 0–9° change in the time-averaged value of ψ ($<\psi>$), between these two temperatures. Changes in $<\psi>$ should also be reflected in the C-1′ chemical shift, although these were considerably less temperature sensitive (<0.3 ppm between 25 and 70°C) than aglycon C-α shifts. All ^1H chemical shifts were essentially independent of temperature over this range.

Characteristics of the optimum model.—The optimum conformation model (Fig. 4) forms an extended ribbon conformation, with glycosidic torsion angles that are comparable to those determined for disaccharides and indicate that the conformation of S9 is a summation of the conformation of the disaccharide elements, with no evidence of distortion in favour of a stable secondary structure. The significance of the topographical features may

become clearer on comparing the S9 PS conformation with those of the other serogroup 9 capsules.

4. Conclusions

Conformational data for the high-molecular-weight polysaccharide from *S. pneumoniae* type 9N were representative of the entire homogeneous sample, and were consistent with low energy conformations determined from molecular mechanics calculations. However, experimental NMR measurements were in closest agreement with time-averaged values from restrained molecular dynamics simulations. The time-averaged conformation was a summation of the conformations of individual disaccharide elements, determined by local interactions, with no evidence that long-range associations stabilise a secondary structure. No significant temperature dependence of the conformation was inferred from NMR experiments recorded, hence the conformation at 70°C is very similar to that at physiological temperature.

Acknowledgements

We thank the S.E.R.C. for a CASE studentship (T.J.R.) and access to 500-MHz NMR facilities (Birkbeck College, U.L.I.R.S.), the M.R.C. Biomedical NMR Centre (Mill Hill) for access to NMR facilities, and Merck, Sharpe, and Dohme for their gift of the type 9N PS. We are also grateful to Miss P. Newland (NIBSC) for molecular-weight determinations, Dr. Alan Bell (Food Science Department, Reading University) for rheological measurements, and Mr. Martin Kipps (ZENECA Agrochemicals) for helpful discussions.

References

[1] C. Jones, B. Mulloy, A. Wilson, A. Dell, and J.E. Oates, *J. Chem. Soc., Perkin Trans. 1*, (1985) 1665–1673.
[2] J.C. Richards, M.B. Perry, and P.J. Kniskern, *Can. J. Biochem.*, 62 (1984) 1309–20.
[3] J.C. Richards and M.B. Perry, in A.M. Wu (Ed.), *The Molecular Immunology of Complex Carbohydrates*, Plenum, 1988, p 594.
[4] M.B. Perry, V. Daoust, and D.J. Carlo, *Can. J. Biochem.*, 59 (1981) 524–533.
[5] T.J. Rutherford, C. Jones, D.B. Davies, and A.C. Elliott, *Carbohydr. Res.*, 218 (1991) 175–184.
[6] S. Szu, C.-J. Lee, D. Carlo, and J. Henrichsen, *Infect. Immun.*, 31 (1981) 371–379.
[7] M. Heidelberger, *Infect. Immun.*, 41 (1983) 1234–1244.
[8] M. Heidelberger, *J. Immunol.*, 91 (1963) 735–739.
[9] E.A. Kabat, *J. Immunol.*, 84 (1960) 82–85.
[10] K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986.
[11] A.D. French and J.W. Brady (Eds.), *Computer Modelling of Carbohydrate Molecules, ACS Symp. Ser.*, 430 (1990).
[12] C. Jones, F. Currie, and M.J. Forster, *Carbohydr. Res.*, 221 (1991) 95–121.
[13] K. Bock, S. Josephson, and D.R. Bundle, *J. Chem. Soc., Perkin Trans. 2*, (1982) 59–70.
[14] J.C. Lindon, J.G. Vintner, M.R. Lifely, and C. Moreno, *Carbohydr. Res.*, 133 (1984) 59–74.
[15] G.M. Lipkind and N.K. Kochetkov, *Bioorg. Khim.*, 11 (1985) 293–301.
[16] P. Newland, B. Bingham, E. Tarelli, and A.H. Thomas, *J. Chromatogr.*, 483 (1989) 406–412.

[17] D.J. States, R.A. Haberkorn, and D.J. Ruben, *J. Magn. Reson.*, 28 (1982) 286.
[18] H. Kessler, C. Griesinger, J. Zarbock, and H.R. Loosli, *J. Magn. Reson.*, 58 (1984) 331–336.
[19] W.F. Reynolds, S. McLean, M. Perpick-Dumont, and R.G. Enriquez, *Magn. Reson. Chem.*, 27 (1989) 162–169.
[20] A. Bax, *J. Magn. Reson.*, 57 (1984) 314–318.
[21] E. Altman, J.R. Brisson, D.R. Bundle, and M.B. Perry, *Can. J. Biochem. Cell. Biol.*, 65 (1987) 876–889.
[22] M.F. Summers, L.G. Marzilli, and A. Bax, *J. Am. Chem. Soc.*, 108 (1986) 4285–4294.
[23] P.J. Hore, *J. Magn. Reson.*, 54 (1983) 539.
[24] P.J. Hore, *J. Magn. Reson.*, 55 (1983) 283–300.
[25] N.L. Allinger and Y.H. Yu, *QCPE*, 13 (1981) 395.
[26] G.A. Jeffrey and R. Taylor, *J. Comp. Chem.*, 1 (1980) 99–109.
[27] I. Tvaroška and S. Pérez, *Carbohydr. Res.*, 149 (1986) 389.
[28] V.C. Singh, P. Weiner, J. Caldwell, and P.A. Kollman, AMBER 3-0, University of California, San Francisco.
[29] S.W. Homans, *Biochemistry*, 29 (1990) 9110–9118.
[30] S.W. Homans and M.J. Forster, *Glycobiology*, 2 (1992) 143–151.
[31] M.J. Forster, C. Jones, and B. Mulloy, *J. Mol. Graphics*, 7 (1989) 196–201.
[32] B. Mulloy, T.A. Frenkiel, and D.B. Davies, *Carbohydr. Res.*, 184 (1989) 39–46.
[33] I. Tvaroska, M. Hricovini, and E. Petrakova, *Carbohydr. Res.*, 189 (1989) 359–362.
[34] A.S. Shashkov, G.M. Lipkind, and N.K. Kochetkov, *Magn. Reson. Chem.*, 26 (1988) 735–747.
[35] H. Baumann, Ph.D. Thesis, *Univ. Stockholm Chem. Comm.*, 9 (1988).
[36] Y. Nishida, H. Hori, H. Ohrui, and H. Meguro, *J. Carbohydr. Chem.*, 7 (1987) 239–250.
[37] O. Jardetzky and G.C.K. Roberts, *NMR in Molecular Biology*, Academic, New York, 1981, p 153.
[38] F.H. Allen, S.A. Bellard, M.D. Brice, B.A. Cartwright, A. Doubleday, H. Higgs, T. Hummelink, B.G. Hummerlink-Peters, O. Kennard, W.D.S. Motherwell, J.R. Rodgers, and D.G. Watson, *Acta Crystallogr., Sect. B*, 35 (1979) 2331.
[39] A.D. French and J.W. Brady (Eds.), *Computer Modelling of Carbohydrate Molecules, ACS Symp. Ser.*, 430 (1990).
[40] T.J. Rutherford, J. Partridge, C.T. Weller, and S.W. Homans, *Biochemistry*, 32 (1993) 12715–12724.
[41] O. Jardetzy, *Biochim. Biophys. Acta*, 221 (1980) 227–232; see also D. Cumming and J.P. Carver, *Biochemistry*, 26 (1987) 6664–6676.
[42] K. Bock, A. Brignole, and B.W. Sigurskjold, *J. Chem. Soc., Perkin Trans. 2*, (1986) 1711–1713.
[43] R.P. Veregrin, C.A. Fyfe, R.H. Marchessault, and M.G. Taylor, *Carbohydr. Res.*, 160 (1987) 41–56.
[44] M.J. Gidley and S.M. Bociek, *J. Am. Chem. Soc.*, 110 (1988) 3820–3829.
[45] C.J. Brown, *J. Chem. Soc., A*, (1966) 927–932.
[46] D.A. Rees, in G.O. Aspinall (Ed.), *Carbohydrates, MTP Int. Rev. Sci., Ser.* 1, Vol. 7, Butterworths, London, 1973.
[47] B.R. Leeflang, J.F.G. Vliegenthart, L.M.J. Kroon–Batenburg, B.P. Eijck, and J. Kroon, *Carbohydr. Res.*, 230 (1992) 41–61.
[48] T.J. Rutherford, C. Jones, D.B. Davies, and A.C. Elliott, *Carbohydr. Res.*, 265 (1994) 97–111, accompanying paper.