ELSEVIER

Carbohydrate Research 337 (2002) 335–344

CARBOHYDRATE RESEARCH

www.elsevier.com/locate/carres

Structural analysis and chemical depolymerization of the capsular polysaccharide of *Streptococcus pneumoniae* type 1

Corné J.M. Stroop,[a] Qiuwei Xu,[b] Mary Retzlaff,[b] Chitranda Abeygunawardana,[b] C. Allen Bush[a,*]

[a]*Department of Chemistry and Biochemistry, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA*
[b]*Bioprocess and Bioanalytical Chemistry, Merck Research Laboratories, West Point, PA 19486, USA*

Received 17 July 2001; accepted 3 December 2001

Abstract

NMR spectroscopy can be used to characterize bacterial polysaccharides such as that of *Streptococcus pneumoniae* type 1 which is a component of the 23-valent pneumococcal vaccine in clinical use. This particular polysaccharide gives NMR spectra with wide lines apparently due to restricted molecular mobility and chain flexibility which leads to rapid dipolar T_2 relaxation limiting the possibility of detailed spectral analysis. Removal of *O*-acetyl groups found on approximately two thirds of the repeating subunits of pneumococcal type 1 capsule leads to narrower NMR lines facilitating a complete assignment of the ^{1}H and ^{13}C NMR spectra. Degradation of the polysaccharide by periodate oxidation followed by base treatment leads to an oligosaccharide fragment of approximately three repeating trisaccharide units. This oligosaccharide has narrow NMR lines and ^{1}H and ^{13}C assignments very similar to those of the *O*-deacetylated polysaccharide. In the native polysaccharide, *O*-acetyl groups are located on the 2- and 3-positions of the 4-linked galacturonic acid residue providing protection against periodate oxidation. Analysis of NOESY spectra combined with molecular modeling of the oligosaccharide shows that flexibility occurs in certain of the saccharide linkages.

Keywords: NMR; Capsular polysaccharide; Depolymerisation; Vaccine; *Streptococcus pneumoniae*

1. Introduction

Streptococcus pneumoniae is a major pathogen responsible for such infections as pneumonia, meningitis, otitis media, scepticemia and sinusitis.[1] Its capsular polysaccharides (CPSs), constituting about 90 different serotypes, are a key factor in the virulence of these bacteria. CPSs are not toxic by themselves, but they inhibit ingestion of the bacterium by host neutrophils. Different serotype structures are associated with differences in immune response, infectivity and disease.[2,3]

Serotype 1 is among the core group of serotypes that is associated with most of the pneumococcal diseases. It is included in the 23-valent pneumococcal vaccine preparation which is in regular clinical use. The CPS of *S. pneumoniae* type 1 consists of a repeating trisaccharide:

→4)-α-GalA-(1→3)-α-GalA-(1→3)

-α-6dGalNAc4N-(1→

with 6dGalNAc4N = 2-acetamido-4-amino-2,4,6-trideoxy-D-galactose, which is, like the two D-galacturonic acid residues, a pyranose. The polysaccharide also contains a nonstoichiometric amount of *O*-acetyl groups, the position of which has not been reported.[4]

NMR spectroscopy is a valuable technique for characterization of the chemical structure of complex bacterial polysaccharides and for confirmation of identity and purity of vaccine preparations.[5] If these capsular polysaccharides tumble as rigid bodies, their molecular weights, well in excess of 100 kDa, would be expected to lead to slow rotational relaxation and extensive dipolar line broadening (short T_2). The fact that such high-molecular-weight polysaccharides give recognizable NMR spectra implies that there must be some

* Corresponding author. Tel.: +1-410-4552506; fax: +1-410-4552608.
E-mail address: bush@umbc.edu (C.A. Bush).

0008-6215/02/$ - see front matter
PII: S0008-6215(01)00318-4

internal flexibility leading to longer T_2 relaxation times. While the exact nature of this internal motion is not completely understood, it is quite variable among polysaccharides of differing chemical structure. For example, the cell-wall polysaccharide of *S. mitis* J22 has relatively narrow NMR lines of just a few Hz, and measured T_2 values are in accord with this observation.[6] On the basis of extensive NMR and molecular modeling studies of this polysaccharide, we have proposed motions about specific hinges that are particularly flexible.[7] In contrast to the J22 polysaccharide, certain polysaccharides, such as the *S. pneumoniae* type 1 CPS, give very broad NMR lines complicating the use of this technique for structural characterization. For such polysaccharides, elevated probe temperatures often improve the NMR spectra, but for some polysaccharides, temperatures of 70 or 80 °C do not render ^{1}H NMR linewidths below 30 Hz.

If higher temperatures do not yield sufficiently narrow spectral lines, chemical modification may provide improved spectra. In this paper, we show that removal of *O*-acetyl groups has a significant effect on the line width. We also report a new method for polysaccharide degradation involving periodate oxidation, followed by base degradation, which provides oligosaccharide fragments from the pneumococcal type 1 polysaccharide. The resulting oligosaccharide fragments were studied by 2D NMR experiments, and molecular modeling results were compared with NMR data of the oligosaccharide fragments to yield a three-dimensional model for the type 1 repeating unit.

2. Results and discussion

NMR analysis.—The 1D ^{1}H NMR spectrum of the *S. pneumoniae* type 1 CPS shows broad and overlapping signals at 60 °C, while the spectrum of the *O*-deacetylated polysaccharide at 50 °C shows lines which are sufficiently narrow ($\sim$3 Hz) to allow for a complete spectral assignment by the customary coherence transfer method (Fig. 1). The completely assigned spectrum of the *O*-deacetylated polysaccharide, given in Table 1, was derived from DQF-COSY, TOCSY and HMQC spectra (not shown).

In chemical degradation studies by periodate oxidation, the only residue in the *S. pneumoniae* type 1 CPS which is expected to be oxidized is the 4-substituted galacturonic acid. (See Fig. 2 for the chemical structure of the polysaccharide.) When the periodate-oxidized polysaccharide is subsequently treated with base, a β elimination reaction is expected to degrade the oxidized uronic acid residue leading to cleavage of the polysaccharide producing oligosaccharides whose nonreducing terminal residue is the 3-substituted galacturonic acid residue (**B**), along with residue **C** glycosidically attached to a fragment of the degraded residue at the reducing terminal.

The spectrum of treated CPS (**Spt1**) shows narrow peaks of approximately equal height at 27 °C (Fig. 1), indicating that the native CPS was successfully depolymerized by periodate oxidation and base degradation, providing a sample that could readily be analyzed by NMR spectroscopy. The ^{1}H and ^{13}C chemical shifts of the degraded **Spt1** sample are fully assigned (Table 1), based on 1D (Fig. 1) and 2D TOCSY (Fig. 2), 2D COSY (not shown), 2D HMQC (Fig. 3) NMR experiments. The three α-anomeric ^{1}H signals at δ 5.22, 5.06, and 4.97 represent the three anomeric protons of the monosaccharides (**A**, **B**, and **C**) in the repeating unit. The H-1 tracks of these resonances in the 100-ms TOCSY spectrum (Fig. 2) show the correlations up to proton H-4, typical for pyranose monosaccharides with a galacto configuration. The DQF-COSY and the short-mixing time TOCSY spectra were used to identify the cross peaks for H-2, H-3 and H-4 protons (Table 1). The H-4 tracks in the TOCSY spectrum (Fig. 2) yield connectivities to the H-5 of residues **A** and **B**, but not to the H-6, which is consistent with residues **A** and **B** being galacturonic acid without a H-6 proton. In the NOESY spectrum to be discussed below, cross peaks between H-4 and H-5 are prominent for all residues, and the HMBC spectrum (not shown) also has strong cross peaks connecting H-1 with C-5 resonances. The H-5 of residue **C** shows a TOCSY cross peak with a methyl (H-6) signal at δ 1.26, indicative of the 6dGalNAc4N residue. All carbon resonances were identified through the HMQC spectrum (Fig. 3). In residue **C**, the signals for C-2 and C-4 at δ 48.2 and 53.7, respectively, are relatively upfield corresponding to a residue with amino groups linked to the C-2 and C-4 carbon. One of the two amino groups is acetylated, as is shown by the *N*-acetyl methyl signals at δ 2.04 (^{1}H) and δ 22.8 (^{13}C). The HMBC spectrum (data not shown) shows, in addition to intraresidue correlations between H-1 and C-3 and C-5 expected for α-sugars, correlations between **A** H-1 and **B** C-3 and between **C** H-1 and **A** C-4. In the anomeric carbon atom resonance region, correlations are seen between **B** C-1 and **C** H-3 and between **C** C-1 and **A** H-4. The NOESY data of Fig. 4 also support the assignment of linkage and sequence indicated in Fig. 2 and originally proposed by Lindberg et al.[4]

The spectral assignment described for the degraded polysaccharide is that of a repeating trisaccharide indicating that the polysaccharide is not completely cleaved at each residue of 4-substituted galacturonic acid. Likewise, the assignment is essentially identical to that of the *O*-deacetylated polysaccharide. Minor differences in the spectra arise since they were recorded at different temperatures. The oligosaccharide product is presumably *O*-deacetylated by the base treatment subsequent to the periodate oxidation.

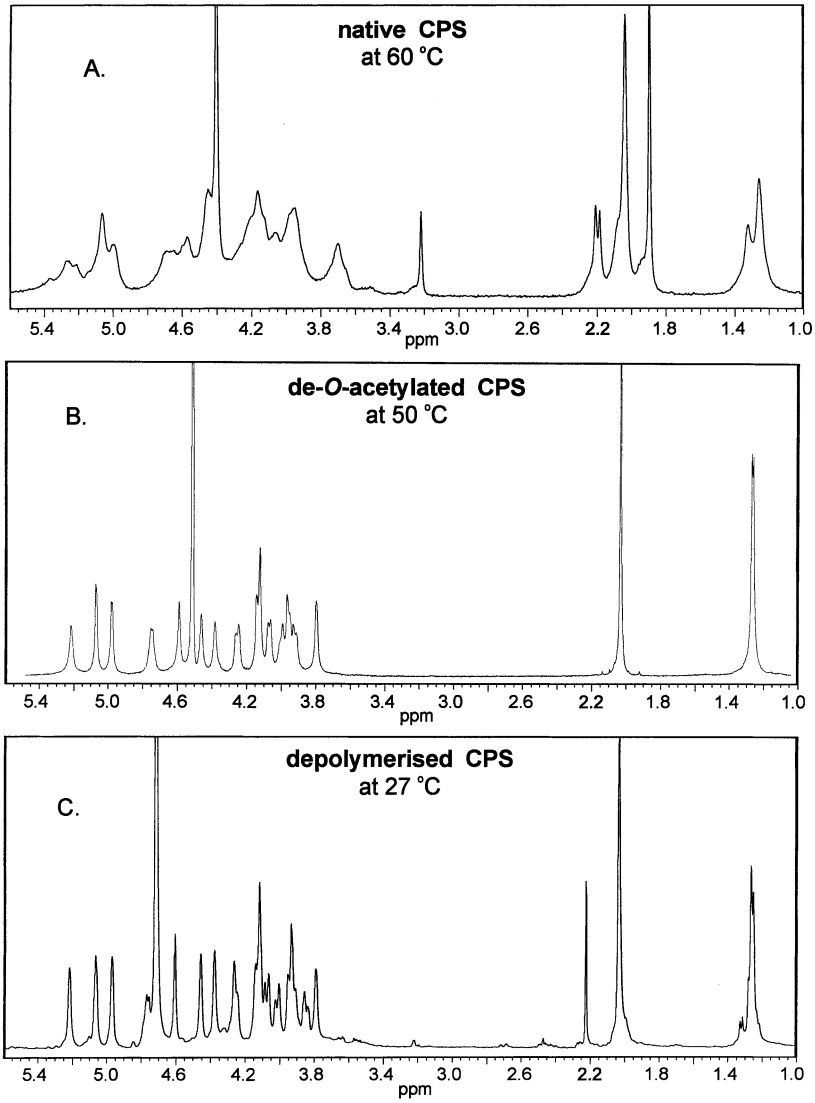

Fig. 1. 1D ^{1}H NMR spectra of native *S. pneumoniae* type 1 CPS (A), *O*-deacetylated CPS (B) and CPS after periodate oxidation and base degradation (C). The resonance at 2.225 ppm in C is acetone as internal standard.

As can be seen in the 2D spectra (Figs. 2–4), signals of lower intensity than those of the residues in the repeating unit are also present. The additional anomeric resonances at δ 5.07 and 98.5 (^{1}H and ^{13}C, respectively, see Fig. 3) for a residue, designated **B′**, has connectivities with signals at δ 3.86 and 4.26 (Fig. 2). The HMQC spectrum shows that the signal at 3.86 ppm stems from two different protons, one with a ^{13}C signal at δ 68.1 and one at δ 70.2. These values (Table 1) are consistent with residue **B′** being the nonreducing terminal of the **Spt1** oligosaccharide fragment as reported by Bock et al.,[8] who gave ^{13}C and ^{1}H assignments for a synthetic disaccharide from the pneumococcal type 1 polysaccharide. Connectivities in the HMBC data (not shown) also confirm this interpretation. Integration of the H-2 and H-3 signals for residue **B** and those for **B′** in the HMQC spectrum gives a ratio of ca. 2:1, suggesting that the average size of the **Spt1** oligosaccharide is a nonasaccharide.

Our data can also be used to determine the positions of *O*-acetylation of the polysaccharide. In Fig. 5(a), an expansion of the methyl group region of the 1D ^{1}H spectrum of the native polysaccharide, two distinct methyl group resonances of approximately equal size are seen at 2.18 and 2.21 ppm, and the integration of these two peaks shows a total of 0.68 *O*-acetyl groups per repeating subunit. Notice that the effect of acetylation shifts the H-6 resonances of the 6dGalNAc4N residue, all of which collapse to a single signal at 1.27 ppm after *O*-deacetylation (Fig. 1).

The 0.68 acetyl residues must be distributed among the two or more out of four possible positions on the GalA residues **A** and **B**. Since the *O*-acetyl survives periodate treatment, as indicated by NMR spectra of

Table 1
^{1}H and ^{13}C chemical shifts of the constituting monosaccharides of depolymerized CPS (**Spt1**) and of *O*-deacetylated CPS

Residue [a]	Position	**Spt1** [b]		*O*-Deacetylated CPS [c]	
		^{1}H (ppm)	^{13}C (ppm)	^{1}H (ppm)	^{13}C (ppm)
A (GalA)	1	5.22	97.4	5.22	97.2
	2	3.92	68.6	3.92	68.6
	3	4.07	69.4	4.07	69.5
	4	4.38	80.1	4.38	79.9
	5	4.61	71.7	4.58	71.8
B (GalA)	1	5.06	99.7	5.07	99.5
	2	3.94	66.5	3.96	66.6
	3	4.02	76.8	3.99	76.6
	4	4.46	68.6	4.47	68.5
	5	4.12	73.1	4.12	72.8
C (6dGalNAc4N)	1	4.97	98.9	4.98	99.1
	2	4.13	48.2	4.14	48.4
	3	4.26	74.1	4.26	73.8
	4	3.79	53.7	3.80	53.6
	5	4.77	63.6	4.74	63.5
	6	1.26	16.0	1.27	15.8
	NAc	2.04	22.8	2.03	22.8
B′ (GalA)	1	5.07	98.5		
	2	3.86	68.1		
	3	3.86	70.2		
	4	4.26	71.4		
	5	4.12	73.1		

[a] For residue labels, see Fig. 2.
[b] Recorded at 27 °C.
[c] Recorded at 30 °C.

the oxidized polysaccharide (not shown), its presence on the 4-substituted GalA residue (**A**) could provide an explanation for the nonasaccharide product of our degradation procedure. A more quantitative estimate of the resistance to periodate can be obtained from Fig. 5(b) showing the anomeric region of the 1D ^{1}H NMR spectrum of the periodate-oxidized and borohydride-reduced CPS. Integration shows that the anomeric resonance of the 4-substituted GalA residue (**A**) is diminished to 0.64 with respect to the anomeric proton of GalA residue **B**. This result suggests that single acetyl groups are on either the 2- or the 3-position, but not on both positions, of residue **A**, a configuration which gives maximum resistance to periodate.

An HSQC spectrum of the native polysaccharide (not shown) features two peaks at chemical shifts typical of acetylated sugars at 4.99/70.4 ppm and at 5.23/77.8 ppm. A DQF-COSY spectrum of this same sample shows that the peak at 4.99 ppm correlates with the anomeric resonance at 5.36 ppm, indicating that this is H-2 of an acetylated residue, while the peak at 5.23 ppm belongs to H-3 of a residue acetylated in that position. These spectra, as well as the 1D spectrum in the methyl region (Fig. 5(a)), indicate that there are just two different types of *O*-acetylated subunits, and we propose that each contains single acetates at either the 2- or the 3-position of residue **A**.

Our procedure for degradation of the polysaccharide differs from classical Smith degradation in which periodate oxidation is followed by borohydride reduction and acid hydrolysis. The wide variation in acid susceptibility of the acetals of glycosidic linkages leads to multiple cleavage points in Smith degradation of some polysaccharides, such as the capsule of *S. pneumoniae* type 1, and it is not useful for polysaccharides with very acid-labile groups such as 3,6-dideoxy sugars. The NMR spectra showing a single nonreducing terminal sugar (**B′**) with the present method indicates cleavage of the *S. pneumoniae* capsule at a single point. The exact identity of the reducing side of this cleavage point is less clear. A possible mechanism for base degradation is β elimination of the oxidized GalA residue, **A**. Abstraction of the C-5 proton would lead to cleavage of the glycosidic linkage of residue **C**, producing a reducing terminal much like β elimination of a uronic acid residue. Our HMQC spectra show no evidence of resonances of a free reducing terminal with intensity comparable to those of the nonreducing terminal residue,

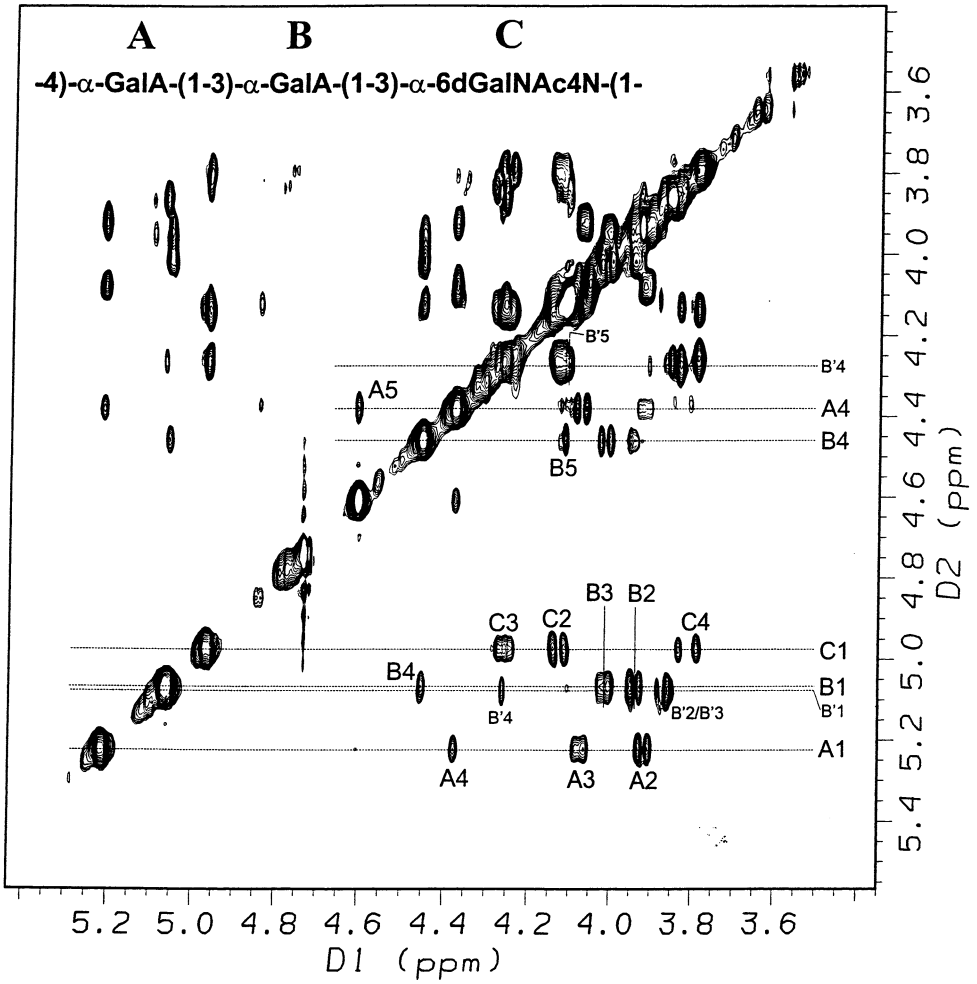

Fig. 2. 500-MHz TOCSY spectrum of the depolymerized *S. pneumoniae* type 1 CPS (**Spt1**), recorded at 27 °C with a mixing time of 100 ms. The letters refer to the residues in the upper left corner. The **A** H-1 peak is labeled A1, etc.

B′. If they are present, the resonances must be obscured by overlap with signals of other anomeric protons. β-Elimination by abstraction of the C-4 proton could lead to a product in which the α-glycosidic linkage of residue **C** is left intact, but we were not able to identify any NMR signals that could be assigned to an aglycone fragment of residue **A**.

Molecular modeling.—Energy contour maps of the glycosidic dihedral angles ϕ and ψ were calculated for the three possible disaccharides: **A**→**B**, **B**→**C**, and **C**→**A**. Whereas the GalA disaccharide showed a relatively wide low-energy area with four distinct minima in the conformational space with an energy up to 10 kcal/mol over the global minimum, the low-energy areas for the other two disaccharides containing 6dGalNAc4N were relatively restricted with two minima identified for **B**→**C** and three minima for **C**→**A**. This feature does not depend on the electrostatic component of the force field; exclusion of the electrostatics gave comparable results. The dihedral angles (ϕ, ψ) of the local minima for the three disaccharides were combined to give 24 different tetrasaccharides. The energy of each of the 24 possible tetrasaccharides was minimized. Any of the resulting tetrasaccharides whose sugar ring puckering was distorted from the $^4C_1$(D) chair was rejected since homonuclear ^{1}H coupling constants indicate all the chairs have normal puckering. This procedure produced eight distinct conformers with energies all within 8 kcal/mol (Table 2).

2D NOESY data at 27 °C with mixing times of 50, 100 and 200 ms (Table 3) provided strong cross peaks at the longer mixing times, while the 50-ms data were helpful in identifying effects of spin diffusion. Since the detailed size distribution of the oligosaccharide degradation products is not known, it is possible that some low-molecular-weight products could contribute positive NOEs that could interfere in analysis of the predominantly negative NOE observed in Fig. 4. Therefore we have also recorded a ROESY spectrum for the sample, in which all cross peaks have the same sign, and it is quite similar to the NOESY data. Table 4 gives internuclear distances for all the proposed models along

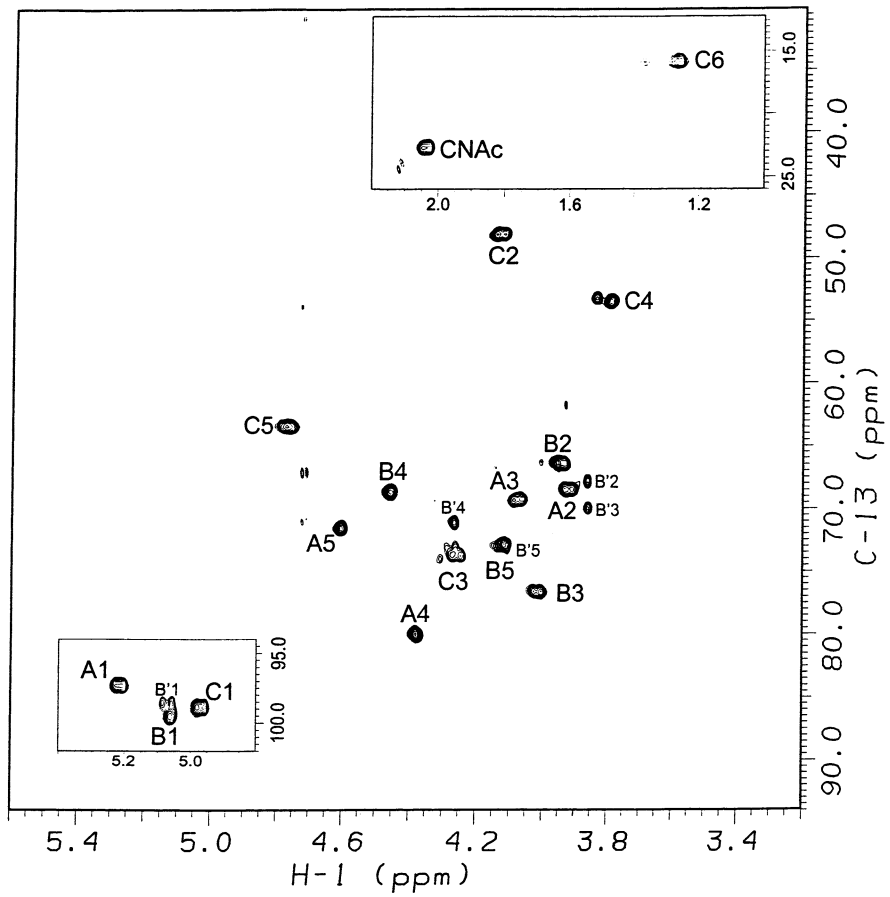

Fig. 3. 500-MHz ^{13}C-decoupled HMQC spectrum of the depolymerized *S. pneumoniae* type 1 CPS (**Spt1**) at 27 °C. The letters refer to the residues in Fig. 2. The **A** H-1–**A** C-1 cross peak is labeled A1, etc. The anomeric and methyl regions are inserts in the lower left and upper right corner, respectively.

with NOE estimates classified as weak, medium or strong. Table 4 shows that although conformer 1 is consistent with most of the NOE data, some cross peaks are not consistent with this conformation, suggesting that certain linkages must be flexible. For the 1→3 linkage between the two GalA residues, conformer 1 predicts strong NOEs between **A** H-1–**B** H-4 as well as for **A** H-1–**B** H-3, but several other conformers do as well. The cross peak between **A** H-1–**B** H-5 is likely to be mediated by the strong cross peak to **B** H-4 (three-spin effect) as is the case for the cross peaks between **B** H-1–**C** H-5 and **C** H-1–**A** H-5, since there are strong cross peaks to the respective H-4 in both cases. The cross peaks between **A** H-5–**B** H-2 and between **A** H-2–**B** H-4 are very weak at 50 ms mixing time and are likely to result from spin diffusion. For the linkage between residues **B** and **C**, we observe strong cross peaks between **B** H-1–**C** H-3 and **B** H-1–**C** H-4. In conformer 1, both distances are long, suggesting that other conformers such as 2, 4 or 5 must contribute. The cross peak between **B** H-5 and the amide methyl, although weak, persists at 50 ms mixing time and cannot arise from three-spin effects arguing that the conformation of the **B**→**C** linkage must contain contributions from conformers such as 4 and 5 in which these groups are closer. Therefore, our results indicate there must be some flexibility about this **B**→**C** linkage.

The linkage **C**→**A** has the same stereochemistry as that of galabiose (α-Gal-(1→4)-Gal) whose conformation has received substantial attention in the literature. The ϕ/ψ pair for conformer 1 (65.3°/−149.9°) resembles those found for the galabiose disaccharide in methyl β-galabioside, 81°/−135°,[9] and in globotriaosyl ceramide (Gb_3), 81°/−134°[10] and 80°/−136°.[11] The ϕ/ψ values for structure 1 also fit those of Gb_3 in complex with its bacterial receptor, as was elucidated by X-ray crystallography, 72.0°/−130.2°.[12] In the case of structure 2, the values for ϕ/ψ for the **C**→**A** linkage, 101.9°/85.0°, are similar to the values for a second conformer found by Cummings et al.,[11] 97°/83°. While our data are consistent with the conformation of the **C**→**A** linkage in conformer 1, they do not rule out contributions from other conformers.

The α-GalA-(1→3)-6dGalNAc4N and the α-GalA-(1→3)-GalA linkages have the same stereochemistry as

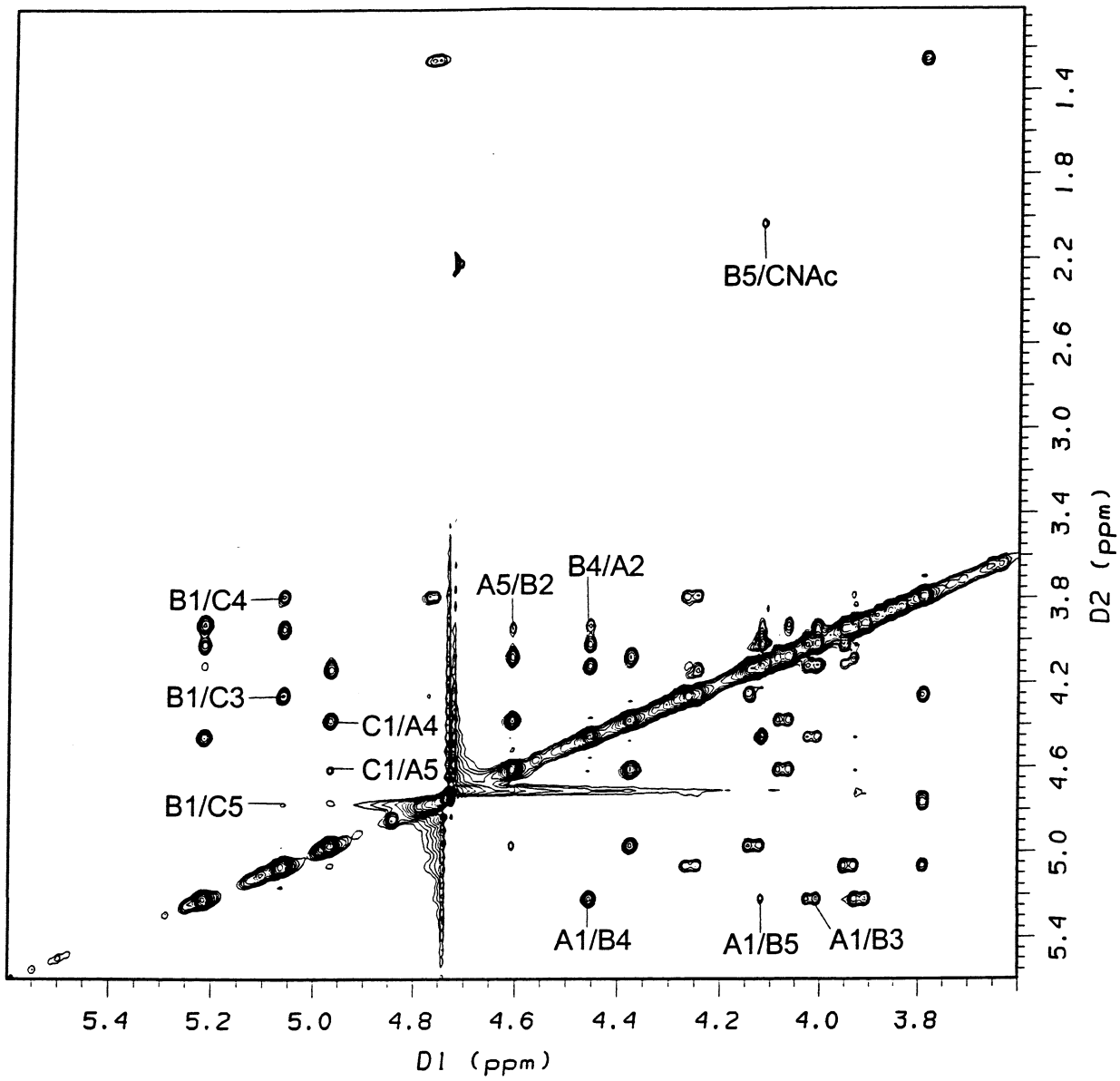

Fig. 4. 500-MHz NOESY spectrum of the depolymerized *S. pneumoniae* type 1 CPS, recorded at 27 °C with a mixing time of 200 ms. Indicated are the inter-residual peaks. The letters refer to the residues in Fig. 2. A1/B4 stands for the **A** H-1–**B** H-4 cross peak, etc.

the Gal-(α1→3)-Gal and the GalNAc-(α1→3)-Gal linkages found in the blood group B and A structures, which have been modeled and analyzed previously.[13,14] Otter et al.[14] have reported molecular modeling calculations for which the ϕ/ψ values (57.0°/−175.4°) fit both the X-ray and NMR data for α-Gal-(1→3)-Gal. These data are in agreement with the ϕ/ψ data for both the **A**→**B** and the **B**→**C** linkage in structure 2, but they are quite different from those in structure 1 (Table 2). HSEA modeling of the disaccharide α-GalA-(1→3)-6dGalNAc4N, which represents our **B**→**C** disaccharide leads to similar ϕ/ψ values of 78°/−152°.[8] We conclude that the *O*-deacetylated polysaccharide is most likely to have flexibility at the **B**→**C** linkage position, but may also exhibit internal motion at other linkages.

It is clear from the data of Fig. 1 that the presence of the *O*-acetyl groups at C-2 and C-3 of residue **A** strongly influences the conformation and dynamics of the polysaccharide, and they also influence the chemical shift of the C-6 methyl group of residue **A**. The models, which were built without acetate, do not provide any clear explanation of these effects. In the lowest energy conformation (number 1 in Table 2), the chain is folded so that residue **A** from one unit is close to that residue in the next repeating unit. If these conformations were present in the acetylated form of the polymer, the

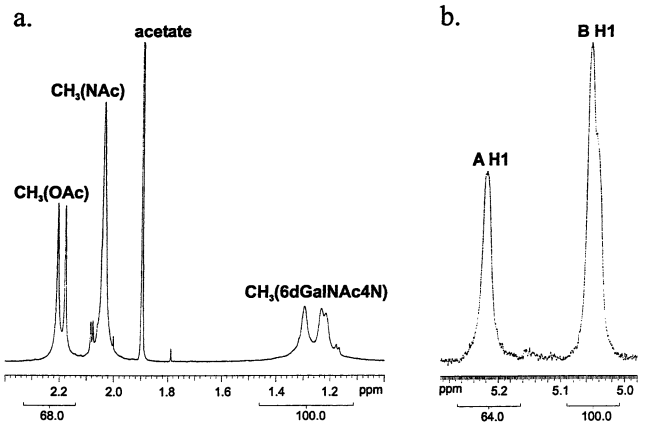

Fig. 5. 1D ^{1}H NMR spectrum expansions of (a) the methyl group region of the native polysaccharide; and (b) the anomeric region of the polysaccharide after periodate oxidation and borohydride reduction. The spectra were collected at 50 °C. Integration of the signals is shown under the axes.

acetates would interact. Conformation 3 shows a similar effect, while conformations 2 and 4 are extended in a more linear way. In none of these conformations does the C-6 methyl group of residue **C** interact with other parts of the model so it may be that the acetate groups in the native polysaccharide would have to be included in the model to describe these effects fully.

3. Experimental

Chemical treatments.—S. pneumoniae type 1 CPS (20 mg) was obtained from Merck Research Laboratories, dissolved in 20 mL of 80 mM NaAc (pH 4) containing 40 mM $NaIO_4$ and was incubated for 144 h at 4 °C in the dark. After reduction of excess periodate with 200 μL of ethylene glycol, the sample was dialyzed overnight against water at rt (dialysis tubing cutoff: 12–14 kDa). The oxidized CPS was lyophilized and subsequently incubated for 2 weeks at rt in 5 mL of 0.2 M NaOH. The resulting oligosaccharide fragments were desalted using a HiTrap desalting columns (Pharmacia, eluent: 5 mM NH_4HCO_3). Oligosaccharides were detected as a void-volume peak by monitoring the effluent at 205 nm.

In a second experiment, the native polysaccharide was oxidized with 100 mM $NaIO_4$ in 50 mM sodium citrate (pH 3.85) and was incubated for 16 h at 5 °C. The excess periodate was destroyed with ethylene glycol before reduction with sodium borohydride (200 mg/mL) for 4 h at rt and dialysis against 1 L of water at 4 °C. After 16 h, the water was replaced with fresh deionized water, and the sample was dialyzed for an additional 3 h. The 1D NMR spectrum was recorded at 50 °C.

O-Deacetylation of the polysaccharide, after cleavage of the contaminating C polysaccharide by HF digestion, was carried out by base hydrolysis with ~70 mM NaOH at 50 °C for 30 min. The sample was neutralized

Table 3
Measured inter-residual NOEs in oligosaccharide **Spt1** at different NOESY mixing times

NOE cross peaks [a]	Relative NOE intensities [b]		
	200 ms	100 ms	50 ms
A H-1–**B** H-4	95.1	88.2	100.0
A H-1–**B** H-3	70.6	59.5	49.6
A H-1–**B** H-5	23.6	12.4	10.5
A H-5–**B** H-2	41.6	19.0	3.8
A H-2–**B** H-4	50.5	22.2	3.1
B H-1–**C** H-3	83.5	70.9	77.2
B H-1–**C** H-4	49.5	46.3	51.7
B H-1–**C** H-5	22.6	12.4	21.1
B H-5–**C** NAc	26.2	19.7	9.3
C H-1–**A** H-4	100.0	100.0	99.1
C H-1–**A** H-5	21.4	22.8	18.0

[a] For residue labels, see Fig. 2.
[b] Highest NOE intensity per mixing time is set at 100.0.

Table 2
Glycosidic dihedral angles of tetrasaccharide models for **Spt1**

Structure	ϕ/ψ for **A**→**B** [a]	ϕ/ψ for **B**→**C**	ϕ/ψ for **C**→**A**	Energy [b] above minimum
1	104.3/−166.9	−59.3/−144.3	65.3/−149.9	0.00
2	50.8/−174.8	46.4/−173.4	101.9/85.0	0.80
3	53.7/−164.1	−36.6/−157.2	98.1/80.6	4.75
4	99.7/−164.4	52.1/−162.3	102.7/83.0	6.04
5	−48.4/−144.0	62.5/−160.6	100.0/85.4	6.24
6	73.6/−105.6	3.1/−173.2	99.1/84.2	6.88
7	102.4/−167.7	−24.4/−155.8	99.9/84.5	7.02
8	73.8/−116.9	56.9/−161.1	101.9/84.4	7.83

[a] For residue labels, see Fig. 2.
[b] In kcal/mol.

Table 4
Inter-residual NOEs in oligosaccharide **Spt1** and distances in the 8 CHARMM-generated tetrasaccharide models [a]

NOE cross peaks [b]	NOE intensities [c]	Distances (in Å)							
		1	2	3	4	5	6	7	8
A H-1–**B** H-4	S	2.14	2.30	2.23	2.17	3.82	3.36	2.16	3.07
A H-1–**B** H-3	M	2.44	3.16	3.01	2.47	3.62	2.19	2.46	2.26
A H-1–**B** H-5	W	3.99	4.61	4.37	4.12	5.65	4.37	4.06	4.43
A H-5–**B** H-2	W	3.92	4.18	3.96	3.99	4.84	3.08	4.18	3.01
A H-2–**B** H-4	W	3.91	4.70	4.66	4.05	5.18	5.03	3.97	4.87
B H-1–**C** H-3	S	3.58	3.22	3.66	3.00	2.83	3.65	3.67	2.90
B H-1–**C** H-4	M	3.91	2.44	3.58	2.36	2.25	3.12	3.50	2.28
B H-1–**C** H-5	W	5.64	4.70	5.58	4.54	4.40	5.28	5.56	4.44
B H-5–**C** NAc	W	6.41	3.27	5.81	2.51	2.65	5.25	5.48	2.47
C H-1–**A** H-4	S	2.76	3.64	3.64	3.64	3.65	3.65	3.66	3.64
C H-1–**A** H-5	W	4.04	3.85	3.87	3.85	3.86	3.86	3.88	3.86

[a] Numbered as in Table 2.
[b] See Fig. 2.
[c] S, strong; M, medium; W, weak.

and dialyzed to remove free acetate before it was lyophilized. The 2D NMR spectra were recorded on the sample in ^2H_2O with ~20 mM NaO^2H at 30 °C.

NMR analysis.—NMR samples were exchanged three times in ^2H_2O (99.9 atom% ^2H, Isotec, Inc.) with intermediate lyophilization, then dissolved in ^2H_2O (99.96 atom% ^2H, Isotec Inc). 1D ^1H NMR and 2D NMR spectra were recorded on a 500 MHz GE Omega PSG spectrometer with a Bruker 5-mm broadband triple-resonance probe or on a Varian Unity 600 MHz instrument. Experimental data were processed on a Silicon Graphics workstation using FELIX 98.0 software (Molecular Simulations, Inc.). The observed chemical shifts (δ) at 27 °C, expressed in ppm, are reported relative to internal acetone (^1H, 2.225 ppm; ^{13}C, 31.07 ppm). The 1D ^1H NMR of the native and of the *O*-deacetylated polysaccharide was recorded at 60 and at 50 °C, respectively, while the 1D and 2D NMR spectra of the oligosaccharide degradation product were collected at 27 °C.

A 2D ^1H–^1H correlation spectrum with solvent presaturation was obtained by the standard pulse sequence,[15] in which coherence is transferred through a double-quantum filter (DQF-COSY). The solvent suppression was performed by a 1.2-s presaturation during the relaxation delay. Phase-sensitive 2D ^1H–^1H-TOCSY (total correlation spectroscopy) spectra were recorded with MLEV-17 isotropic mixing sequence cycles of 15 and 100 ms using the pulse sequence by Bax and Davis.[16] NOE (nuclear Overhauser enhancement) data were obtained from standard 2D ^1H–^1H-NOESY spectra[17,18] with mixing times of 50, 100, and 200 ms. Solvent suppression was achieved by a 1.2-s presaturation during the relaxation delay. All 2D ^1H–^1H-spectra were acquired with data matrices of 256×1024 points representing a spectral width of 4000 Hz in each dimension.

A carbon-decoupled 2D ^1H–^{13}C-HMQC (heteronuclear multiple-quantum coherence) spectrum was obtained by a standard pulse sequence[19] using a WALTZ-16 sequence[20] for carbon decoupling with a field strength of 1295 Hz. The proton carrier frequency was set in the centre of the proton spectrum (4.0 ppm) with a spectral width of 4000 Hz. The ^{13}C carrier frequency was set at 60 ppm, in the centre of a carbon spectrum with a width of 12.5 kHz. A ^1H-detected HMBC (heteronuclear multiple-bond correlation) spectrum[21] was recorded without ^{13}C decoupling during acquisition. Delays of 3.2 ms $[1/(2 \times {}^1J_{CH})]$ and 68 ms $[1/(2 \times {}^nJ_{CH}) - 1/(2 \times {}^1J_{CH})]$ were used with a 1-s relaxation delay between acquisitions. The carrier frequencies and spectral widths were the same as with the HMQC. Both heteronuclear spectra were recorded with 256×1024 data points.

Molecular modeling.—Molecular modeling calculations were carried out with the CHARMM program,[22] version 22, using potential-energy functions and charges as modified for carbohydrates.[23] QUANTA97 software (Molecular Simulations, Inc.) was used for graphical manipulations of the models. All calculations were done in vacuum without inclusion of explicit water molecules and with a distance-dependent dielectric constant and a nonbonded interaction cutoff at 15 Å. The glycosidic dihedral angles ϕ and ψ are defined as O-5–C-1–O-1–C′-X and C-1–O-1–C′-X–C′-(X − 1), respectively, according to IUPAC/IUB convention. Energy contour maps of the dihedral angles ϕ and ψ were constructed for each of the three glycosidic linkages,

using a history-independent grid search with 5 or 10° per step with counter protons for acidic sugars. Each of the grid points was minimized using the steepsest decent method (300 steps), the conjugate gradient method (200 steps) and, finally, 300 steps of the ABNR (adopted basis-set Newton–Raphson) method. The low-energy disaccharide conformations were combined to provide unconstrained tetrasaccharide models (repeating unit plus the next GalA residue). These tetrasaccharides were energy minimized to relax all dihedral angles, and sterically distorted combinations were rejected. These energy minimizations were carried out with the conjugate gradient method, followed by the ABNR method (150 steps for both). The remaining tetrasaccharide conformers were systematically explored to determine which conformation(s) satisfied the inter-residual signals in the NOESY spectrum.

Acknowledgements

We thank Dr John Hennessey of Merck Research Laboratories for encouragement and for helpful discussions. This research was supported by NIH grant GM-57211.

References

1. World Health Organization position paper *Can. Commun. Dis. Rep.* **1999**, *25*, 150–151.
2. Jones, C. *Carbohydr. Eur.* **1998**, *21*, 10–16.
3. Kamerling, J. P. In *Pneumococcal Polysaccharides: A Chemical View*; Tomasz, A., Ed. Streptococcus pneumoniae; Mary Ann Liebert, Inc.: Larchmont, NY, 2000; pp. 81–114.
4. Lindberg, B.; Lindqvist, B.; Lönngren, J.; Powell, D. A. *Carbohydr. Res.* **1980**, *78*, 111–117.
5. Abeygunawardana, C.; Williams, T. C.; Sumner, J. S.; Hennessey, J. P., Jr. *Anal. Biochem.* **2000**, *279*, 226–240.
6. Xu, Q.; Bush, C. A. *Biochemistry* **1996**, *35*, 14512–14520.
7. Martin-Pastor, M.; Bush, C. A. *Biochemistry* **1999**, *38*, 8045–8055.
8. Bock, K.; Lönn, H.; Peters, T. *Carbohydr. Res.* **1990**, *198*, 375–380.
9. Bock, K.; Frejd, T.; Kihlberg, J.; Magnusson, G. *Carbohydr. Res.* **1988**, *176*, 253–270.
10. Nyholm, P.-G.; Magnusson, G.; Zheng, Z.; Norel, R.; Binnington-Boyd, B.; Lingwood, C. A. *Chem. Biol.* **1996**, *3*, 263–275.
11. Cummings, M. D.; Ling, H.; Armstrong, G. D.; Brunton, J. L.; Read, R. J. *Biochemistry* **1998**, *37*, 1789–1799.
12. Ling, H.; Boodhoo, A.; Hazes, B.; Cummings, M. D.; Armstrong, G. D.; Brunton, J. L.; Read, R. J. *Biochemistry* **1998**, *37*, 1777–1788.
13. Yan, Z.-Y.; Bush, C. A. *Biopolymers* **1990**, *29*, 799–811.
14. Otter, A.; Lemieux, R. U.; Ball, R. G.; Venot, A. P.; Hindsgaul, O.; Bundle, D. R. *Eur. J. Biochem.* **1999**, *259*, 295–303.
15. Rance, M.; Sorensen, O. W.; Bodenhausen, G.; Wagner, G.; Ernst, R. R.; Wuthrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *117*, 479–485.
16. Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355–360.
17. Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
18. Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95–117.
19. Bax, A.; Griffey, R. H.; Hawkins, B. L. *J. Magn. Reson.* **1983**, *55*, 301–315.
20. Shaka, A. J.; Keeler, J.; Freeman, R. *J. Magn. Reson.* **1983**, *53*, 313–340.
21. Bax, A.; Summers, M. F. *J. Am. Chem. Soc.* **1986**, *108*, 2093–2094.
22. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187–217.
23. Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohydr. Res.* **1988**, *180*, 207–221.