

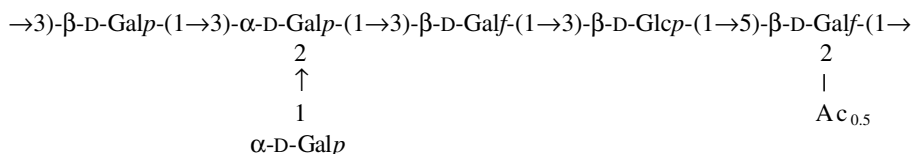
Full assignment of the ^1H and ^{13}C spectra and revision of the O-acetylation site of the capsular polysaccharide of *Streptococcus pneumoniae* Type 33F, a component of the current pneumococcal polysaccharide vaccine

Xavier Lemercinier and Christopher Jones*

Laboratory for Molecular Structure, National Institute for Biological Standards and Control Blanche Lane, South Mimms, Herts EN6 3QG, UK

Received 28 June 2005; received in revised form 16 October 2005; accepted 30 October 2005
Available online 17 November 2005

Abstract—The structure of the capsular polysaccharide from *Streptococcus pneumoniae* Type 33F was originally determined by a combination of chemical methods and limited use of NMR spectroscopy [*Can. J. Biochem. Cell Biol.* **1984**, *62*, 666–677]. We report full ^1H and ^{13}C assignments and confirm the structure of the saccharide repeat unit, but find that the site of O-acetylation is O-2 of the $\rightarrow 5$ - β -D-Galp, rather than the $\rightarrow 3$ - β -D-Galp residue. We find that a slightly higher percentage of the repeat units are O-acetylated:



© 2005 Elsevier Ltd. All rights reserved.

Keywords: Capsular polysaccharide; Pneumococcal; Vaccine

1. Introduction

Streptococcus pneumoniae remains an important bacterial pathogen causing both invasive and non-invasive diseases, mainly in children and the elderly.¹ The World

Health Organisation has estimated that pneumococcal infections cause in excess of 1 million deaths per year in the under-fives, mainly due to acute respiratory infections and in developing countries.² Ninety pneumococcal serotypes have been defined, differing in the structure of their capsular polysaccharide (CPS). The presence of an appropriate level of anti-CPS antibodies is generally considered protective against disease caused by that serotype, and immunological cross-protection between serotypes with CPSs of similar structure occurs. Immunisation of adults has relied on the use of multivalent vaccine comprising capsular polysaccharides (CPS) from the 23 most important serotypes of *S. pneumoniae*.³ Such vaccines, prepared with plain CPS, induce

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; COSY, correlation spectroscopy; CPS, capsular polysaccharide; Hib, *Haemophilus influenzae* type b; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; ROESY, rotating frame Overhauser effect spectroscopy; HSQC, heteronuclear single quantum correlation.

* Corresponding author. Tel.: +44 (0)1707 641211; fax: +44 (0)1707 646730; e-mail: cjones@nibsc.ac.uk

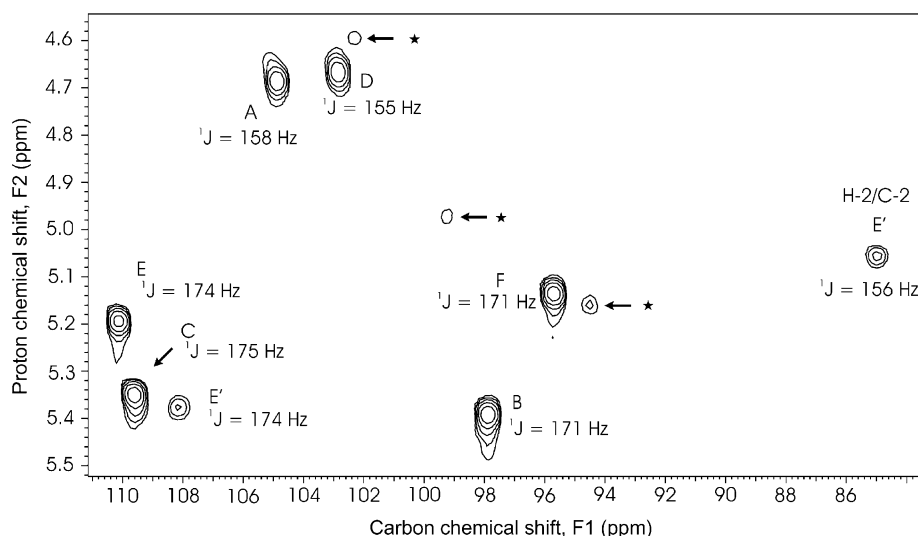


Figure 1. Part of the 500 MHz HSQC spectrum, obtained at 30 °C, showing the cross-peaks from the anomeric positions and from the O-acetylated position. Peaks arising from the Type 33F CPS residues are labelled with letters A–F as explained in the text, and the O-acetylated residue is labelled E'. The values of $^1J_{C-1,H-1}$, obtained from a spectrum obtained without ^{13}C decoupling, are shown for each cross-peak from the CPS. Cross-peaks arising from anomeric resonances of the teichoic acid pneumococcal C-polysaccharide are labelled with an asterisk.

Table 1. Pneumococcal Type 33F CPS 1H and ^{13}C NMR assignments

		$^1J_{H1,C1}$ (Hz)	H-1 C-1	H-2 C-2	H-3 C-3	H-4 C-4	H-5 C-5	H-6 C-6	H-6'	–OCOMe
→3)-β-D-Galp-(1→	A	158	4.69 104.93	3.70 71.08	3.72 82.21	4.09 69.76	3.73 75.85	3.82 62.16	3.84	
→2,3)-α-D-Galp-(1→	B	171	5.40 97.92	4.20 69.93	4.21 77.54	4.30 70.53	4.17 72.02	3.77 62.02	3.80	
→3)-β-D-Galf-(1→	C	155	5.36 109.74	4.44 80.53	4.15 86.41	4.33 82.98	4.00 71.16	3.70 64.33	3.69	
→3)-β-D-Glcp-(1→	D	175	4.68 102.90	3.50 74.58	3.69 83.29	3.50 69.30	3.50 76.74	3.96 61.99	3.96	
→5)-β-D-Galf-(1→	E	174	5.20 110.21	4.22 82.38	4.32 76.92	4.20 82.37	4.10 78.01	3.80 62.15	3.82	
→5)-β-D-Galf(2OAc)-(1→	E'	n.d.	5.39 108.20	5.06 85.01	4.51 76.14	4.29 83.85	n.d.	n.d.	n.d.	2.17 21.54
α-D-Galp-(1→	F	171	5.14 95.75	3.86 69.14	3.96 70.40	3.97 70.55	4.35 71.94	3.78 62.45	3.78	

Data were collected at 500 MHz and at a nominal probe temperature of 30 °C.

To minimise spontaneous de-O-acetylation, the spectra were first obtained at 30 °C: at this temperature, the 1H spectral lines were sometimes too broad to extract coupling constants less than 5 Hz. When higher resolution was needed or to take advantage of minor chemical shift changes induced by a temperature change, additional spectra were run at 70 °C.

For the β-pyranose residues, severe overlap prevented the values of $^3J_{H-2,H-3}$ and $^3J_{H-3,H-4}$ being obtained from the DQFCOSY spectrum. These couplings were instead obtained from an HSQC spectrum recorded with an acquisition time of 820 ms and at 70 °C. The chemical shifts and couplings indicate the presence of one β-Glcp

residue (spin system D δ_{H-3} 3.69 t, $J_{2,3}-J_{3,4} = 7.5$ Hz) and one β-Galp residue (spin system A δ_{H-3} 3.72 d, $J_{2,3} = 8.5$ Hz). This is confirmed by the observation of a strong H-3 to H-4 TOCSY cross-peak for spin system D and its absence for spin system A. The observed ^{13}C chemical shift of the C-3s (which are 8.5 ppm and 6.55 ppm downfield of the value observed for unsubstituted Gal and Glc^{19,20}) indicate glycosylation at O-3 for both residues.

1H and ^{13}C chemical shifts of the α-pyranose residues (spin systems B and F) correspond to Gal residues. In the case of the spin system B, a ^{13}C downfield shift of ~7 ppm relative to the expected value of 70.13 ppm

Table 2. Inter-residue correlations defining the sequence of sugar residues in the repeating unit of pneumococcal Type 33F CPS

Sugar residue and atom	NOESY cross-peaks						Long-range HSQC		HSQC-NOESY spectrum		HSQC-ROESY spectrum	
				δ (ppm)	Intensity ^a	Assign	δ (ppm)	Assign	δ (ppm)	Assign	δ (ppm)	Assign
→3)-β-D-Galp-(1→	A	4.70	H-1	4.21	n.r.	H-3 (B)			77.54	C-3 (B)		
		104.93	C-1				4.21	H-3 (B)	4.21	H-3 (B)	4.21	H-3 (B)
→2,3)-α-D-Galp-(1→	B	5.40	H-1	5.14	Strong	H-1 (F)	86.50	C-3 (C)	86.41	C-3 (C)	86.41	C-3 (C)
		5.40	H-1	3.86	Weak	H-2 (F)			95.75	C-1 (F)	95.75	C-1 (F)
		5.40	H-1	3.96	Weak	H-3 (F)						
		5.40	H-1	4.15	Strong	H-3 (C)						
		97.92	C-1						4.15	H-3 (C)		
		97.92	C-1						5.14	H-1 (F)		
→3)-β-D-Galp-(1→	C	5.37	H-1	3.70	Strong	H-3 (D)					83.29	C-3 (D)
		109.74	C-1				3.70	H-3 (D)				
→3)-β-D-Glcp-(1→	D	4.68	H-1	4.10	Medium	H-5 (E)						
		4.68	H-1	3.82	Medium	H-6 (E)						
		102.90	C-1								4.10	H-5 (E)
→5)-β-D-Galp-(1→	E	5.20	H-1	3.72	n.r.	H-3 (A)			82.22	C-3 (A)	82.21	C-3 (A)
		110.21	C-1						3.69	H-3 (A)	3.69	H-3 (A)
→5)-β-D-Galp(2OAc)-(1→	E'	5.39	H-1	3.75	n.r.	H-3 (A')						
→α-D-Galp-(1→	F	5.14	H-1	4.20	n.r.	H-2 (B)			97.92	C-1 (B)	97.92	C-1 (B)
		5.14	H-1	5.40	Strong	H-1 (F)			69.92	C-2 (B)		
		95.75	C-1						5.40	H-1 (B)		
		95.75	C-1						4.20	H-2 (B)		

^a n.r. = not resolved, because overlap of resonances from adjacent protons prevents discrimination of possible contributions.

for C-3 indicates a glycosylation at this position. Whilst the chemical shift of C-2 (69.93 ppm) corresponds to that of an unsubstituted position, data from NOESY, ROESY and long-range HSQC spectra confirm that this position is glycosylated (Table 2). Vicinal 2,3-disubstitution imposes strong steric hindrance that can restrict motion or alter bond angles compared to monosubstituted sugars, resulting in different glycosylation shift values.²¹ For spin system B, H2 and H3 have identical chemical shifts and same multiplicities at 30 °C, and an HSQC-TOCSY spectrum at 70 °C was necessary to distinguish them. The ¹³C chemical shifts of the other α-pyranose (spin system F) are all within 0.8 ppm of the expected values for an unsubstituted sugar,^{19,20} indicating a terminal side-chain residue. The ¹H and ¹³C chemical shifts of the two furanose residues correspond to galactose. Comparing the ¹³C chemical shifts of these two β-D-Galp clearly show different substitution patterns. In residue C, the C-3 resonance is ~10 ppm downfield and C-2 1.15 ppm upfield of the corresponding carbons in residue E, indicating a C-3 linkage. On the other hand, in residue E linkage at C-5 is indicated by a downfield shift of 6.8 ppm of the C-5 resonance and an upfield shift of 2.4 ppm for the C-6 resonance.

2.2. Inter-residue linkage positions

The sequence of linkages between the sugar residues were determined by combining data from a NOESY spectrum, HSQC-NOESY and HSQC-ROESY (Fig. 2)

spectra and long-range ¹H–¹³C correlations observed in an HSQC experiment tuned to observe small ¹H–¹³C coupling constants. These data are summarized in Table 2. These correlations establish unequivocally the connectivities between the six sugars residues and confirm the structure determined previously.¹⁷

2.3. Comparison with results previously published

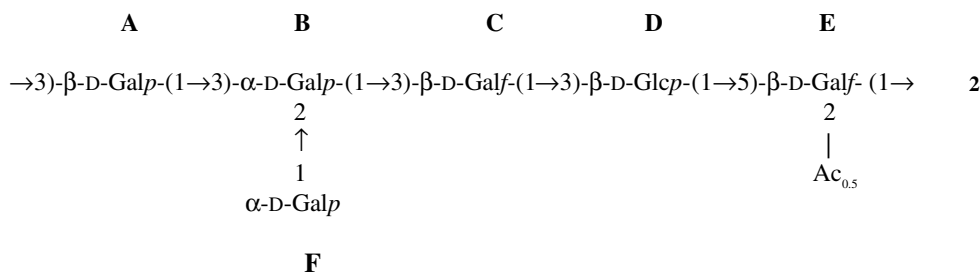
The structure obtained on the basis of our results is identical to that proposed previously for the de-O-acetylated CPS.¹⁷ In that paper, the authors give ¹³C and ¹H assignments (anomeric protons only) based on the 1D ¹H and ¹³C spectra of the native CPS and of the various oligomers obtained by degradation. Two-dimensional spectra, such as HSQC (see Fig. 1) highlight inconsistencies in their assignments; for example, the lowest field proton anomeric, arising from α-D-Galp(**B**), is assigned to one of the two β-D-Galp residues by Richards et al. Our ¹³C assignments agree with those of Richards et al.¹⁷ for both β-pyranosyl residues and the side-chain α-Galp residue. The most important differences arise for the ¹³C resonances **B** C-2, **C** C-2, **D** C-2 and **E** C-2.

2.4. O-Acetylation site

Richards et al. noted¹⁷ that a signal present at 85.1 ppm in the ¹³C spectra of native O-acetylated material is absent in spectra of de-O-acetylated material. The C-3 signal of non-O-acetylated →3)-β-D-Galp-(1→ residue,

at 86.5 ppm, is present in both cases but is more intense for the de-O-acetylated material. On the basis of peak intensities and because the chemical shift difference of 1.4 ppm between these signals corresponds to the upfield shift normally experienced by carbons β to an O-acetylated position,²² the authors tentatively assigned the peak at 85.1 ppm as C-3 of a putative C-2 O-acetylated $\rightarrow 3$ - β -D-Galp-(1 \rightarrow) residue. In our study, the HSQC spectrum shows, at a chemical shift typical of an anomeric signals of a β furanose residue, a third low intensity cross-peak (δ_{H} 5.39 ppm, δ_{C} 108.20 ppm, labelled E' in Fig. 1) which correlates in COSY and TOCSY spectra to a weak resonance at 5.06 ppm (δ_{C} 85.01 ppm in the HSQC spectrum). This low field proton chemical shift confirms that one of the furanose residues is O-acetylated at C-2. Correlations in DQFCOSY and HSQC spectra locate the H-3 and C-3 resonances of this residue at δ_{H} 4.51 and δ_{C} 76.14 ppm, and the H-4 at δ_{H} 4.29 ppm. These values are compatible only with the $\rightarrow 5$ - β -D-Galp-(1 \rightarrow) residue. The H-2 and C-2 resonances of the O-acetylated $\rightarrow 5$ - β -D-Galp-(1 \rightarrow) residue are 0.8 ppm and 2.4 ppm downfield of those of the non-O-acetylated residue, with a downfield shift of 0.2 ppm for H-3 and an upfield shift of approximately 0.8 ppm for C-3. These values are similar to those found in the literature for chemical shift modifications induced by O-acetylation.²² The NOESY spectrum contains a strong correlation between H-1 of the non-O-acetylated $\rightarrow 5$ - β -D-Galp-(1 \rightarrow) (E) and H-3 of $\rightarrow 3$ - β -D-Galp-(1 \rightarrow) (A), whilst the H-1 of the 2-O-acetylated $\rightarrow 5$ - β -D-Galp-(1 \rightarrow) correlates to a resonance with a slightly different chemical shift value, 0.04 ppm downfield, consistent with a secondary chemical shift on H-3 of $\rightarrow 3$ - β -D-Galp-(1 \rightarrow) adjacent residue. Such a situation has been previously observed, in the case of the meningococcal CPS.²³ Other shifted resonances of the $\rightarrow 3$ - β -D-Galp-(1 \rightarrow) aglycone of the 2-O-acetylated $\rightarrow 5$ - β -D-Galp-(1 \rightarrow) were not identified.

The following revised structure, **2**, for the repeating unit of the pneumococcal Type 33F CPS is therefore proposed:



Four samples of pneumococcal Type 33F CPS from different sources were analysed and all showed initial degrees of O-acetylation of approximately 50%. During our study, we noted that the spectra of material in solution at room temperature changed with time, suggesting

either migration or a loss of the O-acetyl group. O-Acetyl group migration has been observed in other CPSs,²³ and is known to happen easily for furanose residues.²⁴ As no evidence for O-acetyl migration was detected for the Type 33F CPS, and as the decrease in the intensities of the resonances from the C-2 O-acetylated $\rightarrow 5$ - β -D-Galp-(1 \rightarrow) matched increases in the intensities of the free acetate and non-O-acetylated $\rightarrow 5$ - β -D-Galp-(1 \rightarrow), facile loss of the O-acetyl group rather than migration seems likely (Fig. 2).

2.5. Identification of minor resonances

Pneumococcal teichoic acid (called pneumococcal C-polysaccharide) is an ubiquitous contaminant of pneumococcal CPS samples.^{18,25} Minor anomeric signals, identified by an asterisk in Figure 1, correspond to data published earlier for C-polysaccharide.¹⁸ Observation of a phosphocholine resonance at 3.2 ppm and the methyl group on the 6-deoxysugar at 1.25 ppm confirm this, whilst ³¹P resonances indicate the presence of only one phosphocholine substituent per repeat unit. The proportion of teichoic acid present in the samples was estimated from the relative intensities of the phosphocholine and four lowest field anomeric resonances of the CPS as approximately 7% (molar ratio of repeat units).

3. Experimental

Three CPS samples were material destined for vaccine production and the fourth was from the ATCC (American Type Culture Collection, Manassas, USA). Polysaccharide (~5 mg for ¹H spectra, ~20 mg for heteronuclear studies) was lyophilised twice from 400 μ L of deuterated water (>99.5% ²H, Apollo...) and redissolved in 700 μ L of deuterated water for analysis. NMR spectra were obtained on Varian Unity 500 or Inova 500 spectrometers using 5 mm indirect detection probes and running VNMR versions 6.1A, 6.1B

or 6.1C. Spectral widths were 6000 Hz (¹H) or 20,000 Hz (¹³C). The pulse programs used were those in the Varian library, except for the introduction of spin-echo sequences into the TOCSY and ROESY programs. The mixing times for the TOCSY²⁶ experiments

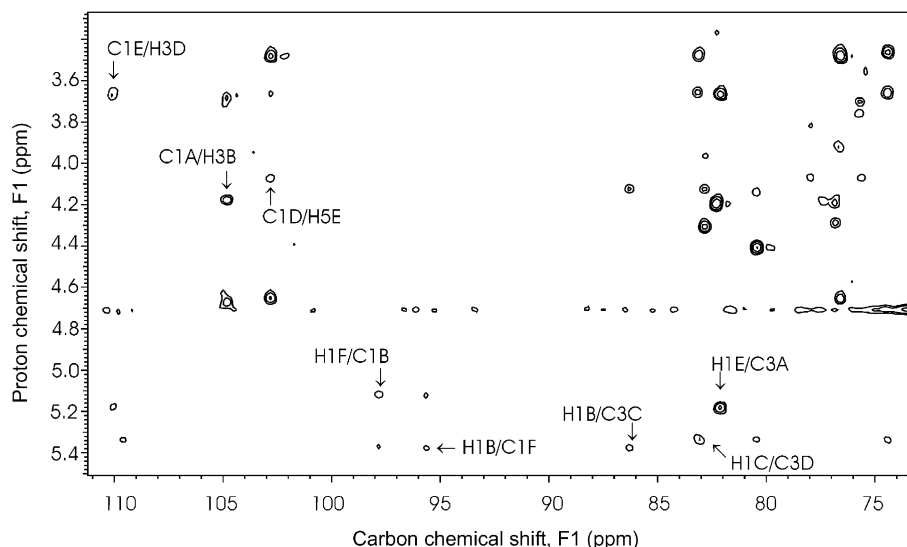


Figure 2. Part of the 500 MHz HSQC-ROESY spectrum, obtained at 30 °C with a 100 ms mixing time, showing trans-glycosidic correlations which were used to establish the sequence of and connectivities between the sugar residues. Residue labeling A–F is as explained in the text.

were 25 and 60 ms, 75 ms for the NOESY experiment,²⁷ and 100 ms for the ROESY experiment.²⁸ The HSQC sequence was that of Wider and Wüthrich,²⁹ with and without ¹³C decoupling. The HSQC-TOCSY and HSQC-ROESY sequences were variants of those of Crouch et al.,³⁰ modified to use the Wider and Wüthrich HSQC sequence. The mixing times in the HSQC-ROESY experiments were 25, 60 and 100 ms. Inter-residue ¹H–¹³C correlations were observed in ‘long range’ HSQC experiments with delays of 20 and 40 ms. Spectra were collected at an indicated probe temperature of 30 °C, apart from a limited set of 2D spectra which were obtained at 70 °C. ¹H and ¹³C chemical shifts are referenced to internal TSP-*d*₄ at zero (¹H) and –1.80 ppm (¹³C), respectively, and ³¹P to external 85% phosphoric acid at zero ppm.

Acknowledgements

We are grateful to Merck and Co. Inc. for the samples of polysaccharides, and Dr T. Frenkiel (N.I.M.R., Mill Hill) for the HSQC pulse sequence program.

References

- Fedson, D.; Anthony, J.; Scott, G. *Vaccine* **1999**, *17*, S11–S18.
- World Health Organisation *Weekly Epidemiological Record* **2003**, *78*, 110–119. Available at www.who.int/vaccines-documents/PP-WER/wer7423.pdf.
- Robbins, J. B.; Austrian, R.; Lee, C. J.; Rastogi, S. C.; Schiffman, G.; Henrichsen, J.; Mäkelä, P. H.; Broome, C. V.; Facklam, R. R.; Tiesjema, R. H.; Parke, J. C. *J. Infect. Dis.* **1983**, *148*, 1136–1159.
- Jacobson, R. M.; Poland, G. A. *Minerva Pediatr.* **2002**, *54*, 295–303.
- Fedson, D. *Vaccine* **1999**, *17*, S85–S90.
- Lockhart, S. *Expert Rev. Vaccines* **2003**, *2*, 633–648.
- Black, S.; Shinefield, H.; Fireman, B.; Lewis, E.; Ray, P.; Hansen, J. R.; Elvin, L.; Ensor, K. M.; Hackell, J.; Siber, G.; Malinoski, F.; Madore, D.; Chang, I.; Kohberger, R.; Watson, W.; Austrian, R.; Edwards, K.; Malinoski, F.; Madore, D.; Chang, I.; Kohberger, R.; Watson, W.; Austrian, R.; Edwards, K. *Pediatr. Infect. Dis. J.* **2000**, *19*, 187–195.
- Dagan, R.; Kayhty, H.; Wuorimaa, T.; Yaich, M.; Bailleux, F.; Zamir, O.; Eskola, J. *Pediatr. Infect. Dis. J.* **2004**, *23*, 91–98.
- Nurkka, A.; Joensuu, J.; Henckaerts, I.; Peeters, P.; Poolman, J.; Kilpi, T.; Käyhty, H. *Pediatr. Infect. Dis. J.* **2004**, *23*, 1008–1014.
- Miernyk, K. M.; Parkinson, A. J.; Rudolph, K. M.; Petersen, K. M.; Bulkow, L. R.; Greenberg, D. P.; Ward, J. I.; Brenneman, G.; Reid, R.; Santosham, M. *Clin. Infect. Dis.* **2000**, *31*, 34–41.
- Porat, N.; Barkai, G.; Jacobs, M. R.; Treffer, R.; Dagan, R. *J. Infect. Dis.* **2004**, *189*, 385–392.
- Biagini, R. E.; Schlottmann, S. A.; Sammons, D. L.; Smith, J. P.; Snawder, J. C.; Striley, C. A. F.; MacKenzie, B. A.; Weissman, D. N. *Clin. Diagn. Lab. Immunol.* **2003**, *10*, 744–750.
- European Pharmacopoeia* 5th Ed. EDQM, Strasbourg (2005).
- Lemercinier, X.; Jones, C. *Biologicals* **2000**, *28*, 175–183.
- Jones, C.; Lemercinier, X. *J. Pharmaceut. Biomed.* **2002**, *30*, 1233–1247.
- Lemercinier, X.; Martinez-Cabrera, I.; Jones, C. *Biologicals* **2000**, *28*, 17–24.
- Richards, J. C.; Perry, M. B.; Kniskern, P. J. *Can. J. Biochem. Cell Biol.* **1984**, *62*, 666–677.
- Karlsson, C.; Jansson, P. E.; Sørensen, U. B. S. *Eur. J. Biochem.* **1999**, *265*, 1091–1097.
- Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
- Jansson, P. E.; Kenne, L.; Widmalm, G. *Carbohydr. Res.* **1989**, *188*, 169–191.

21. Jansson, P. E.; Kenne, L.; Wehler, T. *Carbohydr. Res.* **1988**, *179*, 359–368.
22. Rutherford, T. J.; Jones, C.; Davies, D. B.; Elliott, A. C. *Carbohydr. Res.* **1991**, *218*, 175–184.
23. Lemercinier, X.; Jones, C. *Carbohydr. Res.* **1996**, *296*, 83–96.
24. Jackson, M.; Denu, J. *J. Biol. Chem.* **2002**, *277*, 18535–18544.
25. Vialle, S.; Sepulcri, P.; Dubayle, J.; Talaga, P. *Carbohydr. Res.* **2005**, *340*, 91–96.
26. Griesinger, C.; Otting, G.; Wüthrich, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1988**, *110*, 7870–7872.
27. Jeener, J.; Meier, B. H.; Bachman, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546–4553.
28. Bax, A.; Davis, D. G. *J. Magn. Res.* **1985**, *65*, 207–213.
29. Wider, G.; Wüthrich, K. *J. Magn. Res., Ser. B* **1993**, *102*, 239–241.
30. Crouch, R. C.; McFadyen, R. B.; Daluge, S. M.; Martin, G. E. *Magn. Res. Chem.* **1990**, *28*, 792–797.