ELSEVIER

Carbohydrate Research 302 (1997) 79–84

CARBOHYDRATE
RESEARCH

The structure of the capsular polysaccharide from *Klebsiella* type 52, using the computerised approach CASPER and NMR spectroscopy

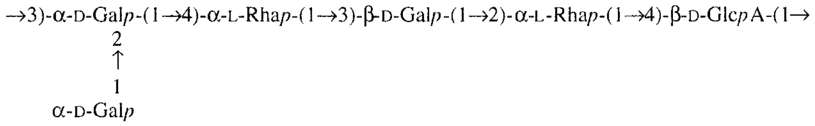
Roland Stenutz [a], Bertil Erbing [a], Göran Widmalm [a],
Per-Erik Jansson [b,*], Wolfgang Nimmich [c]

[a] *Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden*
[b] *Clinical Research Centre, Analytical unit, Karolinska Institute, Huddinge Hospital, NOVUM, S-141 86 Huddinge, Sweden*
[c] *Institut für Medizinische Mikrobiologie, Universität Rostock, D-18055 Rostock, Germany*

Received 11 November 1996; accepted 27 February 1997

Abstract

The structure of the capsular polysaccharide from *Klebsiella* type 52 has been elucidated using an improved and extended version of the computerised approach CASPER and NMR spectroscopy as principal methods. A previous suggestion to the structure but without the anomeric prefixes, could be shown correct [H. Björndal et al., *Carbohydr. Res.*, 31 (1973) 93–100]. The polysaccharide has a hexasaccharide repeat with the following structure:



$$\begin{array}{l} \rightarrow 3)\text{-}\alpha\text{-D-Gal}p\text{-}(1\rightarrow 4)\text{-}\alpha\text{-L-Rha}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rha}p\text{-}(1\rightarrow 4)\text{-}\beta\text{-D-Glc}p\text{A-}(1\rightarrow \\ \qquad\quad 2 \\ \qquad\quad \uparrow \\ \qquad\quad 1 \\ \quad \alpha\text{-D-Gal}p \end{array}$$

© 1997 Elsevier Science Ltd.

Keywords: *Klebsiella* type 52; CASPER; Capsular polysaccharide

1. Introduction

The capsular polysaccharides from *Klebsiella* have been fairly well documented structurally. Most of the suggestions have stood the test of time but some have been modified, e.g. K4 [1–3], K38 [4,5], and K82 [6,7]. Others have not been characterised completely. The structure of the capsular polysaccharide from *Klebsiella* type 52 (K52) has been published [8] but without anomeric configurations. We now report on the full structure of K52 using both a further extended version of CASPER [9], a computerised approach that interprets unassigned NMR data, and

---

[*] Corresponding author.

0008-6215/97/$17.00 © 1997 Elsevier Science Ltd. All rights reserved.
PII S0008-6215(97)00106-7

NMR spectroscopy using inter alia ^{1}H,^{1}H-NOEs and long range ^{13}C,^{1}H-correlations.

The CASPER program has been successfully used for oligo- and poly-saccharides with both known and unknown structures [10,11]. The program has now been extended so that it can interpret spectra which are not fully resolved or difficult to analyse because of excessively broadened and/or overlapped signals.

2. Results and discussion

Hydrolysis of K52 with trifluoroacetic acid gave galactose and rhamnose in the molar ratio of 3:2. On hydrolysis of a methylated sample methyl ethers corresponding to terminal galactose, 2- and 4-substituted rhamnose, 3-substituted galactose, and 2,3-disubstituted galactose were obtained. An additional ether, corresponding to 4-substituted glucuronic acid, was obtained when the methylated material was reduced with lithium triethyl borodeuteride before hydrolysis. The absolute configurations of the sugars were also determined [12,13] and were shown to be D for all sugars but rhamnose which was L. This is all in accordance with previous data, and strongly indicates that it is the same polysaccharide that was investigated.

The ^{1}H and ^{13}C NMR (Fig. 1) spectra of K52 were recorded together with the ^{13}C,^{1}H-correlation spectrum which was recorded both as the ^{1}H-detected, coupled and proton decoupled, and the ^{13}C-detected spectrum. These spectra were, even at high field, 600 MHz (for ^{1}H), not entirely resolved and also contained a number of weak broad signals. The elucidation with NMR spectroscopy was therefore far from straightforward.

The ^{1}H NMR spectrum showed a resolved anomeric region and signals for rhamnose methyl groups but apart from that only a crowded region for signals from ring protons. Signals in the anomeric region at δ 5.40, 5.20, 4.66, and 4.55 were clearly attributable to glucose/galactose-type sugar residues and signals at δ 5.11 and 5.06 to rhamnose residues as the latter had much smaller $J_{H\text{-}1,H\text{-}2}$ coupling constants. From the ^{13}C,^{1}H-correlation spectrum it was clear that the signals at δ 105.2, 104.1, 96.8, and 96.6 corresponded to the *gluco/galacto* residues and thesignals at δ 102.4 and 100.4 to the rhamnose residues. The ^{13}C and the ^{13}C,^{1}H-correlation spectra formed the basis for the input to the CASPER program.

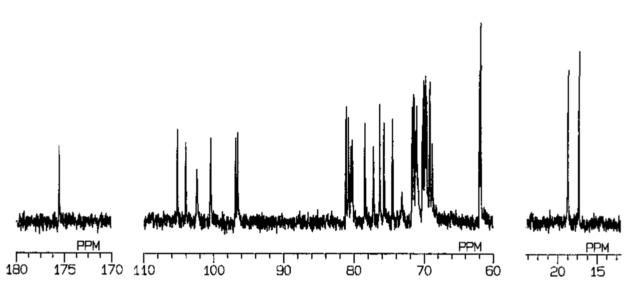
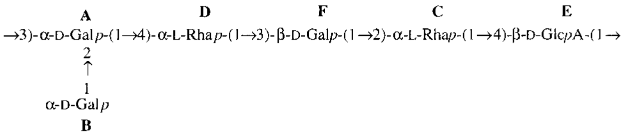


Fig. 1. The ^{13}C NMR spectrum of the capsular polysaccharide from *Klebsiella* K52.

The CASPER program works with data from sugar and methylation analysis in conjunction with unassigned ^{1}H or ^{13}C NMR data or the ^{13}C,^{1}H-correlation data, or a combination of these. The computer program generates, from the data on components and linkages, all possible repeating units or oligosaccharide structures by permutation of the constituent sugar residues, and anomeric configurations. If only structures that have compatible $^{1}J_{C\text{-}1,H\text{-}1}$ values and $^{3}J_{H\text{-}1,H\text{-}2}$ values are allowed, the number of possible structures is decreased. The NMR spectra for all structures are then calculated using monosaccharide chemical shifts to which disaccharide glycosylation shifts are added. If spectra of oligo- and poly-saccharides with branching at vicinal positions are simulated, correction values may have to be added [14]. The fit between experimental and simulated spectra, normally from a peak-by-peak comparison, is calculated and a ranking list according to fit is generated. The sum of the chemical shift differences between the experimental and the simulated spectrum is referred to as the deltasum.

Hitherto, comparison between spectra has required that all experimental signals are identified. This is not always easy or even possible and a fitting procedure without this requirement is desirable. The situation is the same for any of the spectra, the ^{1}H, the ^{13}C and the ^{13}C,^{1}H-correlated spectrum. We have now extended the program so that it can work with any number of experimental peaks. The fitting procedure developed is exhaustive, i.e. it compares the spectra in all possible ways and takes the fit with lowest deltasum. It has to be taken into consideration, however, that the fewer the signals, the lesser the difference between the deltasums. It should also be borne in mind that in general, the larger the oligosaccharide the smaller the differences in deltasums, unless restraints like J information is used. If reliable J information is available it is recommended that it is used.

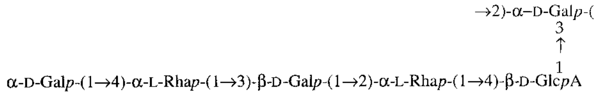
In the ^{13}C NMR spectrum of K52, 33 out of 36 signals were readily discernible. Two additional signals, at δ_C 69.8 and δ_C 71.1, were easily found by ^{13}C,^{1}H correlation spectroscopy. One signal was either overlapping or of very low intensity and was set to δ_C 69.8 and δ_H 4.03, as the corresponding H,C correlation peak was relatively strong. The input to the computer program was first only the 33 readily recognisable ^{13}C NMR signals and secondly, all tentatively identified ^{13}C NMR signals. Thirdly, the full ^{13}C,^{1}H-correlation spectrum was used. The residue information as obtained from methylation analysis was put in as usual. The results from all runs are given in Table 1 and Scheme 1. No large differences were found between the different simulated structures and therefore several simulated structures were taken into consideration. Shown in Scheme 1 are the four best structures, and in addition structure **5**, which is of interest as it only differs from structure **1** at the branching point and has a good fit in the ^{13}C,^{1}H-correlation simulation. The results show that already with 33 signals the correct structure is chosen (see below). The capacity to sort it out from the other structures is limited, however. In effect, if no more signals can be found, several structures on the list have to be considered as candidates and the differentiation based on other evidence. When using all ^{13}C NMR signals it turned out that three of the four first from the previous simulation were present but in another order. Also a higher ability to distinguish structure **1** from the others in terms of difference in deltasum were found when using 36 signals. The ^{13}C,^{1}H-correlated spectrum has a ranking order of structures, with the best preference for structural suggestion structure **1**. An apparent discrepancy between experimental and simulated chemical shifts for H-6 protons in the 3-substituted β-D-Gal residue is of unknown origin but affects all spectra alike. Interestingly, structure **5** has a good fit and therefore cannot be ruled out completely from the discussion. From the different runs with CASPER the following structure is thus concluded as the most likely alternative (for symbols **A**–**F** see below)



$$\underset{\substack{2\\ \uparrow \\ 1 \\ \alpha\text{-D-Gal}p \\ \text{B}}}{\overset{\text{A}}{\rightarrow 3)\text{-}\alpha\text{-D-Gal}p}}\text{-}(1\rightarrow 4)\text{-}\overset{\text{D}}{\alpha\text{-L-Rha}p}\text{-}(1\rightarrow 3)\text{-}\overset{\text{F}}{\beta\text{-D-Gal}p}\text{-}(1\rightarrow 2)\text{-}\overset{\text{C}}{\alpha\text{-L-Rha}p}\text{-}(1\rightarrow 4)\text{-}\overset{\text{E}}{\beta\text{-D-Glc}p\text{A}}\text{-}(1\rightarrow$$

1

and possible though less likely:



$$\alpha\text{-D-Gal}p\text{-}(1\rightarrow 4)\text{-}\alpha\text{-L-Rha}p\text{-}(1\rightarrow 3)\text{-}\beta\text{-D-Gal}p\text{-}(1\rightarrow 2)\text{-}\alpha\text{-L-Rha}p\text{-}(1\rightarrow 4)\text{-}\overset{\substack{\rightarrow 2)\text{-}\alpha\text{-D-Gal}p\text{-}(1\rightarrow \\ 3 \\ \uparrow \\ 1}}{\beta\text{-D-Glc}p\text{A}}$$

5

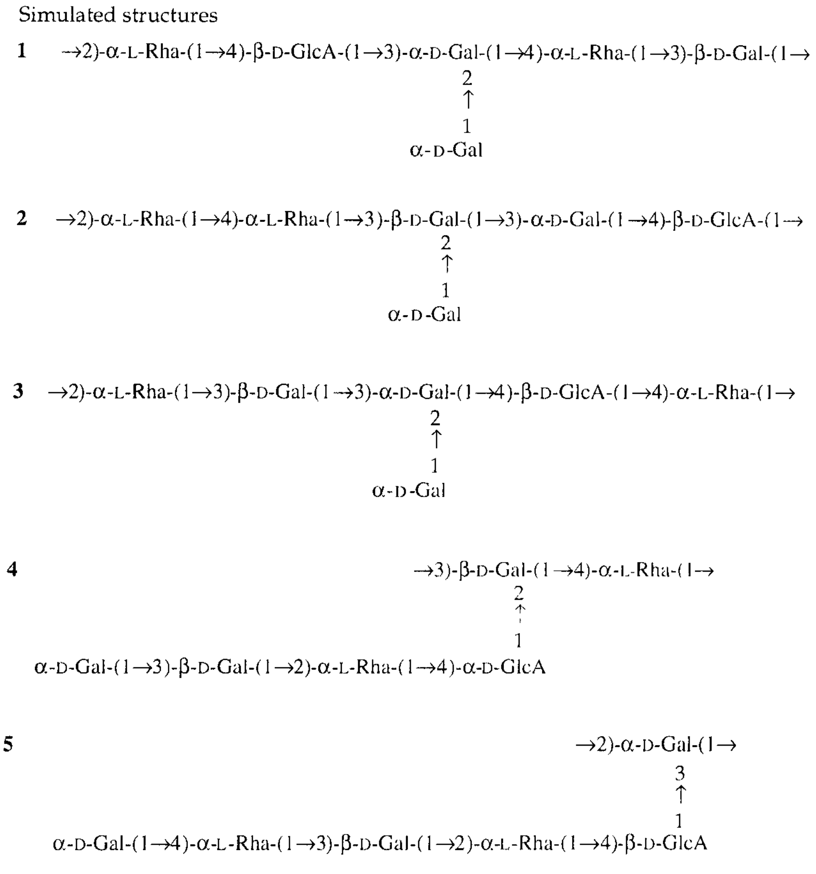
The structure of K52 was also ascertained by NMR methods and a differentiation between structure **1** and **5** could be made. Assignments of most of the signals in the ^{1}H and ^{13}C NMR spectra of K52 were obtained from a number of 2D NMR spectra and are given in Table 2. The residues are named **A**–**F**

Table 1
Simulated ^{13}C and ^{1}H NMR spectra of structure **1**

	H-1/C-1	H-2/C-2	H-3/C-3	H-4/C-4	H-5/C-5	H-6a/C-6	H-6b
A → 2,3)-α-D-Gal-(1 →	5.28	4.10	4.03	4.28	4.23	3.73	3.71
	97.6	73.5	78.3	69.7	71.4	61.5	
B α-D-Gal-(1 →	5.06	3.80	3.87	3.98	4.11	3.70	3.70
	97.8	69.1	70.1	70.1	71.8	61.7	
C → 2)-α-L-Rha-(1 →	5.13	4.03	3.86	3.52	3.99	1.32	
	100.6	80.5	71.1	73.6	69.9	17.4	
D → 4)-α-L-Rha-(1 →	5.09	3.97	4.01	3.53	4.00	1.38	
	102.4	71.8	69.9	82.1	69.0	17.8	
E → 4)-β-D-GlcA-(1 →	4.68	3.47	3.61	3.67	3.84		
	104.1	74.3	75.0	79.9	76.0	175.7	
F → 3)-β-D-Gal-(1 →	4.53	3.74	3.65	3.97	3.67	3.61	3.69
	105.5	71.6	81.5	69.6	76.0	61.6	

according to decreasing chemical shift of the anomeric proton signals. The glycosylation shifts, given in parentheses, were calculated by comparison with the respective monosaccharide. The two rhamnose residues were easily assigned from their H-1 and C-6 resonances and the fact that the signal for C-2 in one and C-4 in the other residue was shifted downfield to $\delta \sim 80$, made it possible to assign them to the 2- and 4-substituted rhamnose residue, named **C** and **D**, respectively. The residue with δ_{H-1} 5.40, (**A**), is assigned to the disubstituted galactosyl residue though it has only one large ^{13}C glycosylation shift, for the C-3 signal. It has been shown, however, that at vicinal disubstitution the glycosylation shifts may be small [14]. Furthermore, residue **F** has a negative glycosylation shift not compatible with disubstitution. Residue **A** also has significant ^{1}H NMR glycosylation shifts corroborating the assignment. The residue with δ_{H-1} 4.66 (**E**) is assigned to the 4-substituted glucuronic acid residue as the whole spin system is characterised by large $^{1}H,^{1}H$ couplings. There was also a significant downfield shift for the C-4 signal, in accord with 4-substitution. The remaining two residues are both galactose residues, one is terminal and one is 3-substituted, one α and one β. The glycosylation shift of the C-3 signal indicated **F** to be the 3-substituted Gal residue, and consequently **B** should be the terminal Gal residue.

NOESY experiments were employed to obtain sequential information. The cross-peaks from the



	(33 signals)		(36 signals)		(38 signals)
Structure	^{13}C deltasum	Ranking	^{13}C deltasum	Ranking	C,H deltasum
1	10.8	1	13.5	1	31.2
2	11.0	2	15.7	4	39.8
3	12.3	3	15.5	3	39.5
4	12.5	4	17.7	17	44.5
5	14.0	43	16.8	7	33.7

Scheme 1. The results of a simulation with the ^{13}C NMR spectrum of K52 using first 33 then 36 signals, and finally the $^{13}C,^{1}H$-correlation spectrum. The list shows the four best suggestions and number **5**, which is of interest as it differs only at the branch point and has a high rank in the C,H-correlation list.

Table 2
^{1}H and ^{13}C NMR chemical shifts for the *Klebsiella* K52 CPS

Sugar residue	Chemical shift (δ)					
	H-1/C-1	H-2/C-2	H-3/C-3	H-4/C-4	H-5/C-5	H-6/C-6
→ 2,3)-α-D-Gal*p*-(1 →	5.40	4.23	4.14	4.29 [b]	4.35 [b]	~ 3.76 [b]
A	(0.18) [a]	(0.45)	(0.33)	(0.34)	(0.32)	
	96.6 [173] [c]	71.1	78.4	69.8	71.7	61.8
	(3.4)	(1.8)	(8.3)	(−0.48)	(0.4)	(−0.2)
α-D-Gal*p*-(1 →	5.20	3.87	3.95	4.03 [b]	4.34 [b]	~ 3.72 [b]
B	(−0.02)	(0.09)	(0.14)	(0.08)	(0.31)	
	96.8 [174]	69.2	70.1	69.8	71.5	62.0
	(3.6)	(−0.2)	(0.0)	(−0.5)	(0.2)	(0.0)
→ 2)-α-L-Rha*p*-(1 →	5.11	4.06	3.86	3.51	4.02	1.28
C	(−0.01)	(0.14)	(0.05)	(0.06)	(0.16)	(0)
	100.4 [176]	80.8	71.1	73.2	69.6	17.3
	(5.6)	(9.0)	(0.1)	(0.0)	(0.5)	(0.4)
→ 4)-α-L-Rha*p*-(1 →	5.06	4.09	4.02 [b]	3.68	3.98	1.40
D	(−0.06)	(0.17)	(0.21)	(0.23)	(0.12)	(0.12)
	102.4 [174]	71.2	70.2	80.4	68.8	18.7
	(7.6)	(−0.6)	(−0.8)	(7.2)	(−0.3)	(1.0)
→ 4)-β-D-Glc*p*A-(1 →	4.66	3.44	3.59	3.64	3.76	–
E	(0.01)	(0.04)	(0.07)	(0.10)	(0.04)	
	104.1 [164]	74.5	75.7	80.2	77.2	175.6
	(7.3)	(−0.5)	(−0.8)	(7.5)	(0.3)	(0.0)
→ 3)-β-D-Gal*p*-(1 →	4.55	3.72	3.73	4.03 [b]	3.73 [b]	~ 3.87 [b]
F	(0.02)	(0.27)	(0.14)	(0.14)	(0.08)	
	105.2 [163]	71.5	81.1	69.8	76.3	62.0
	(7.8)	(−1.5)	(7.3)	(0.1)	(0.4)	(0.0)

[a] Chemical shift differences compared to monomers.
[b] Tentative assignments.
[c] Values for $^{1}J_{C,H}$ in square brackets.

anomeric protons were examined for inter- and intra-residue information and the results are given in Table 3. The results are all in agreement with six disaccharide elements, which are in accord with structure **1**. The NOEs observed for the **A**–**D** and **B**–**A** elements should not be present in structure **5** and it can therefore be ruled out.

Long-range $^{13}C,^{1}H$-correlations obtained from an HMBC spectrum (Table 3) corroborated the assigned disaccharide elements deduced from the NOESY experiment. Thus, inter-residue correlations from anomeric protons, giving sequential information, were observed for all six disaccharide elements. Correlations from anomeric carbons were also useful in further giving information, for three of the six disaccharide elements. A correlation **A**–**D** is in support for structure **1** but not for **5**. It can thus be concluded that structure **1** should be the correct one for K52.

Table 3
Interglycosidic NOE and HMBC contacts from anomeric protons and carbons observed for the *Klebsiella* K52 CPS

Residue	$\delta_{H-1}/\delta_{C-1}$	NOE contacts	HMBC contacts
A	5.40	3.68 (H-4, **D**)	80.4 (C-4, **D**)
	96.6		3.68 (H-4, **D**)
B	5.20	5.40 (H-1, **A**)	71.1 (C-2, **A**)
		4.23 (H-2, **A**)	
C	5.11	3.64 (H-4, **E**)	80.2 (C-4, **E**)
	100.4		3.64 (H-4, **E**)
D	5.06	3.73 (H-3, **F**)	81.1 (C-3, **F**)
E	4.66	4.14 (H-3, **A**)	78.4 (C-3, **A**)
F	4.55	4.06 (H-2, **C**)	80.8 (C-2, **C**)
	105.2		4.06 (H-2, **C**)

3. Experimental

Bacterial strains.—Strain 5759/50 was obtained from the type culture collection at State Serum Institute, Copenhagen, Denmark.

General methods.—Concentrations were performed under diminished pressure at < 40 °C or under a stream of air or N_2. For GLC, a Hewlett-Packard 5890 instrument fitted with a flame-ionisation detector was used. GLC–EIMS was performed on a Hewlett-Packard 5970 MSD instrument.

Alditol acetates and partially methylated alditol acetates were analysed on an HP-5 capillary column (25 m × 0.20 mm) using the temperature program 180 °C (1 min) → 250 °C at 3 °C/min. Analysis of the acetylated (+)-2-butyl glycosides [12,13] were performed on the same column but the temperature program 130 °C (1 min) → 220 °C at 3 °C/min was used.

Computer analysis with CASPER.—The original CASPER program, written in Pascal, was further enhanced and rewritten in C for use on any unix workstation. All simulation and plotting modules have been implemented. The calculations with chemical shifts and glycosylation shifts were implemented as previously described. The data base which previously was quadratic, i.e. grew quadratically on inclusion of new monomers, has been made compact through the use of cross references.

NMR spectroscopy.—NMR spectra of solutions in D_2O were recorded at 70 °C using either a JEOL GSX-270 or a Varian Unity Plus 600 instrument. Chemical shifts are reported in ppm, using sodium 3-trimethylsilylpropanoate-d_4 (TSP, δ_H 0.00) or acetone (δ_C 31.0), as internal references. ^{1}H,^{1}H-COSY, relayed and double relayed ^{1}H,^{1}H-COSY, NOESY, and ^{13}C,^{1}H-COSY were performed using standard pulse-sequences. The mixing time in the NOESY experiment was 250 ms. The $^1J_{C\text{-}1,H\text{-}1}$ values were determined from a coupled inverse-detected HSQC spectrum, and the ^{1}H,^{13}C long-range couplings were investigated with an inverse detected HMBC experiment using a delay time of 60 ms.

Sugar and methylation analysis.—Native and methylated K52 were hydrolysed by treatment with 4 M HCl at 100 °C for 2 h. The sugars in the hydrolysates were converted into alditol acetates and partially methylated alditol acetates by conventional means. The methylated polysaccharide in THF was carboxyl reduced by treatment with lithium triethyl borodeuteride in tetrahydrofuran (1 M) at 0 °C for 1 h.

Acknowledgements

This work was supported by grants from the Swedish Natural Science Research Council, and the Swedish Research Council for Engineering Sciences. The skilled technical assistance of Mrs Helena Liedgren is gratefully acknowledged.

References

[1] E.H. Merrifield and A.M. Stephen, *Carbohydr. Res.*, 96 (1981) 113–120.

[2] Y.A. Knirel, V.A. Mamontova, N.A. Kocharova, A.S. Shashkov, T.F. Solov'eva, and N.K. Kochetkov, *Bioorg. Khim.*, 14 (1988) 1583–1585.

[3] Y.A. Knirel, N.A. Kocharova, A.S. Shashkov, N.K. Kochetkov, V.A. Mamontova, and T.F. Solov'eva, *Carbohydr. Res.*, 188 (1989) 145–155.

[4] B. Lindberg, K. Samuelsson, and W. Nimmich, *Carbohydr. Res.*, 30 (1973) 63–70.

[5] P.-E. Jansson, B. Lindberg, M.C. Manca, W. Nimmich, and G. Widmalm, *Carbohydr. Res.*, 261 (1994) 111–118.

[6] G. Dutton and A. Lim, *Carbohydr. Res.*, 123 (1983) 247–257.

[7] P.-E. Jansson, B. Lindberg, G. Widmalm, G. Dutton, A. Lim, and I. Sutherland, *Carbohydr. Res.*, 175 (1988) 103–109.

[8] H. Björndal, B. Lindberg, J. Lönngren, M. Meszaros, and J. Thompson, *Carbohydr. Res.*, 31 (1973) 93–100.

[9] P.-E. Jansson, L. Kenne, and G. Widmalm, *Pure Appl. Chem.*, 61 (1989) 1181–1192.

[10] K. Hermansson, P.-E. Jansson, L. Kenne, G. Widmalm, and F. Lindh, *Carbohydr. Res.*, 235 (1992) 69–81.

[11] A. Johansson, P.-E. Jansson, and G. Widmalm, *Carbohydr. Res.*, 253 (1994) 317–322.

[12] K. Leontein, B. Lindberg, and J. Lönngren, *Carbohydr. Res.*, 62 (1978) 359–362.

[13] G.J. Gerwig, J.P. Kamerling, and J.F.G. Vliegenthart, *Carbohydr. Res.*, 77 (1979) 1–7.

[14] H. Baumann, B. Erbing, P.-E. Jansson, and L. Kenne, *J. Chem. Soc., Perkin Trans. 1*, (1989) 2153–2165.