Available online at www.sciencedirect.com
ScienceDirect
Carbohydrate Research 343 (2008) 977–981
Carbohydrate RESEARCH

Note

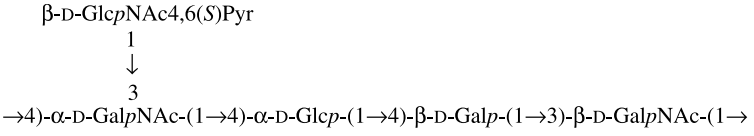
The O-polysaccharide of *Escherichia coli* O112ac has the same structure as that of *Shigella dysenteriae* type 2 but is devoid of O-acetylation: a revision of the *S. dysenteriae* type 2 O-polysaccharide structure

Andrei V. Perepelov,[a,*] Andrej Weintraub,[b] Bin Liu,[c,d] Sof'ya N. Senchenkova,[a] Alexander S. Shashkov,[a] Lu Feng,[c,d] Lei Wang,[c,d,e,f] Göran Widmalm[g] and Yuriy A. Knirel[a]

[a]*N. D. Zelinsky Institute of Organic Chemistry, Russian Academy of Sciences, 119991 Moscow, Russian Federation*
[b]*Karolinska Institute, Department of Laboratory Medicine, Division of Clinical Microbiology, Karolinska University Hospital, S-141 86 Stockholm, Sweden*
[c]*TEDA School of Biological Sciences and Biotechnology, Nankai University, TEDA, Tianjin 300457, PR China*
[d]*Tianjin Key Laboratory for Microbial Functional Genomics, TEDA College, Nankai University, TEDA, Tianjin 300457, PR China*
[e]*The Engineering and Research Center for Microbial Functional Genomics and Detection Technology, Ministry of Education, PR China*
[f]*The Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, PR China*
[g]*Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden*

Received 28 December 2007; received in revised form 1 February 2008; accepted 4 February 2008
Available online 10 February 2008

Abstract—The O-antigen structure of *Shigella dysenteriae* type 2 was reinvestigated using chemical modifications along with high-resolution 2D ^{1}H and ^{13}C NMR spectroscopy. The O-antigen was found to contain a pyruvic acid acetal, which was overlooked in an early study, and the following revised structure of the pentasaccharide repeating unit was established:



where ~70% GlcNAc residues bear an *O*-acetyl group at position 3. The O-antigen of *Escherichia coli* O112ac was found to have the same carbohydrate structure but to lack O-acetylation.

Keywords: *Escherichia coli*; *Shigella dysenteriae*; O-Antigen; Bacterial polysaccharide structure; Pyruvic acid acetal

Shigella are human pathogens that cause bacilliary dysentery (shigellosis) and diarrhea, continuing to threaten public health mainly in less developed countries with conditions of poor sanitation. Although bacteria *Shigella* are closely related to *Escherichia coli* and both are in fact one species, *Shigella* strains are put into their own genus and subgrouped into four species: *Shigella boydii*, *Shigella dysenteriae*, *Shigella flexneri*, and *Shigella sonnei*. Based on structure variation in their O-antigens, the first three species are typed into multiple serotypes. The existence of multiple O-antigen forms within a species can confer selective advantage to

* Corresponding author. Tel.: +7 495 1376148; fax: +7 499 1355328; e-mail: perepel@ioc.ac.ru

0008-6215/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2008.02.002

bacteria for adaptation in their econiches.[1] More than half of the O-antigen forms of *Shigella* are shared with *E. coli* strains.[2]

The structure elucidation of *Shigella* O-antigens (O-polysaccharides) has a long history and has been completed only recently (for a survey see Ref. 2). A number of the O-antigen structures, for which older methods were used, have been re-examined using modern NMR techniques and some of them revised.[3–7]

In 1977, the structure of *S. dysenteriae* type 2 O-antigen was established[8] (Chart 1). In the present work, this structure has been revised and close relationships of the O-antigens of *S. dysenteriae* type 2 and *E. coli* O112ac are demonstrated.

The lipopolysaccharide (LPS) of *S. dysenteriae* type 2 was isolated from dried bacterial cells by the phenol–water procedure and degraded with dilute acetic acid to give a high-molecular-mass O-polysaccharide (OPS), which was separated from low-molecular-mass materials by GPC on Sephadex G-50. The ^{13}C NMR spectrum indicated an irregularity of the OPS, probably owing to the presence of pyruvic acid acetal (Pyr) and *O*-acetyl groups in non-stoichiometric quantities (minor signals at δ 25.6 and 21.5). Therefore, LPS and OPS were treated with aq ammonia to give O-deacylated polymers LPSOH and DPS, respectively.

The ^{1}H NMR spectrum of LPSOH contained, inter alia, signals for five anomeric protons at δ 4.56–5.30, a methyl group of Pyr at δ 1.44, three *N*-acetyl groups at δ 2.00–2.02. The ^{13}C NMR spectrum of LPSOH (Fig. 1A) showed signals for five anomeric carbons in the region δ 99.7–106.6, five hydroxymethylene groups (C-6 of hexoses) at δ 61.3–62.2 and 65.8 (data of a DEPT experiment), three nitrogen-bearing carbons (C-2 of amino sugars) at δ 50.0, 52.5, and 57.0, 17 oxygen-bearing non-anomeric sugar ring carbons in the region δ 67.1–81.8, methyl groups of Pyr (C-3) at δ 26.1 and three *N*-acetyl groups at δ 23.6–23.9, CO groups of Pyr (C-1) and *N*-acetyl groups at δ 174.9–176.7. The position of the acetal carbon (C-2) and the carboxyl group (C-1) of Pyr was found following the correlations of Pyr H-3,C-2 and H-3,C-1 in the HMBC spectrum at δ 1.44/102.4 and 1.44/175.8, respectively.

The ^{1}H and ^{13}C NMR spectra of LPSOH (Table 1) were assigned using 2D ^{1}H,^{1}H COSY, TOCSY, ROESY, ^{1}H,^{13}C HSQC, and HMBC experiments. The ^{1}H,^{1}H experiments revealed five spin systems for residues of β-Glc*p*NAc (**A**), α-Gal*p*NAc (**B**), α-Glc*p* (**C**), β-Gal*p* (**D**), and β-Gal*p*NAc (**E**) (the sugars were denoted according to their sequence in the repeating unit shown below). The anomeric configurations of the sugar residues were ascribed based on $J_{1,2}$ coupling constant

α-D-Glc*p*NAc3/4Ac
1
↓
3
→4)-α-D-Gal*p*NAc-(1→4)-α-D-Glc*p*-(1→4)-β-D-Gal*p*-(1→3)-α-D-Gal*p*NAc-(1→

Chart 1. Structure of the O-polysaccharide of *S. dysenteriae* type 2[8].

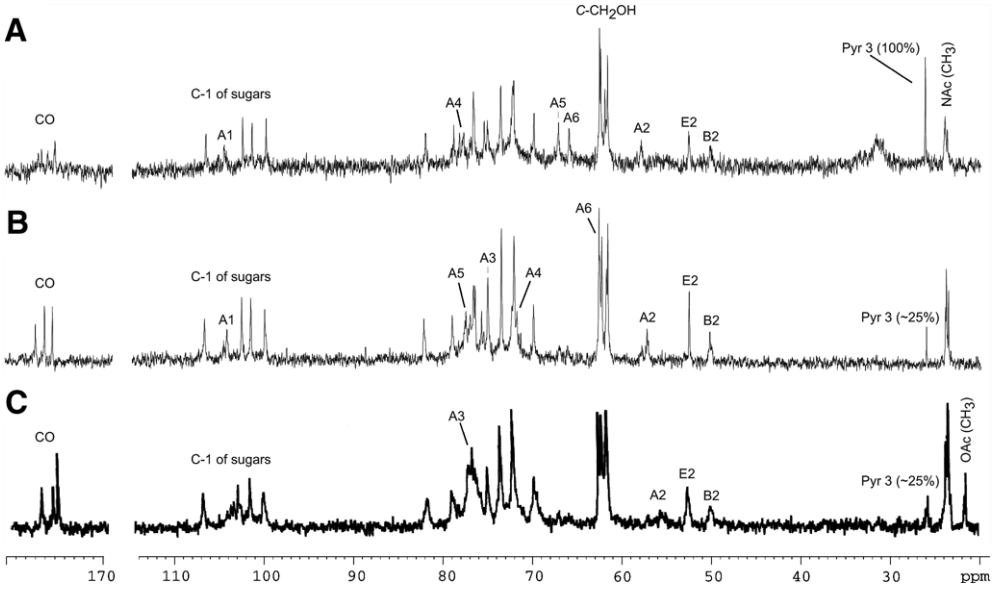


Figure 1. ^{13}C NMR spectrum of LPSOH (A), DPS (B), and OPS (C) from *S. dysenteriae* type 2. Arabic numerals refer to atoms in sugar residues denoted by letters as shown in Table 1.

Table 1. ^{1}H and ^{13}C NMR chemical shifts of LPSOH and OPS from *S. dysenteriae* type 2 (δ, ppm)

Sugar residue	Nucleus	1	2	3	4	5	6 (6a, 6b)[a]
LPSOH[b]							
β-D-Glc*p*NAc-(1→	^{1}H	4.62	3.83	3.68	3.44	3.42	3.64; 4.02
A	^{13}C	104.2	57.0	72.3	78.1	67.1	65.8
→3,4)-α-D-Gal*p*NAc-(1→	^{1}H	5.30	4.14	3.86	4.29	3.96	
B	^{13}C	99.7	50.0	77.6	75.3	72.3	
→4)-α-D-Glc*p*-(1→	^{1}H	4.86	3.47	3.87	3.60	4.18	
C	^{13}C	101.3	73.5	74.9	76.7	72.0	
→4)-β-D-Gal*p*-(1→	^{1}H	4.56	3.52	3.70	3.98	3.73	
D	^{13}C	106.6	72.0	73.4	78.6	76.5	
→3)-β-D-Gal*p*NAc-(1→	^{1}H	4.81	4.08	3.84	4.09	3.54	
E	^{13}C	102.3	52.5	81.8	69.9	76.6	
Pyr	^{1}H			1.44			
	^{13}C	175.8	102.4	26.1			
DPS[c]							
β-D-Glc*p*NAc-(1→	^{1}H	4.55	3.80	3.55	3.42	3.43	3.78; 3.93
A	^{13}C	104.1	57.2	75.1	71.6	77.3	62.5
→3,4)-α-D-Gal*p*NAc-(1→	^{1}H	5.34	4.20	3.97	4.45	4.05	3.72; 3.80
B	^{13}C	99.7	50.1	77.5	75.7	72.7	61.5
→4)-α-D-Glc*p*-(1→	^{1}H	4.89	3.51	3.92	3.64	4.20	3.68; 3.88
C	^{13}C	101.4	73.4	74.9	76.8	72.0	62.2
→4)-β-D-Gal*p*-(1→	^{1}H	4.59	3.56	3.73	4.01	3.68	
D	^{13}C	106.5	72.1	73.5	79.0	76.4	
→3)-β-D-Gal*p*NAc-(1→	^{1}H	4.96	4.14	3.88	4.14	3.74	
E	^{13}C	102.4	52.5	82.1	69.8	76.5	

[a] H-6/C-6 correlations of units **B**–**E** in LPSOH and **D**–**E** in DPS are not-resolved in the region δ_H/δ_C 3.65–3.87/61.3–62.2.
[b] Chemical shifts for the *N*-acetyl groups are δ_H 2.00–2.02, δ_C 23.6–23.9 (Me) and 174.9–176.7 (CO).
[c] Chemical shifts for *N*-acetyl groups are δ_H 2.04–2.07, δ_C 23.5–23.9 (Me) and 175–176.7 (CO).

values (<4 Hz for α-linked and >6 Hz for β-linked monosaccharides). The absence of any ^{13}C NMR signals in the region δ 83–88 demonstrated the pyranose form of all sugar residues. The signals for C-3 and C-4 of unit **B**, C-4 of units **C** and **D**, and C-3 of unit **E** were shifted downfield to δ 75.3–81.8, as compared with their positions in the spectra of the corresponding non-substituted monosaccharides at δ 68.4–72.0.[9] These displacements were due to glycosylation effects and defined the modes of substitution of the monosaccharides in the OPS.

A 2D ROESY experiment revealed interresidue cross-peaks between the following anomeric protons and protons at the linkage carbons: **A** H-1,**B** H-3; **B** H-1,**C** H-4; **C** H-1,**D** H-4; **D** H-1,**E** H-3; and **E** H-1,**B** H-4 at δ 4.62/3.86; 5.30/3.60; 4.86/3.98; 4.56/3.84; and 4.81/4.29, respectively (Table 1). These data are in agreement with the glycosylation pattern as determined by the ^{13}C NMR chemical shift data and established the sequence of the monosaccharide residues in the O-unit.

The ROESY experiment showed also Pyr H-3,**A** H-4 and Pyr H-3,**A** H-6a correlations at δ 1.44/3.44 and 1.44/3.64, respectively, which indicated the location of the pyruvic acid acetal at positions 4 and 6 of unit **A**. The (*S*)-configuration of the acetal carbon was inferred by the Pyr C-3 chemical shift of δ 26.1 compared with published data for the corresponding (*R*)- and (*S*)-isomers.[10] Hence, the polysaccharide chain of LPSOH has the structure shown in Chart 2.

The ^{1}H and ^{13}C (Fig. 1B) NMR spectra of the DPS showed that the majority of the repeating units lack pyruvic acid. The major series of the signals in the spectra were assigned in the same manner as for LPSOH (Table 1). A significantly lower-field position of the signals for C-4 and C-6 of unit **A** at δ 78.1 and 65.8 and a higher-field position of the signals for C-3 and C-5 of unit **A** at δ 72.3 and 67.1 in the ^{13}C NMR spectrum of LPSOH, as compared with their positions at δ 71.6, 62.5, 75.1, and 77.3, respectively, in the spectrum

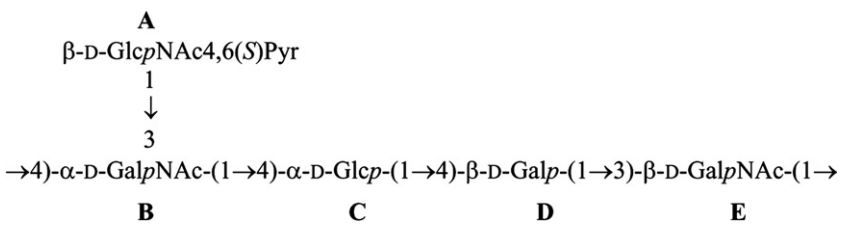


Chart 2. Structure of the polysaccharide chain of LPSOH from *S. dysenteriae* type 2 (this work).

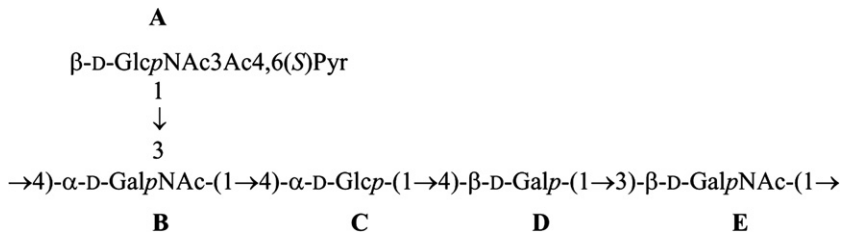


Chart 3. Structure of the O-polysaccharide of *S. dysenteriae* type 2 (this work).

of DPS, confirmed the location of the pyruvic acid acetal.

The position of the *O*-acetyl group was determined by a $^{1}H,^{13}C$ HMQC experiment with the OPS (its ^{13}C NMR spectrum is shown in Fig. 1C). As compared with the $^{1}H,^{13}C$ HMQC spectrum of the DPS, this showed a significant downfield displacement of ~70% of the **A** H-3,C-3 cross-peak from δ 3.55/75.1 to 5.05/76.9. This displacement is due to a deshielding effect of the *O*-acetyl group and indicated partial O-acetylation of unit **A** at position 3.

Therefore, based on the data obtained, it is concluded that the O-polysaccharide of *S. dysenteriae* type 2 has the structure shown in Chart 3. This structure differs from the published structure in the presence of a pyruvic acid acetal, which was overlooked in the earlier studies of the *S. dysenteriae* type 2 O-polysaccharide,[8] and in the configuration of two glycosidic linkages (units **A** and **E** are β-linked rather than α-linked).

It is known that the O-antigens of *S. dysenteriae* type 2 and *E. coli* O112ac are serologically identical.[11] In order to reveal the structural basis for the serological relationship, we studied the O-antigen structure of *E. coli* O112ac. The ^{1}H NMR spectrum of the *E. coli* O112ac LPS showed, inter alia, a resonance at δ 1.49 and resonances from *N*-acetyl groups at $\delta \sim 2$. However, resonances from *O*-acetyl groups at a slightly higher chemical shift were absent. The O-polysaccharide was obtained by mild acid degradation of the LPS of *E. coli* O112ac at pH 4.2. Sugar analysis by GLC of the alditol acetates after full acid hydrolysis of the polysaccharide revealed Glc, Gal, GlcN, and GalN. GLC analysis of the acetylated (*S*)-2-butyl glycosides demonstrated the D configuration of these monosaccharides.

The $^{1}H,^{13}C$ HSQC NMR spectrum of the *E. coli* O112ac polysaccharide (Fig. 2) contains cross-peaks for five sugar residues and a cross-peak at δ_H/δ_C 1.49/25.4 for a pyruvic acid 4,6-acetal of a hexopyranosyl residue.[12] The HSQC spectrum was in full agreement with that of the LPSOH from *S. dysenteriae* type 2, thereby confirming the structural identity of the two O-polysaccharides, except for the presence of *O*-acetyl groups in the latter (vide supra).

The O-polysaccharides of *S. dysenteriae* type 2 and *E. coli* O112ac show a structural similarity to the O-polysaccharides of another pair, *S. boydii* type 15 and *E. coli* O112ab, which are also identical structurally[2,13] and serologically.[11] The latter pair differ in having an L-IdoA residue in place of D-Gal and lacking the pyruvate acetal. Nevertheless, the similarity of the O-antigens provides a serological relatedness significant enough for considering *E. coli* O112ab and O112ac as variants of the serogroup.

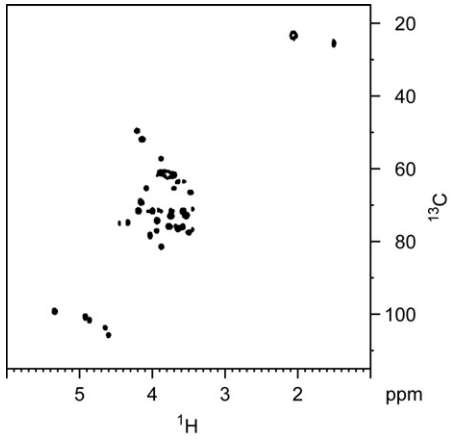


Figure 2. $^{1}H,^{13}C$ HSQC NMR spectrum of the O-polysaccharide from *E. coli* O112ac.

1. Experimental

1.1. Bacterial strains and isolation of lipopolysaccharides

S. dysenteriae type 2 (G1252) was from the Institute of Medical and Veterinary Science, Adelaide, Australia, and *E. coli* O112ac:K66:H^{-} (CCUG 11414) was from the Culture Collection, University of Gothenburg, Sweden. Bacteria were grown to late log phase using a fermentor under constant aeration at 37 °C and pH 7.0 as described for *S. dysenteriae* type 4;[6] bacterial cells were washed and dried. Lipopolysaccharides were isolated from dried bacterial cells by the phenol–water method[14] in a yield of 7% and purified as described.[6] *E. coli* O112ac was grown and extracted as previously described.[15]

1.2. Degradation of lipopolysaccharides

Delipidation of the LPS of *S. dysenteriae* type 2 was performed by treatment with aq 2% HOAc or 0.1 M sodium acetate pH 4.2, respectively, at 100 °C for 5 h. The precipitate of lipid A was removed by centrifugation (13,000*g*, 20 min), and the high-molecular-mass OPS was recovered from the supernatant by fractionation by GPC on a column (56 × 2.6 cm) of Sephadex G-50 (S) (Amersham Biosciences, Sweden) in 0.05 M pyridinium acetate buffer pH 4.5, monitored by a differential refractometer (Knauer, Germany).

The LPS and OPS of *S. dysenteriae* type 2 were treated with aq 12.5% ammonia at 37 °C for 16 h, the solution was desalted by gel filtration on Sephadex G-50 (S) as described above to give the O-deacylated polymers LPSOH and DPS, respectively. *E. coli* O112ac was delipidated and purified as previously described.[15]

1.3. Chemical analyses

The OPS of *E. coli* O112ac was hydrolyzed with 2 M CF_3CO_2H (120 °C, 2 h). Monosaccharides were identified by GLC of the alditol acetates on a Hewlett–Packard 5890 Series II instrument equipped with an HP fused silica column (0.25 mm × 30 m) using a temperature program of 180 °C (1 min), a temperature increase to 210 °C (3 °C min^{-1}) and 210 °C (10 min). The absolute configuration of the monosaccharides was determined by GLC of the acetylated (*S*)-2-butyl glycosides.[16]

1.4. NMR spectroscopy

NMR spectra of the DPS and LPSOH of *S. dysenteriae* type 2 OPS were measured at 30 and 40 °C, respectively, using a Bruker DRX-500 MHz spectrometer. NMR spectra of the OPS of *E. coli* O112ac were recorded at 55 °C using a Bruker AVANCE 500 MHz spectrometer equipped with a 5 mm PFG triple-resonance CryoProbe. Samples were freeze-dried from 99.9% D_2O and studied as solutions in 99.96% D_2O using acetone (δ_H 2.225) or sodium 3-trimethylsilyl-[2,2,3,3-$^2H_4$]propanoate (δ_H 0.00) as the internal reference and acetone (δ_C 31.45) or dioxane (δ_C 67.4) in D_2O as the external reference. Data processing was performed using vendor-supplied software.

Acknowledgments

This work was supported by the Russian Foundation for Basic Research (Project 05-04-48992), the Russian Science Support Foundation, the Council on Grants at the President of the Russian Federation for Support of Young Russian Scientists (MK-157.2007.4), the Swedish Research Council, the Swedish Agency for Research Cooperation with Developing Countries, Tianjin Municipal Special Fund for Science and Technology Innovation Grant 05FZZDSH00800, the NSFC Key Programs Grants 30530010 and 20536040, NSFC General Program Grant 30670038, the Chinese National Science Fund for Distinguished Young Scholars (30788001), and the National 863 program of China Grants 2006AA020703 and 2006AA06Z409. We thank Mrs. M. Sörensson for technical assistance.

References

1. Reeves, P. R.; Wang, L. *Microbiol. Immunol.* **2002**, *264*, 109–135.
2. Liu, B.; Knirel, Y. A.; Feng, L.; Perepelov, A. V.; Senchenkova, S. N.; Wang, Q.; Reeves, P.; Wang, L., in press.
3. Senchenkova, S. N.; Feng, L.; Yang, J.; Shashkov, A. S.; Cheng, J.; Liu, D.; Knirel, Y. A.; Reeves, P.; Jin, Q.; Ye, Q.; Wang, L. *J. Bacteriol.* **2005**, *187*, 2551–2554.
4. Jonsson, K. H. M.; Weintraub, A.; Widmalm, G. *Carbohydr. Res.* **2006**, *341*, 2986–2989.
5. Zhao, G.; Perepelov, A. V.; Senchenkova, S. N.; Shashkov, A. S.; Feng, L.; Li, X.; Knirel, Y. A.; Wang, L. *Carbohydr. Res.* **2007**, *342*, 1275–1279.
6. Perepelov, A. V.; Liu, B.; Senchenkova, S. N.; Shashkov, A. S.; Feng, L.; Knirel, Y. A.; Wang, L. *Carbohydr. Res.* **2007**, *342*, 2676–2681.
7. Perepelov, A. V.; Senchenkova, S. N.; Shashkov, A. S.; Liu, B.; Feng, L.; Wang, L.; Knirel, Y. A. *Russ. J. Bioorg. Chem.* **2008**, *34*, 110–117.
8. Dmitriev, B. A.; Knirel, Y. A.; Kochetkov, N. K.; Hofman, I. L.; Čapek, K. *Eur. J. Biochem.* **1977**, *76*, 433–440.
9. Jansson, P.-E.; Kenne, L.; Widmalm, G. *Carbohydr. Res.* **1989**, *188*, 169–191.
10. Garegg, P. J.; Jansson, P.-E.; Lindberg, B.; Lindh, F.; Lönngren, J.; Kvarnström, I.; Nimmich, W. *Carbohydr. Res.* **1980**, *78*, 127–132.
11. Ewing, W. H. In *Identification of Enterbacteriaceae*; Edwards, P. R., Ewing, W. H., Eds.; Elsevier: New York, 1986; pp 135–172.
12. Jansson, P.-E.; Lindberg, J.; Widmalm, G. *Acta Chem. Scand.* **1993**, *47*, 711–715.
13. Perepelov, A. V.; Liu, B.; Senchenkova, S. N.; Shashkov, A. S.; Feng, L.; Knirel, Y. A.; Wang, L. *Carbohydr. Res.* **2008**, *343*, 571–575.
14. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
15. Larsson, E. A.; Urbina, F.; Yang, Z.; Weintraub, A.; Widmalm, G. *Carbohydr. Res.* **2004**, *339*, 1491–1496.
16. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1979**, *77*, 1–7.