A Novel Open-Chain Nononic Acid Linked by an Ether Bond to Glucose as a Polysaccharide Constituent

Elisabeth J. Faber, J. Albert van Kuik, Koen M. Halkes, Johannis P. Kamerling,* and Johannes F. G. Vliegenthart[a]

Abstract: A novel sugar constituent was isolated from the heteropolysaccharide excreted by *Streptococcus thermophilus* 8S when grown in skimmed milk. The structure and absolute configuration were determined by means of chemical analysis, mass spectrometry, NMR spectroscopy, along with molecular dynamics simulations, and was shown to be 6-*O*-(3′,9′-dideoxy-D-*threo*-D-*altro*-nononic acid-2′-yl)-D-glucopyranose.

Keywords: carbohydrates · chirality · exopolysaccharides · *Streptococcus thermophilus* · structure elucidation

Introduction

The exopolysaccharides (EPSs) produced by lactic acid bacteria are promising candidates for a new generation of food thickeners because of their physical properties in combination with their GRAS (generally recognized as save) status. For this reason, detailed structural studies have been performed on EPSs produced by various species of the *Lactobacillus*, *Lactococcus*, and *Streptococcus* genera, as summarized in reference [1]. The EPSs characterized so far mainly consist of combinations of glucose, galactose, and/or rhamnose, and to a minor extent of *N*-acetylglucosamine, *N*-acetylgalactosamine, or glucuronic acid. In some cases, non-carbohydrate substituents, such as phosphate, acetyl, and glycerol groups, are present.

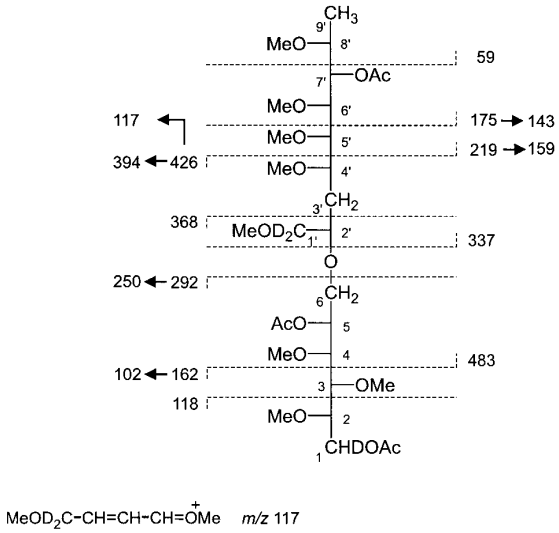
Here, we report on the structural elucidation of a novel sugar constituent found in the EPS excreted by the lactic acid bacterium *Streptococcus thermophilus* 8S when grown in skimmed milk. The absolute configuration of this sugar constituent was determined by NMR analysis and comparison to reference compounds, supplemented with MD calculations.

Results and Discussion

Isolation, purification, and composition of the exopolysaccharide: The EPS produced by *S. thermophilus* 8S in reconstituted pasteurized skimmed milk was isolated and purified by means of ethanol precipitation, followed by acetone precipitation and gelfiltration on Sephacryl S-500. Monosaccharide analysis of the native EPS (n-EPS), including the determination of absolute configurations, showed the presence of D-Gal, D-Glc, D-GalNAc, and D-Rib in a molar ratio of 2:1:1:1, as well as an unknown sugar component in a molar ratio of 0.7 in terms of peak areas relative to Glc. Methylation analysis of n-EPS revealed the presence of an unknown sugar component along with 4-substituted Gal*p*, 4-substituted Glc*p*, 4-substituted Gal*p*NAc, and 2-substituted Rib*f* in a molar ratio of 2:1:1:1. According to NMR experiments (data not shown), the Gal and Glc residues are in the pyranose ring form and the Rib residue is in the furanose ring form. No conclusive structural information could be deduced from the [1-D]alditol mass spectrum of the unknown sugar component. To investigate whether the identification problems might be caused by the presence of a carboxyl group, n-EPS was subjected to a carboxyl reduction ($NaBD_4$; cr-EPS) prior to methylation analysis. EIMS of the partially methylated [1-D]alditol acetate originating from the unknown sugar component in cr-EPS revealed a fragmentation pattern in accordance with a polyhydroxy C_9 chain linked at C2′ to C6 of a [1-D]Hex-ol residue through an ether bond (Figure 1). The *O*-acetyl groups at C1 and C5 of [1-D]Hex-ol indicated a pyranose ring for the Hex moiety of the unknown sugar component, whereas the *O*-acetyl group at C7′ suggested a glycosylation site. The doubly deuterium-labeled C1′ atom reflects a carboxyl group, originally present. Furthermore, the fragmentation pattern indicated deoxygenation at C3′ and C9′ to give rise to a dideoxy sugar. To determine the identity of the Hex residue, the carboxyl-reduced component was isolated from a hydrolyzate of cr-EPS and subjected to treatment with BBr_3 in order to cleave the ether bond

[a] Prof. Dr. J. P. Kamerling, Dr. E. J. Faber, Dr. J. A. van Kuik, Dr. K. M. Halkes, Prof. Dr. J. F. G. Vliegenthart
Bijvoet Center, Department of Bio-Organic Chemistry
Section of Glycoscience and Biocatalysis
Padualaan 8, 3584 CH Utrecht (The Netherlands)
Fax: (+31) 30-2540980
E-mail: j.p.kamerling@chem.uu.nl

  0947-6539/02/0819-4498 $ 20.00+.50/0



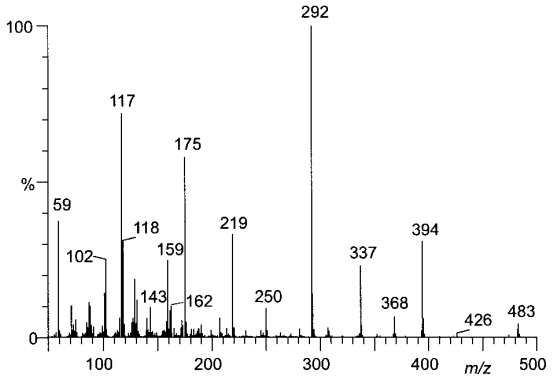
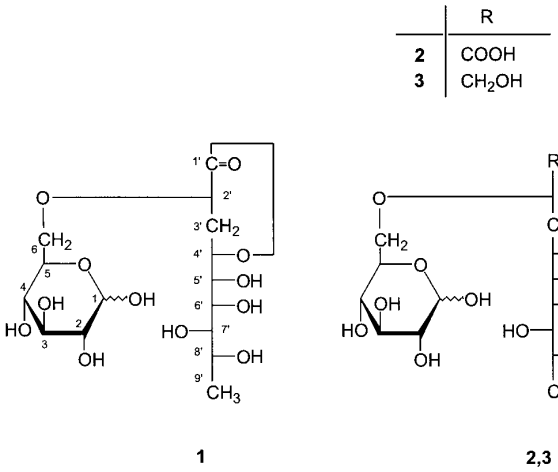


Figure 1. EIMS spectrum of the partially methylated [1-D]alditol acetate of the novel sugar constituent isolated from cr-EPS.



Scheme 1. The lactonized form (**1**), the acid form (**2**) and the reduced form (**3**) (C_9 moiety) of 6-*O*-(3′,9′-dideoxynononic acid-2′-yl)-D-glucopyranose.

between C2′ and C6. Monosaccharide analysis of the obtained material and determination of the absolute configuration indicated the Hex residue to be D-Glc. On the basis of this information, the unknown sugar component is proposed to be a 6-*O*-(3′,9′-dideoxynononic acid-2′-yl)-D-glucopyranose.

Identification of the unknown sugar component: For the structural characterization of the unknown exopolysaccharide constituent, the nonderivatized and the carboxyl-reduced ($NaBH_4$) unknown sugar components were isolated from acid hydrolyzates of n-EPS and cr-EPS, respectively, by means of graphitized carbon. MALDI-TOF analysis of the compound isolated from n-EPS revealed a [M^++Na] pseudomolecular ion at $m/z = 421$, corresponding to the mass of structure **1**, which is the lactonized form (C_9 moiety) of 6-*O*-(3′,9′-dideoxynononic acid-2′-yl)-D-glucopyranose (Scheme 1).

In the anomeric region ($\delta = 4.4 - 5.4$ ppm) of the one-dimensional ^{1}H NMR spectrum of **1** (Figure 2A) doublets were observed at $\delta = 5.205$ ($^3J_{1,2} = 3.8$ Hz; α-pyranose) and 4.629 ppm ($^3J_{1,2} = 7.9$ Hz; β-pyranose), indicating the two anomers of the Glc moiety of **1**. Furthermore, signals for the methylene ($\delta \approx 2.71$, 2.23 ppm) and methyl ($\delta = 1.180$ ppm) protons were observed that originated from the C_9 moiety of **1**. Comparison of the two-dimensional TOCSY spectra of **1**

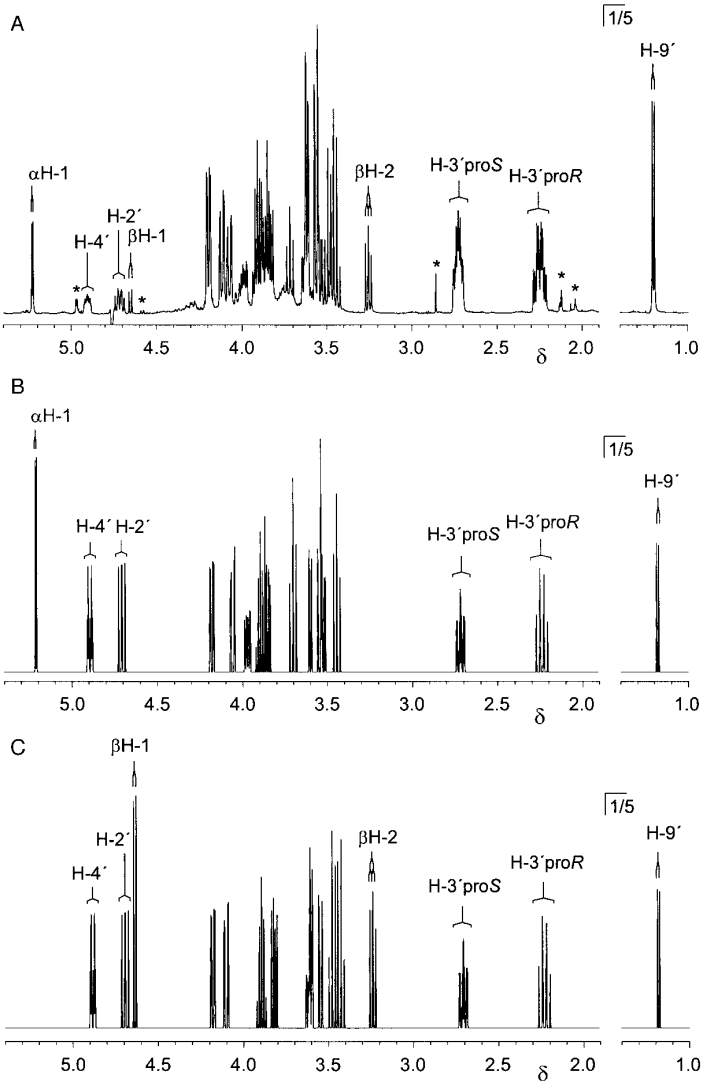


Figure 2. A) One-dimensional ^{1}H NMR spectrum of **1**, recorded at 500 MHz and 27 °C, and simulated spectra of **1** with B) α or C) β configurations of the Glc moiety, respectively. Asterisks (*) in spectrum A indicate impurities.

  0947-6539/02/0819-4499 $ 20.00+.50/0

Table 1. ^{1}H NMR chemical shifts[a] of **1**, **2**, and **3**, and ^{13}C NMR chemical shifts[b] of **1**, recorded in D_2O at 27 °C. Coupling constants (J [Hz]) are included in parentheses.

	Proton		**1**[c]		**2**[d]	**3**	Carbon	**1**
α-D-Glc*p*	H1		5.205	($J_{1,2}=3.8$)	5.296	5.224	C1	92.9
	H2		3.518	($J_{2,3}=9.8$)	3.345	3.50	C2	72.2
	H3		3.694	($J_{3,4}=9.3$)	n.d.	3.70	C3	73.6
	H4		3.437	($J_{4,5}=10.0$)	n.d.	3.52	C4	70.3
	H5		3.964	($J_{5,6proS}=2.0$)	n.d.	3.93	C5	71.0
	H6pro*S*		4.051	($J_{5,6proR}=5.1$)	n.d.	3.98	C6	70.0
	H6pro*R*		3.846	($J_{6proS,6proR}=-11.0$)	n.d.	3.78		
β-D-Glc*p*	H1		4.629	($J_{1,2}=7.9$)	4.685	4.643	C1	96.8
	H2		3.231	($J_{2,3}=9.5$)	3.065	3.24	C2	74.9
	H3		3.468	($J_{3,4}=9.3$)	n.d.	3.47	C3	76.5
	H4		3.420	($J_{4,5}=10.0$)	n.d.	3.47	C4	70.3
	H5		3.606	($J_{5,6proS}=1.9$)	n.d.	3.56	C5	75.4
	H6pro*S*		4.095	($J_{5,6proR}=6.0$)	n.d.	4.03	C6	70.4
	H6pro*R*		3.812	($J_{6proS,6proR}=-11.2$)	n.d.	3.74		
C_9 moiety	H2′	α	4.699	($J_{2',3'proS}=8.7$)	3.971	3.77	C1′	n.d.
		β	4.686	($J_{2',3'proR}=10.7$)			C2′	76.5
	H3′pro*S*	α	2.710	($J_{3'proS,3'proR}=-12.3$)	1.90	1.77	C3′	28.0
		β	2.699	($J_{3'proS,4'}=2.7$)				
	H3′pro*R*	α	2.233	($J_{3'proR,4'}=10.3$)	1.79	1.60		
		β	2.223					
	H4′	α	4.886	($J_{4',5'}=2.6$)	4.07	4.12	C4′	79.5
		β	4.876					
	H5′		4.173	($J_{5',6'}=9.3$)	3.68	3.82	C5′	69.3
	H6′		3.539	($J_{6',7'}=1.4$)	3.66	3.60	C6′	71.0
	H7′		3.594	($J_{7',8'}=7.1$)	3.58	3.62	C7′	74.5
	H8′		3.887	($J_{8',9'}=6.5$)	3.90	3.91	C8′	69.4
	H9′		1.180		1.179	1.200	C9′	18.9

[a] Relative to the signal of internal acetone at $\delta=2.225$. [b] Relative to the α-anomeric signal of external [1-^{13}C]glucose at $\delta=92.9$. [c] Data were refined by simulation of the one-dimensional ^{1}H NMR spectrum. [d] n.d. = not determined.

with increasing mixing times (25 – 250 ms) allowed the assignment of the ^{1}H signals belonging to each of the spin systems of each of the anomers (Table 1). The αH1 and βH1 TOCSY tracks showed spin systems characteristic of α- and β-Glc*p*, respectively, and the H3′pro*R*, H3′pro*S*, and H9′ TOCSY tracks revealed the complete spin system of the C_9 moiety (Figure 3). Stereospecific assignments of H6pro*R* and H6pro*S* in Glc were performed by analogy with ^{1}H NMR studies of stereospecifically deuterated D-Glc*p*,[2] and are in agreement with the assignments of H6pro*R* and H6pro*S* of 6-substituted D-Glc*p* in gentiobiose.[3] The methylene protons of the C_9 moiety were assigned as pro*R* and pro*S* in accordance with the configuration of the C_9 moiety of compound **1** (vide infra). The two-dimensional ^{13}C,^{1}H HMQC spectra of **1** (Figure 4) allowed the assignment of the ^{13}C signals (Table 1). The downfield chemical shifts of αC6 ($\delta=70.0$ ppm) and βC6 ($\delta=70.4$ ppm) in comparison with NMR data of Glc*p* (α-D-Glc*p*, $\delta_{C6}=61.6$; β-D-Glc*p*, $\delta_{C6}=61.7$ ppm[4]) confirmed the 6-substitution of the Glc*p* moiety of **1**.

As indicated by NMR and MS analysis of **1**, this compound was released as an 1′,4′-lactone by hydrolysis of n-EPS. The lactone ring of **1** was opened (→**2**) by adding NaOD prior to NMR analysis. The H1 and H2 signals of the Glc*p* moiety and all the ^{1}H resonances of the C_9 moiety of **2** were assigned by means of TOCSY experiments (mixing times 25 – 250 ms), and were consistent with structure **2** depicted in Scheme 1. In addition, the ^{1}H signals of the component isolated from cr-EPS (Table 1) were essentially assigned as described for **1**, and were indicative of structure **3** (Scheme 1).

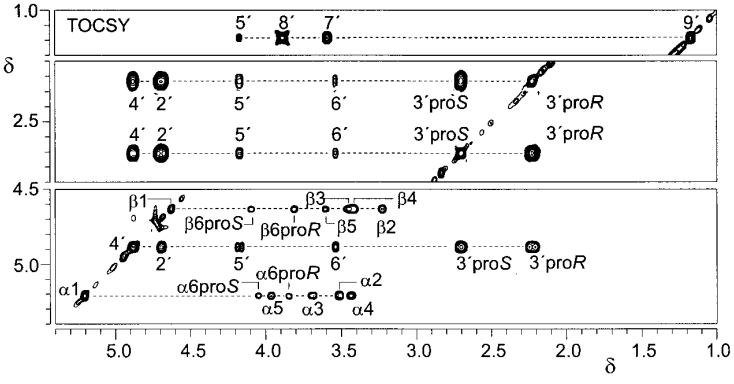


Figure 3. Parts of the 500 MHz two-dimensional TOCSY spectrum of **1**, recorded in D_2O at 27 °C. Diagonal peaks of the anomeric protons of the Glc*p* moiety, and of H3′pro*R*, H3′pro*S*, H4′, and H9′ of the C_9 moiety are indicated. Labels near crosspeaks refer to the protons of the scalar-coupling network belonging to the diagonal peak. The absence of the H2′ track in F2 is caused by HOD suppression.

^{1}H,^{1}H coupling constants of compound **1** (Table 1) were determined by simulation of ^{1}H subspectra for the α- and β-anomers of **1** (Figure 2B and C). Two sets of chemical shifts are observed for H2′, H3′pro*R*, H3′pro*S*, and H4′, induced by the two anomeric configurations of Glc*p*. The vicinal coupling constants $J_{4',5'}$ (2.6), $J_{5',6'}$ (9.3), $J_{6',7'}$ (1.4), and $J_{7',8'}$ (7.1 Hz) in the C_9 moiety indicate that the dihedral angles χ^1 (H4′-C4′-C5′-H5′), χ^2 (H5′-C5′-C6′-H6′), χ^3 (H6′-C6′-C7′-H7′), and χ^4 (H7′-C7′-C8′-H8′) are predominantly in the regions of ±60°, 180°, ±60°, and 180°, respectively.[5] The coupling values of 2.6 Hz and 7.1 Hz imply noticeable contributions of other conformations.

For the determination of the *R*/*S* configurations at the chiral C2′ and C4′ atoms, distance information for the intraresidual protons of **1** (1′,4′-lactone) was essential. In the two-dimensional ROESY spectrum (Figure 5) H3′pro*S* showed strong crosspeaks with H2′ and H4′, whereas H3′pro*R* showed weak crosspeaks with H2′ and H4′. This proves that H2′, H3′pro*S*, and H4′ are located on the same side of the lactone ring. This observation excludes the 2′*S*,4′*R* and 2′*R*,4′*S* configurations; this means that it is necessary to differentiate between the 2′*S*,4′*S* and 2′*R*,4′*R* configurations.

ROE interactions between protons of D-Glc*p* and the lactone ring, detected in the spectrum of **1**, allow the determination of the absolute configuration of the chiral

0947-6539/02/0819-4500 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 19

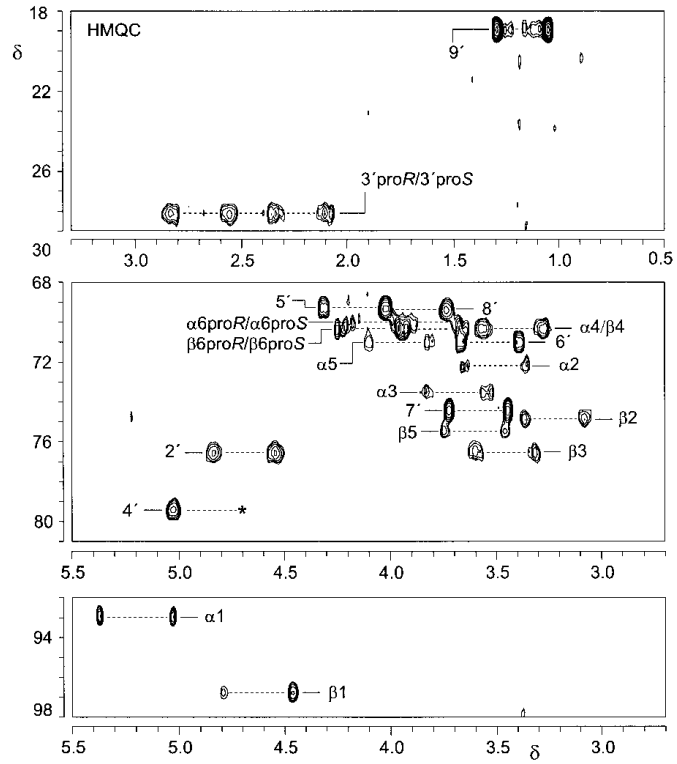


Figure 4. Part of the 500 MHz ^{13}C, ^{1}H HMQC spectrum of **1**, recorded in D_2O at 27 °C. α1 stands for the set of crosspeaks between H1 and C1 of the α-D-Glc*p* moiety, etc. The asterisk (*) indicates a signal lost by HOD suppression.

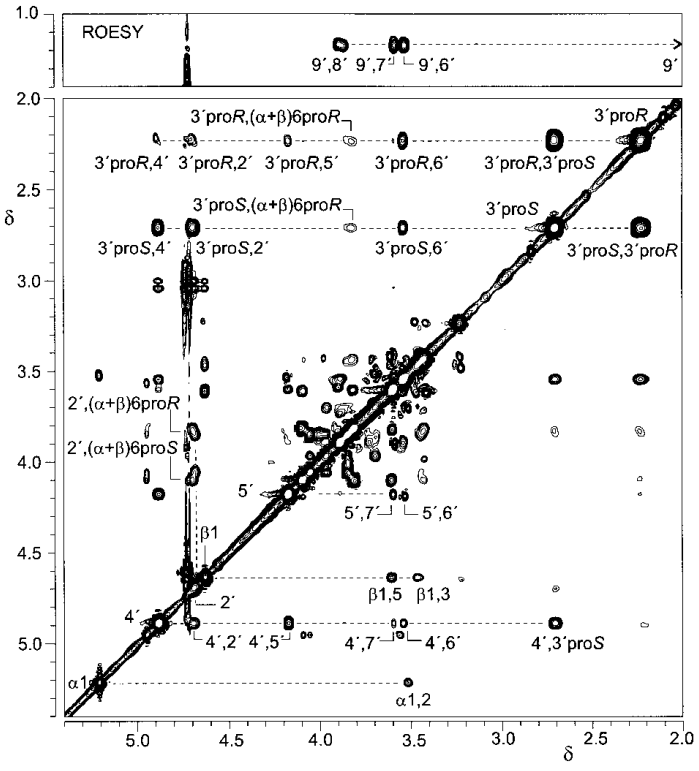


Figure 5. 500 MHz two-dimensional ROESY spectrum of **1**, recorded in D_2O at 27 °C. α1 corresponds to the diagonal peak belonging to residue αH1; α1,2 refers to an intraresidual crosspeak between αH1 and αH2, and 3′pro*S*,(α+β)6pro*R* means an interresidual connectivity from H3′pro*S* to αH6pro*R* and βH6pro*R*, etc. The absence of the H2′ track in F2 is caused by HOD suppression.

atoms in the lactone ring. In the case of compound **1**, it is expected that the conformation at C2′-O6-C6-C5 is similar to a (1→6)-glycosidic linkage in oligosaccharides. The conformation of a (1→6) linkage is determined by three glycosidic dihedral angles, ϕ (O5′-C1′-O6-C6), ψ (C1′-O6-C6-C5), and ω (O6-C6-C5-C4). The vicinal coupling constants $J_{5,6proR}$ and $J_{5,6proS}$ (≈6 and ≈2 Hz) in **1** indicate conformational behavior for ω as expected for a (1→6)-glycosidic linkage.[3] Crystal structures and molecular mechanics (MM) modeling results of monosaccharides that are (1→6)-linked to glucose, as in raffinose (α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-β-D-Fru*f*),[6] melibiose (α-D-Gal*p*-(1→6)-α-D-Glc*p*),[7] and panose (α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-α-D-Glc*p*),[8] show an extended planar conformation for C1′-O6-C6-C5. Crosspeaks of equal intensity observed on the H2′ ROESY track of **1**, H2′,H6pro*R* and H2′,H6pro*S*, indicated a similar extended planar conformation for C2′-O6-C6-C5 (Figure 6A). The

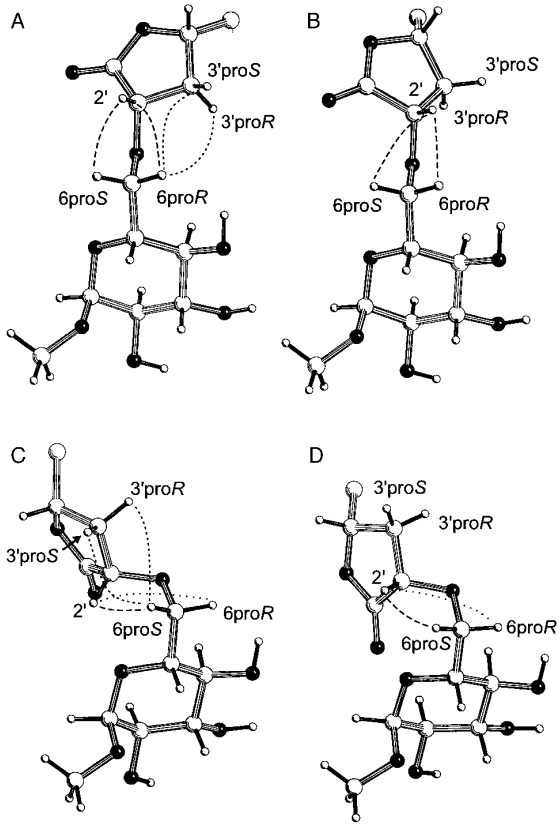


Figure 6. The four main (ϕ, ψ)-conformations of compound **1**. A) ϕ, ψ = 95°, 175°. B) ϕ, ψ = 155°, 175°. C) ϕ, ψ = 95°, 85°. D) ϕ, ψ = 155°, 85°.

crosspeaks observed on the H3′pro*R* and H3′pro*S* ROESY tracks to the H6pro*R* signal show that these methylene protons are predominantly in the proximity of H6pro*R*, but not in that of H6pro*S*. These findings indicate the *S* configuration (2′*S*,4′*S*) rather than the *R* configuration (2′*R*,4′*R*) occurs at both chiral atoms in the lactone ring. In oligosaccharides, the flexibility of a (1→6)-glycosidic linkage results in multiple conformational regions. Similarly, ROESY crosspeaks related to the extended planar conformation in **1** are most likely the result of conformational averaging, and,

  0947-6539/02/0819-4501 $ 20.00+.50/0

therefore, the ϕ, ψ conformational space of **1** was explored to estimate the contribution of other conformations to these ROESY crosspeaks. It should be noted that the ω dihedral angle in compound **1** does not influence the ROESY crosspeak intensities between D-Glc*p* and the lactone ring. The dihedral angles ϕ (C3′-C2′-O6-C6) and ψ (C2′-O6-C6-C5) in compound **1** were examined by molecular dynamics (MD) calculations with the method of adaptive umbrella sampling of the potential-of-mean-force.[9] Free-energy profiles, obtained by sampling the individual dihedral angles, were translated into rotamer population distributions (Figure 7).

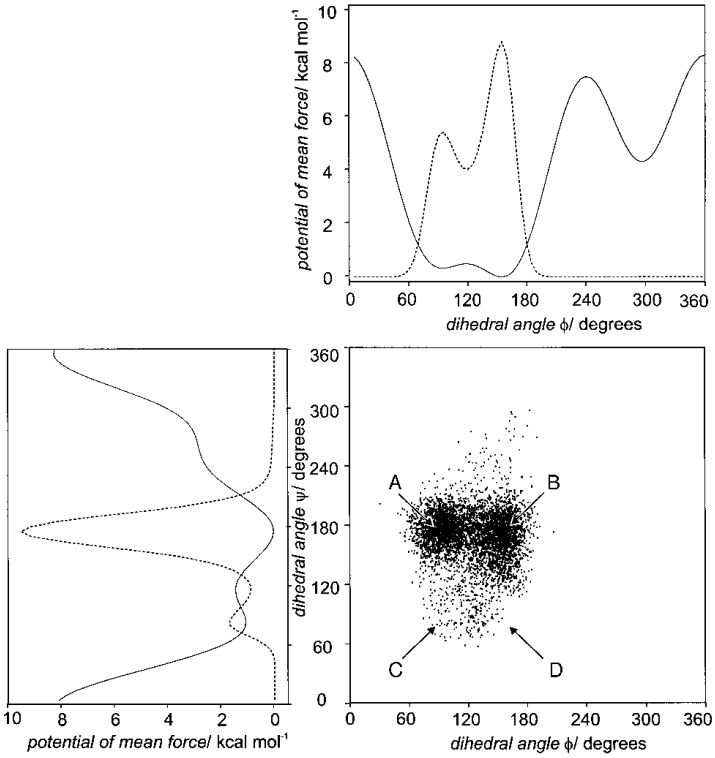


Figure 7. Potential-of-mean-force profiles (solid lines) and probability distribution profiles (dotted lines) of the orientation of the dihedral angles ϕ (C3′-C2′-O6-C6) (top panel) and ψ (C2′-O6-C6-C5) (left panel). Scatter plot (bottom-right panel) of the orientations of the dihedral angles ϕ and ψ of a MD simulation of compound **1** in water. The orientations of the dihedral angles of the four main (ϕ, ψ)-conformations (Figure 6) are indicated by arrows.

The ϕ, ψ conformational space occupies a single large region with local minima at 95° and 155° for ϕ and at 85° and 175° for ψ. This region was explored further in a 5 ns MD simulation. The two most populated areas were at ϕ, $\psi = 95°$, 175° and ϕ, $\psi = 155°$, 175°. The ϕ, ψ set 95°, 175° has also been found for the crystal structures of raffinose $\phi = 71.9°$, $\psi = -169.6°$; melibiose $\phi = 76.5°$, $\psi = -173.9°$; and panose $\phi = 71.4°$, $\psi = 165.2°$. Figure 6 shows models constructed according to the four main ϕ, ψ conformations. Only for conformations with ϕ, $\psi = 95°$, 175° and ϕ, $\psi = 95°$, 85° (Figure 6A and 6) are ROE contacts expected between H3′ and H6 (H3′pro-*R*,H6pro*R*, H3′pro*S*,H6pro*R*, H3′pro*R*,H6pro*S* or H3′pro*S*,H6pro*S*). In Figure 5, the ROESY crosspeaks H3′pro*R*,H6pro*R* and H3′pro*S*,H6pro*R*, together with the similar intensities of the H2′,H6pro*R* and H2′,H6pro*S* ROESY crosspeaks, are in agreement with the calculated preference for $\psi = 175°$ (>85 %, Figure 7 bottom-right figure, A and B areas). A preference for $\psi = 85°$ would have been revealed by a strong H2′,H6pro*S* ROE contact in combination with a weak H2′,H6pro*R* contact (Figure 6C). The free-energy profiles of the mirror image of the lactone ring, the (2′*R*,4′*R*)-configuration, gave essentially opposite results. Therefore, only the (2′*S*,4′*S*) configuration fits the NMR data.

To determine the *R*/*S* configurations at the chiral C5′, C6′, C7′, and C8′ atoms, model structures with all possible combinations for these chiral atoms were computer-generated and investigated. The model structures were created with $\chi^1 = \pm 60°$, $\chi^2 = 180°$, $\chi^3 = \pm 60°$, and $\chi^4 = 180°$ (vide supra), in agreement with the vicinal coupling constants (Table 1). This resulted in 64 possibilities. The structures were labeled according to the configurations of the chiral C5′, C6′, C7′, and C8′ atoms as follows: $\mathbf{C}_{RRRS}$ = model structure with a 5′*R*,6′*R*,7′*R*,8′*S* configuration, etc.

The strong crosspeaks in the ROESY spectrum of **1** (Figure 5), originating from atoms separated by more than four bonds (H3′pro*R*,H6′, H3′pro*S*,H6′ and H6′,H9′), were used to check the 64 structures for violations of $^1H-^1H$ distances. To compensate for idealized values of the dihedral angles in the model structures, the upper proton distance was set to 3.6 Å, which is approximately 1 Å larger than the distances between H3′pro*R*,H6′, H3′pro*S*,H6′, and H6′,H9′, calculated from the ROE intensities. Note that assuming conformational averaging will make these distances shorter. This resulted in eight possible configurations: $\mathbf{C}_{RRRS}$ ($\chi^1 = -60$, $\chi^3 = -60$), $\mathbf{C}_{RRSR}$ ($\chi^1 = -60$, $\chi^3 = 60$), $\mathbf{C}_{RSRS}$ ($\chi^1 = -60$, $\chi^3 = -60$), $\mathbf{C}_{RSSR}$ ($\chi^1 = -60$, $\chi^3 = 60$), $\mathbf{C}_{SRRS}$ ($\chi^1 = -60$, $\chi^3 = -60$), $\mathbf{C}_{SRSR}$ ($\chi^1 = -60$, $\chi^3 = 60$), $\mathbf{C}_{SSRS}$ ($\chi^1 = -60$, $\chi^3 = -60$), and $\mathbf{C}_{SSSR}$ ($\chi^1 = -60$, $\chi^3 = 60$). Evaluation of the computer-generated structures with a 5′*S* configuration showed that, because of the significant differences in distances between H3′pro*R*,H6′ (≈2.6 Å) and H3′pro*S*,H6′ (≈3.6 Å) within these structures, the ROESY crosspeak H3′pro*R*,H6′ is expected to be much stronger than the crosspeak H3′pro*S*,H6′ (Figure 8A). This is in contrast to the computer-generated structures with 5′*R* configuration, whereby the two crosspeaks are expected to have the same intensity. The crosspeaks observed in the ROESY spectrum (Figure 5) have equal intensity and therefore suggest the 5′*R* configuration, an assignment that will be reassessed below. This reasoning leaves four possible configurations for the C5′–C8′ fragment: $\mathbf{C}_{RRRS}$, $\mathbf{C}_{RRSR}$, $\mathbf{C}_{RSRS}$, and $\mathbf{C}_{RSSR}$. The use of additional ROESY crosspeaks did not produce conclusive evidence for further reduction of the number of allowed model structures.

To select from the remaining four configurations, four deoxy-alditol model compounds were prepared, in which the coupling constants $J_{2,3}$, $J_{3,4}$, $J_{4,5}$ correspond to the coupling constants $J_{7',8'}$, $J_{6',7'}$, $J_{5',6'}$ in **1**, respectively, and a possible H1,H4 ROESY crosspeak to the H9′,H6′ crosspeak in **1**. The model compounds were 1-deoxy-D-Alt-ol, 1-deoxy-L-Glc-ol, 1-deoxy-L-Ido-ol, and 1-deoxy-D-Gal-ol, representing $\mathbf{C}_{RRRS}$, $\mathbf{C}_{RRSR}$, $\mathbf{C}_{RSRS}$, and $\mathbf{C}_{RSSR}$, respectively (Scheme 2). Of the vicinal coupling constants of the C5–C1 fragment in the 4 deoxy-alditols (Table 2), only those of 1-deoxy-L-Glc-ol match the corresponding data in compound **1**. This leaves

0947-6539/02/0819-4502 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 19

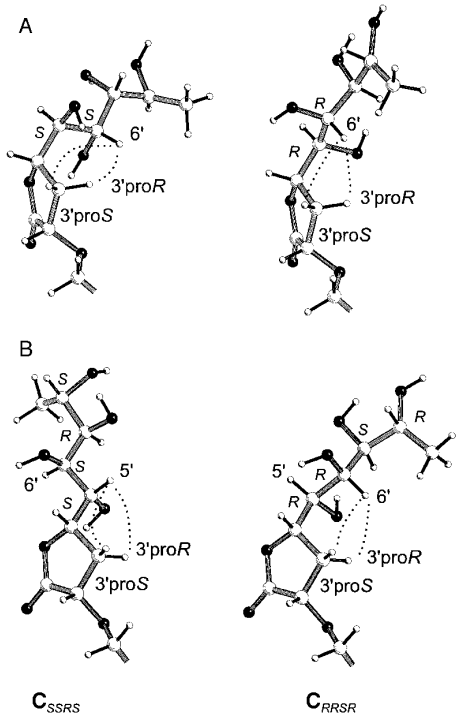
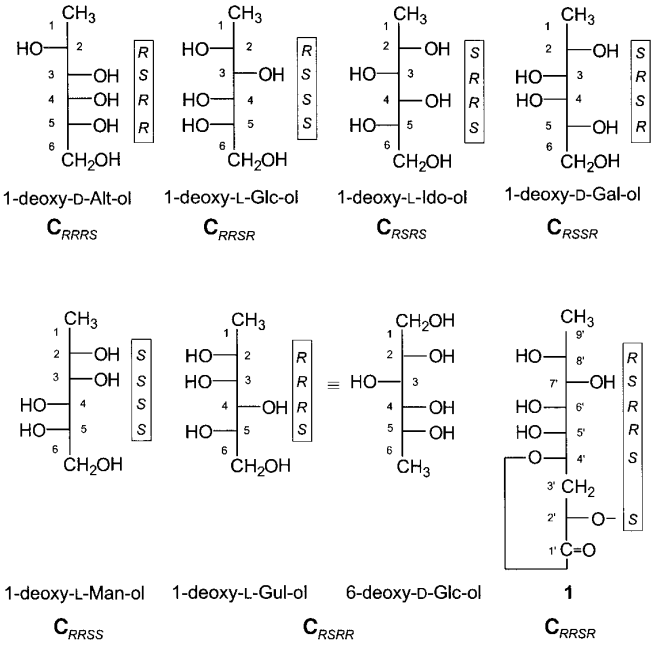


Figure 8. A) ROE contacts for configurations of compound **1** with the 5′*S* configuration (left) and the 5′*R* configuration (right) and with $\chi^1 = -60^\circ$ for both configurations. B) ROE contacts in the minimal energy conformations of compound **1** with the $\mathbf{C}_{SSRS}$ ($\chi^1 = 60^\circ$) and the $\mathbf{C}_{RRSR}$ ($\chi^1 = -60^\circ$) configuration.



Scheme 2. Fischer projection of six deoxy-alditol model compounds together with the C1′-C9′ fragment of **1**. The configurations of the chiral C5, C4, C3, and C2 atoms in the deoxy-alditols correspond to the configurations of the chiral C5′, C6′, C7′, and C8′ atoms in **1**, respectively. The deoxy-alditols were labeled according to the configurations of the chiral C5′, C6′, C7′, and C8′ atoms in **1** as follows: $\mathbf{C}_{RRRS}$ = model structure of **1** with 5′*R*,6′*R*,7′*R*,8′*S* configuration, etc. The actual *R*/*S* assignments of the deoxy-alditols are depicted in the boxes on the right side of the projections. The D-alditols, 1-deoxy-D-Alt-ol and 1-deoxy-D-Gal-ol, were used because the starting compounds for the respective L-alditols, L-Alt and D-Fuc, were unavailable. 1-Deoxy-D-Alt-ol and 1-deoxy-D-Gal-ol, actually represent the mirror image of the configurations $\mathbf{C}_{RRRS}$ and $\mathbf{C}_{RSSR}$, respectively.

$\mathbf{C}_{RRSR}$ to represent the correct configuration. Because the H7′,H8′ coupling constant ($J_{7',8'} = 7.1$ Hz) in compound **1** indicates a notable contribution of $\chi^4 = \pm 60^\circ$, the assumption that χ^4 is predominantly 180° might be too strict. Therefore, a new set of compounds was generated without the constraint $\chi^4 = 180^\circ$, but with $\chi^1 = \pm 60^\circ$, $\chi^2 = 180^\circ$, $\chi^3 = \pm 60^\circ$, and the H3′pro*R*,H6′, H3′pro*S*,H6′, and H6′,H9′ upper proton distance set to 3.6 Å. This allows $\mathbf{C}_{RRSS}$ and $\mathbf{C}_{RSRR}$ to become candidates. These configurations are represented by 1-deoxy-L-Man-ol and 6-deoxy-D-Glc-ol (1-deoxy-L-Gul-ol), respectively (Scheme 2 and Table 2). $\mathbf{C}_{RSRR}$ could be rejected on the basis of the vicinal coupling constants. The coupling constants of 1-deoxy-L-Man-ol are essentially identical to those of 1-deoxy-L-Glc-ol and match the corresponding data of compound **1**. The strong H9′,H6′ crosspeak in the ROESY spectrum of compound **1** was used to select between $\mathbf{C}_{RRSS}$ and $\mathbf{C}_{RRSR}$. ROESY spectra were recorded for both 1-deoxy-L-Man-ol and 1-deoxy-L-Glc-ol to assess the intensity of the H1,H4 crosspeak. The ROESY crosspeak intensities are given in Table 3, expressed relative to the crosspeak intensity of the methyl group and its nearest neighbor proton. The low intensity of the H1,H4 crosspeak in the spectrum of 1-deoxy-L-Man-ol, and the high intensity of this crosspeak in the spectrum of 1-deoxy-L-Glc-ol, justifies the elimination of $\mathbf{C}_{RRSS}$ and confirms $\mathbf{C}_{RRSR}$ to represent the correct configuration. The selection of 1-deoxy-L-Glc-ol over 1-deoxy-L-Man-ol is supported by the chemical shift of the methyl group, observed at $\delta = 1.195$ and 1.180 ppm for 1-deoxy-L-Glc-ol and compound **1**, respectively, and at $\delta = 1.268$ ppm for 1-deoxy-L-Man-ol (Table 2). Finally, the mirror image of this configuration, $\mathbf{C}_{SSRS}$, was evaluated to reassess the 5′*R* configuration. For both $\mathbf{C}_{RRSR}$ and $\mathbf{C}_{SSRS}$ the population distribution of the dihedral angle χ^1 was analyzed by the method of adaptive umbrella sampling of the potential-of-mean-force.[9] It appears that the favored conformation for $\mathbf{C}_{RRSR}$ is $\chi^1 = -60^\circ$ (-60°:60°:180° = 85:9:6), while for $\mathbf{C}_{SSRS}$ $\chi^1 = 60^\circ$ (-60°:60°:180° = 1:98:1). This means that

Table 2. ^{1}H NMR chemical shifts and ^{1}H,^{1}H coupling constants (J[Hz]) of deoxyalditols and the C5′–C9′ fragment of compound **1**, in D_2O.

	H1 ($J_{1,2}$)	H2 ($J_{2,3}$)	H3 ($J_{3,4}$)	H4 ($J_{4,5}$)	H5	H6
1-deoxy-D-Alt-ol	1.230 (6.6)	4.075 (2.3)	3.478 (7.6)	3.765 (5.2)	3.918	3.806/3.674
1-deoxy-L-Glc-ol	1.195 (6.4)	3.906 (7.4)	3.602 (1.5)	3.613 (8.8)	3.755	3.836/3.634
1-deoxy-L-Ido-ol	1.214 (6.5)	3.943 (5.4)	3.488 (4.2)	3.708 (4.7)	3.840	3.722/3.640
1-deoxy-D-Gal-ol	1.243 (6.6)	4.090 (1.9)	3.479 (9.1)	3.639 (1.6)	3.963	3.687/3.677
1-deoxy-L-Man-ol	1.268 (6.4)	3.862 (7.7)	3.591 (1.6)	3.785 (8.4)	3.740	3.844/3.658
1-deoxy-L-Gul-ol[a]	1.223 (6.4)	3.890 (6.5)	3.515 (3.1)	3.783 (5.4)	3.795	3.718/3.619
	H9′ ($J_{8',9'}$)	H8′ ($J_{7',8'}$)	H7′ ($J_{6',7'}$)	H6′ ($J_{5',6'}$)	H5′	
1	1.180 (6.5)	3.887 (7.1)	3.594 (1.4)	3.539 (9.3)	4.173	

[a] To simplify comparison, 6-deoxy-D-Glc-ol is written as 1-deoxy-L-Gul-ol.

  0947-6539/02/0819-4503 $ 20.00+.50/0

Table 3. ROESY crosspeak intensities for deoxyalditols and the C5′ – C9′ fragment of compound **1**, in D_2O.

	H1,H2	H1,H3	H1,H4
1-deoxy-L-Man-ol	1.00	0.62	0.18
1-deoxy-L-Glc-ol	1.00	0.52	1.06
	H9′,H8′	H9′,H7′	H9′,H6′
1	1.00	0.76	0.96

for $\mathbf{C}_{SSRS}$, the ROESY crosspeaks H3′pro*R*,H6′ and H3′pro*S*,H6′ are expected to be much weaker than the ROESY crosspeaks H3′pro*R*,H5′ and H3′pro*S*,H5′ (Figure 8B). This is not in agreement with the ROESY crosspeaks in Figure 5, and therefore $\mathbf{C}_{SSRS}$ could be rejected.

In conclusion, the absolute configurations at the chiral centers in the C_9 moiety of **1** (1′,4′-lactone) are determined to be 2′*S*,4′*S*,5′*R*,6′*R*,7′*S*,8′*R*. Thus, compound **1** is the lactonized form (C_9 moiety) of 6-*O*-(3′,9′-dideoxy-D-*threo*-D-*altro*-nononic acid-2′-yl)-D-glycopyranose. A low-energy conformation of the methyl glycoside of **1**, together with intraresidual ROEs, is presented in Figure 9. The nononic acid linked

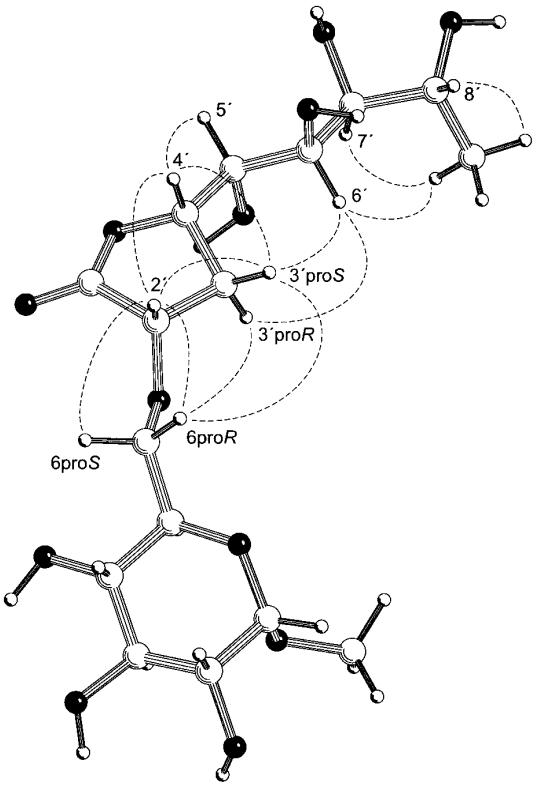


Figure 9. Snap shot from a MD simulation of the methyl α-glycoside of compound **1**. Dotted lines indicate important ROEs observed in the 2D ROESY spectrum. Aliphatic protons were reattached.

through an ether-bond to a monosaccharide presents a novel constituent of naturally occurring polysaccharides. Detailed analysis of the complete EPS repeating unit will be reported elsewhere.

Experimental Section

Culture conditions of the microorganism and isolation of the exopolysaccharide: *S. thermophilus* 8S, obtained from NIZO food research (Ede, The Netherlands), was cultivated in reconstituted skimmed milk, and the EPS was isolated and purified as described previously.[1]

Isolation of the unknown sugar component: The exopolysaccharide (5 mg) was hydrolyzed with trifluoroacetic acid (2 M, 2 h, 100 °C), and the obtained mixture lyophilized. The residue was dissolved in water and fractionated on graphitized carbon.[10] The unknown component, eluted with 40 % acetonitrile that contained 0.05 % trifluoroacetic acid, was lyophilized prior to analysis. A similar protocol was followed for the carboxyl-reduced ($NaBH_4$) polysaccharide.

Carboxyl reduction: Carboxyl reduction of the EPS was performed as described previously.[11] A solution of polysaccharide (2 mg) in 2-(4-morpholino)-ethanesulfonic acid (0.2 M, 1 mL, pH 4.75), containing *N*-ethyl-*N*-(3-dimethylaminopropyl)carbodiimidehydrochloride (30 mg), was stirred for 90 min at room temperature. After reduction with $NaBD_4$ (10 mg, 1 h), the obtained material was neutralized with HCl (1.5 M), desalted on graphitized carbon, and lyophilized prior to analysis. To obtain a complete carboxyl reduction the procedure was repeated twice.

Ether bond cleavage: The unknown sugar component was isolated from the reduced EPS as described. Freshly distilled BBr_3 was added to a cooled solution of the isolated sugar component in dry dichloromethane (1 mg mL^{-1}, −80 °C). The mixture was kept at −80 °C for 30 min, then slowly brought to room temperature. After 18 h, the excess of reagent was decomposed with water, the solution evaporated to dryness, and boric acid removed by coevaporation with methanol.

Monosaccharide and methylation analysis: Monosaccharide analysis, including the determination of the absolute configurations, and methylation analysis were performed as described previously.[1] Obtained derivatives were analyzed by GLC and GLC-EIMS as described previously.[1]

Preparation of deoxy-alditols: 1-Deoxy-D-galactitol, 1-deoxy-L-mannitol, and 6-deoxy-D-glucitol (1-deoxy-L-gulitol) were prepared by borohydride reduction of L-fucose, L-rhamnose, and D-quinovose, respectively.

1-Deoxy-L-glucitol, 1-deoxy-L-iditol, and 1-deoxy-D-altritol: These compounds were prepared[12] from L-glucose, L-idose, and D-altrose, respectively, as follows: The aldose (100 mg) was dissolved in fuming hydrochloric acid (2 mL), cooled to 0 °C on an ice/water bath, and ethanethiol (2 mL) was added. After 1 h, the reaction mixture was slowly poured into a saturated sodium hydrogen carbonate solution. Solid sodium hydrogen carbonate was carefully added until an alkaline solution was obtained. The aqueous solution was extracted with ethyl acetate (5 × 25 mL). The combined organic layers were dried (Na_2SO_4), filtered, and concentrated. The aldose diethyl dithioacetal was dissolved in pyridine (7.5 mL) and acetic anhydride (7.5 mL) was added. The reaction mixture was stirred overnight, then concentrated and co-concentrated with toluene, ethanol, and finally dichloromethane. The diethyl dithioacetal pentaacetate was purified by vacuum-line chromatography (dichloromethane/acetone 99:1).

The aldose diethyl dithioacetal pentaacetate was boiled under reflux in ethanol for 16 h with approximately four times its weight of Raney nickel. The Raney nickel was filtered off and washed carefully with methanol and ethyl acetate. The organic phase was concentrated and re-*O*-acetylated with pyridine (5 mL) and acetic anhydride (5 mL). After 2 h, the reaction mixture was concentrated and co-concentrated with toluene, ethanol, and finally dichloromethane. The 1-deoxyalditol pentaacetate was purified by vacuum-line chromatography (dichloromethane/ethyl acetate 98:2). The pure compound was dissolved in methanol and sodium methoxide was added until a pH of 9 was reached. The reaction mixture was stirred for 2 h, then Dowex H^+ was added until a neutral pH was obtained. The reaction mixture was filtered and concentrated.

Mass spectrometry: Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) experiments were performed as described previously.[1]

NMR spectroscopy: One- and two-dimensional NMR spectra were recorded in D_2O on a Bruker AMX500 spectrometer (Bijvoet Center, Department of NMR Spectroscopy). NMR spectra were recorded and processed as described.[1] The two-dimensional off-resonance ROESY

0947-6539/02/0819-4504 $ 20.00+.50/0 *Chem. Eur. J.* **2002**, *8*, No. 19

spectrum was recorded with an adiabatic spin-lock pulse of 350 ms at a field strength corresponding to 6.1 kHz. The spin-lock frequency was alternatively placed 3520 Hz upfield or downfield of the centre of the spectrum, thus obtaining an average spin-lock angle ν of 60°. Proton distances were estimated from crosspeak intensities in the two-dimensional ROESY spectrum and were calibrated to distances on the basis of distances of the geminal methylene protons H6pro*R*,H6pro*S* and H3′pro*R*,H3′pro*S* (1.8 Å). Spectral simulations were performed with in-house software based on the LAOCOON program.[13] Vicinal and geminal coupling constants were taken positively and negatively, respectively.

Computer-generated structures: Theoretical *R*/*S* configurations of the C_9 moiety of compound **1** were generated by computer. To this end, the Glc residue was replaced by an *O*-methyl group. All *R*/*S* configurations for the chiral C5′,6′,7′,8′ atoms were generated with dihedral angles (H_{n-1}-C_{n-1}-C_n-H_n) in agreement with the vicinal coupling constants ($+60°$ and $-60°$ for $J_{n-1,n} < 3$, 180° for $J_{n-1,n} > 7$) and tested for compliance with NOE-derived distance constraints. All possible combinations of dihedral angles $\pm 10°$ were also evaluated.

Adaptive umbrella sampling of the potential-of-mean-force: Potential-of-mean-force (PMF) calculations[9] were performed on theoretical *R*/*S* configurations of compound **1** or the C_9 moiety of compound **1**, with the method of adaptive umbrella sampling. The calculations were performed with the GROMOS program[14] with the standard force field for carbohydrates.[15] All simulations were divided into jobs of 10 ps. Dihedral angle values were partitioned into 72 classes, each having a width of 5°. The derivative of the PMF was evaluated as a 12-term Fourier series. Each system was simulated for 5–10 ns, and the final PMF for each system was used to obtain the rotamer population distributions of the sampled dihedral angle.

Molecular dynamics simulations: MD simulations in water were performed with the GROMOS program[14] with the standard force field for carbohydrates.[15] Molecules were placed in a truncated octahedron with periodic boundary conditions containing approximately 500 water molecules by means of the SPC/E model.[16] All bond lengths were kept fixed by means of the SHAKE procedure.[17] Simulations were performed at constant temperature (300 K) and pressure (1 atm) with relaxation times of 0.1 and 0.5 ps, respectively. For the simulations, a cut-off radius of 0.8 nm, a time step of 2 fs, and a total simulation time of 5 ns were used.

Acknowledgements

This work was supported by the PBTS Research Program with financial aid from the Ministry of Economic Affairs and by the Integral Structure Plan for the Northern Netherlands from the Dutch Development Company. The authors thank F. Kingma (NIZO food research, Ede, The Netherlands) for cultivation of *S. thermophilus* 8S, Dr. G. J. Gerwig for borohydride reduction to obtain three reference compounds, Dr. B. R. Leeflang for recording the NMR spectra of the reference compounds, and Dr. S. R. Haseley for stimulating discussions.

[1] E. J. Faber, J. P. Kamerling, J. F. G. Vliegenthart, *Carbohydr. Res.* **2001**, *331*, 183–194.
[2] Y. Nishida, H. Ohrui, H. Meguro, *Tetrahedron Lett.* **1984**, *25*, 1575–1578.
[3] L. Poppe, *J. Am. Chem. Soc.* **1993**, *115*, 8421–8426.
[4] K. Bock, C. Pedersen, *Adv. Carbohydr. Chem. Biochem.* **1983**, *41*, 27–66.
[5] C. A. G. Haasnoot, F. A. A. M. de Leeuw, C. Altona, *Tetrahedron* **1980**, *36*, 2783–2792.
[6] H. M. Berman, *Acta Crystallogr. Sect. B* **1970**, *26*, 290–299.
[7] J. A. Kanters, G. Roelofsen, H. M. Doesburg, T. Koops, *Acta Crystallogr. Sect. B* **1976**, *32*, 2830–2837.
[8] A. Imberty, S. Pérez, *Carbohydr. Res.* **1988**, *181*, 41–55.
[9] R. W. W. Hooft, B. P. van Eijck, J. Kroon, *J. Chem. Phys.* **1992**, *97*, 6690–6694.
[10] N. H. Packer, M. A. Lawson, D. R. Jardine, J. W. Redmond, *Glycoconjugate J.* **1998**, *15*, 737–747.
[11] R. L. Taylor, H. E. Conrad, *Biochemistry* **1972**, *11*, 1383–1388.
[12] D. T. Williams, J. K. N. Jones, *Can. J. Chem.* **1966**, *44*, 412–415.
[13] S. Castellano, A. A. Bothner-By, *J. Chem. Phys.* **1964**, *41*, 3863–3869.
[14] W. F. van Gunsteren, GROMOS, Groningen Molecular Simulation Package; University of Groningen (The Netherlands), **1987**.
[15] S. A. H. Spieser, J. A. van Kuik, L. M. J. Kroon-Batenburg, J. Kroon, *Carbohydr. Res.* **1999**, *322*, 264–273.
[16] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, *Intermolecular Forces*, Reidel, Dordrecht, **1981**, pp. 331–342.
[17] J. P. Ryckaert, G. Giccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.

Received: March 5, 2002 [F3927]

  0947-6539/02/0819-4505 $ 20.00+.50/0