Eur. J. Biochem. *205*, 799–808 (1992)

Novel oligosaccharide constituents of the cellulase complex of *Bacteroides cellulosolvens*

Gerrit J. GERWIG[1], Johannis P. KAMERLING[1], Johannes F. G. VLIEGENTHART[1], Ely MORAG[2], Raphael LAMED[2] and Edward A. BAYER[3]

[1] Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, The Netherlands
[2] Department of Biotechnology, The George S. Wise Faculty of Life Sciences, Tel Aviv University, Israel
[3] Department of Biophysics, The Weizmann Institute of Science, Rehovot, Israel

(Received November 18, 1991) – EJB 91 1547

The multiple cellulase-containing protein complex, isolated from the cellulolytic bacterium *Bacteroides cellulosolvens*, contains oligosaccharides which are O-linked mainly to a 230-kDa subunit. The oligosaccharide chains were liberated by alkaline-borohydride treatment and fractionated as oligosaccharide alditols via gel-permeation chromatography and HPLC. The fractions were investigated by one- and two-dimensional (correlation, homonuclear Hartmann-Hahn, rotating-frame nuclear Overhauser enhancement) 500-MHz ^{1}H-NMR spectroscopy in combination with monosaccharide and methylation analyses and with fast-atom-bombardment mass spectrometry. The following carbohydrate structures could be established:

D-Gal*f*-β(1 → 3)
\
D-Gal*f*-α(1 → 3)-D-Glc*p*NAc-α(1 → 2)-D-Gal*f*-α(1 → 2)-D-GalOH (**1**)

D-Gal*f*-β(1 → 3)
\
D-Glc*p*NAc-α(1 → 2)-D-Gal*f*-α(1 → 2)-D-GalOH (**2**)

D-Glc*p*NAc-α(1 → 2)-D-Gal*f*-α(1 → 2)-D-GalOH (**3**)

D-Gal*f*-β(1 → 3)-D-GalOH (**4**)

D-Gal*f*-α(1 → 2)-D-GalOH (**5**) and D-GalOH (**6**).

The results indicate an interesting similarity between the oligosaccharide moieties of the cellulase complex of *B. cellulosolvens* and of *Clostridium thermocellum* [Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F. G., Morag (Morgenstern), E., Lamed, R. & Bayer, E. A. (1991) *Eur. J. Biochem. 196*, 115–122], having **3**, **5** and **6** as common elements.

The furanose form of a terminal α-D-galactose residue demonstrated an inhibitory effect on the interaction of *Griffonia simplicifolia* I isolectin B4 with the cellulosome-like entity of *B. cellulosolvens*.

The study of the production of cellulolytic enzymes and their utilization in the hydrolysis of cellulosic materials are considered to be of practical importance in the search for alternative energy sources and renewable sources of food or chemicals. In this context, knowledge of the biochemical nature of the cellulases should eventually prove useful for the design of processes that will lead to the efficient degradation of cellulose.

One of the cellulolytic bacteria which is known for the production of high levels of cellulases is the Gram-positive, thermophilic anaerobe, *Clostridium thermocellum*, belonging to the family of *Bacillaceae*. Its extracellular and cell-associated multicellulase complex, termed the cellulosome, has been subject of intensive study by our groups [1, 2]. One of its remarkable features in the association of carbohydrate with at least one of the cellulosomal subunits. The primary structure of these novel O-linked carbohydrate chains has recently been elucidated [3, 4], but their role in cellulosomal functioning is still unknown.

Correspondence to J. F. G. Vliegenthart, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P.O. Box 80.075, NL-3508 TB Utrecht, The Netherlands

Abbreviations. FAB-MS, fast-atom-bombardment mass spectrometry; DQF^{1}H–^{1}H COSY, double-quantum-filtered ^{1}H–^{1}H correlation spectroscopy; HOHAHA, homonuclear Hartmann-Hahn spectroscopy; ROESY, rotating-frame nuclear Overhauser enhancement spectroscopy; GalOH, galactitol; GSI-B_4, *Griffonia simplicifolia* I isolectin B_4; 2D, two-dimensional; PNGase-F, peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F.

Enzyme. Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (EC 3.5.1.52).

Comparative lectin-binding experiments have suggested the occurrence of similar carbohydrate structures in the cellulase systems of other cellulolytic bacteria [5, 6]. Nearly all cellulose-degrading strains appear to exhibit a strong interaction with the α-Gal-specific lectin from *Griffonia simplicifolia*. Another feature shared by most of the cellulolytic strains tested, is the immunochemical cross-reactivity with anti-(*C. thermocellum* cellulosome) antibodies.

One of these strains, *Bacteroides cellulosolvens*, a Gram-negative, non-spore-forming, mesophilic anaerobe, belonging to the family of *Bacteroidaceae*, is claimed to express cellulose-degrading capacity, equal to that of *C. thermocellum* [7, 8]. *B. cellulosolvens* also produces extracellular and cell-associated cellulases which possess endoglucanase, exoglucanase and xylanase activities [9]. The majority of cellulase activity in this bacterium appears to be cell-associated in the form of exocellular protuberances, probably similar to those of *C. thermocellum* [6, 10, 11], and the cellulases are organized in a multicomponent cellulosome-like complex. In view of the interaction with the α-Gal-specific lectin, in conjunction with the other properties shared by the cellulase systems of *C. thermocellum* and *B. cellulosolvens*, we were interested in determining the extent of similarity in their cellulosome-borne oligosaccharide moieties. In this paper, the finding of novel O-linked carbohydrate chains in the cellulase complex of *B. cellulosolvens* is described.

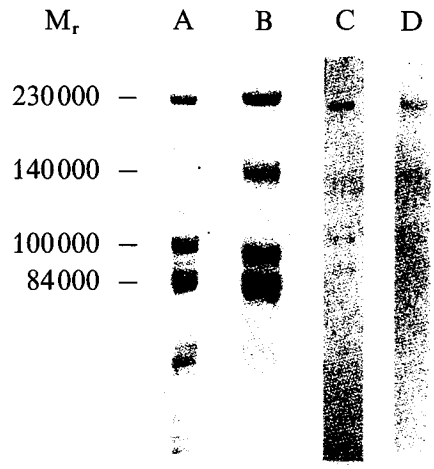


Fig. 1. SDS/PAGE pattern and saccharide distribution of the cellulosome-like entity in *B. cellulosolvens*. The cell-free growth culture (lane A) was treated with microcrystalline cellulose and the adsorbed fraction (lanes B–D) was eluted from the resin. Lane C, the separated material was blotted and labeled with GSI-B_4–peroxidase, the α-Gal-specific lectin-enzyme conjugate. Lane D was stained for carbohydrate using the periodate-Schiff's reagent. Lanes A and B were stained for protein using Coomassie brilliant blue R.

Table 1. Molar carbohydrate composition of the cellulase complex of *B. cellulosolvens* (A), alkaline-borohydride-treated cellulase complex (AB) and derived fractions. For the molar ratio GlcNAc was taken as 1.0 for A; for the other fractions GalOH was taken as 1.0. All samples were contaminated with a small amount of Glc (0.1–0.3). GalOH was corrected for anhydro-GalOH [23].

Monosaccharide	Carbohydrate content						
	A	AB	I.1	II.1	III.2	IV.1	IV.2
	mol/mol						
Gal	4.4	3.5	3.0	2.0	1.1	1.1	1.0
GalOH	–	1.0	1.0	1.0	1.0	1.0	1.0
GlcNAc	1.0	0.8	1.0	0.8	0.9	–	–

EXPERIMENTAL PROCEDURES

General methods

Monosaccharide analysis, absolute configuration determination, SDS/PAGE, alkaline-borohydride treatment, methylation analysis, HPLC, fast-atom-bombardment mass spectrometry (FAB-MS) and 500-MHz ^{1}H-NMR spectroscopy were performed as described previously [3, 4].

Preparation of the cellulase complex from *B. cellulosolvens*

The method employed was essentially identical to that described for the purification of the cellulosome of *C. thermocellum* [2]. *B. cellulosolvens* (SPB 25) NRCC 2944] cells were grown for 24 h on cellobiose medium [8].

Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (PNGase-F) treatment

The cellulase complex (20 µl, 2 mg/ml) was boiled for 3 min in aqueous 0.5% SDS, containing 0.1 M 2-mercaptoethanol. The sample was diluted fourfold with 0.2 M sodium phosphate buffer (pH 7.2), containing 10 mM 1,10-phenanthroline hydrate and 1.25% Nonidet P-40. An aliquot (5 µl) of *N*-glycanase from *Flavobacterium meningosepticum* (250 U/ml, Genzyme Corp., Boston, MA) was added and the reaction was allowed to proceed overnight at 37 °C. A sample of fetuin was used as a positive control. Initially, it was tried to analyze the incubation mixture by SDS/PAGE as described earlier for the cellulosome of *C. thermocellum* [3], but the glycosylated cellulosome-like complex in *B. cellulosolvens* tended to aggregate following exposure to the above incubation conditions. Therefore, to test the extent of deglycosylation, a dot-blot assay system was designed, which proved efficient for the fetuin control. Using this assay, a lectin-enzyme conjugate, consisting of the α-Gal-specific isolectin B_4 from *Griffonia simplicifolia* (GSI-B_4) coupled to horseradish peroxidase (GSI-B_4–peroxidase, Sigma), was used to label protein-bound carbohydrate. Samples (0.5 µl) were blotted in twofold dilutions onto nitrocellulose membranes, and the latter were blocked for 30 min at 23 °C with 2% (mass/vol.) bovine serum albumin in 0.15 M NaCl and 10 mM potassium phosphate, pH 7.2 (NaCl/P_i). The membrane-immobilized samples were then incubated for 1 h with 2 µg/ml GSI-B_4 peroxidase, washed three times with saline, and color was developed using 4-chloro-1-naphthol in the presence of hydrogen peroxidase as described previously [12].

Specificity of GSI-B_4

Samples of the purified *B. cellulosolvens* cellulase complex were dotted onto nitrocellulose paper and blocked for 30 min with bovine serum albumin dissolved in NaCl/P_i. The methyl glycosides of α-D-Gal*f*, β-D-Gal*f*, and α-D-Gal*p*, at a concentration of 7 µg in 1 µl NaCl/P_i, were each combined with 0.1 µg GSI-B_4–peroxidase in 1 µl NaCl/P_i, incubated for 60 min at 23 °C and 2-µl aliquots were layered individually over the dotted samples in a moist chamber. After 60 min, the

nitrocellulose membranes were washed with $NaCl/P_i$, and the colored precipitate was developed as described above.

RESULTS

As reported earlier [12], affinity purification of the cellulosome-like entity from *B. cellulosolvens* resulted in four major components which exhibited molecular masses of 230, 140, 100 and 84 kDa and in some minor bands (Fig. 1). Most of the carbohydrate appeared to be associated with the 230-kDa component, as demonstrated by staining procedures based on the use of GSI-B_4 – peroxidase and periodate/Schiff (Fig. 1C and D, respectively). The other major bands also contained carbohydrate, but in much lower concentrations. No detectable amounts of carbohydrate could be split off from the cellulosome-like entity of *B. cellulosolvens* by PNGase-F, suggesting the absence of N-linked carbohydrate chains in the complex, similar to the observation for the cellulosome of *C. thermocellum* [3].

Monosaccharide analysis [13] of the cellulase complex showed a total carbohydrate content of 4.2% (by mass) with Gal and GlcNAc as the constituents (Table 1A). Determination of the absolute configuration of the monosaccharides as their (–)-2-butyl glycosides showed the D configuration in each case [14, 15]. Monosaccharide analysis of alkaline-borohydride-treated cellulase complex (Table 1A and B) afforded besides Gal and GlcNAc, also galactitol (GalOH), due to conversion of Gal, which is involved in the linkage to the protein. The alkaline-borohydride-treated material was separated on Bio-Gel P-2 (Fig. 2), and five carbohydrate-containing fractions (I – V) were pooled. The preceding fractions contained minor amounts of carbohydrate, and monosaccharide analysis revealed a composition identical to that of the cellulase complex, indicating the presence of residual protein-bound carbohydrate.

Fraction I was subfractionated by HPLC on Lichrosorb-NH_2 (Fig. 3A), yielding two major peaks, denoted I.1 and II.1. The latter peak originated from overlap of Bio-Gel fraction I with fraction II, and will be discussed below. Monosaccharide analysis of I.1 suggested the presence of a pentasaccharide alditol consisting of GalOH, GlcNAc and three Gal residues (Table 1). Positive FAB-MS of underivatized I.1 (spectrum not depicted) showed the presence of pseudo-molecular ions $(M + H)^+$ and $(M + Na)^+$ at *m/z* 872 and at *m/z* 894, respectively, pointing to a molecular mass of 871 Da, in accordance with the assumed pentasaccharide alditol. Further support for such a compound was obtained from the positive FAB mass spectrum of permethylated I.1 (spectrum not depicted), giving rise to a pseudo-molecular ion $(M + H)^+$ at *m/z* 1124. The presence of intense fragment ions at *m/z* 668 (Gal_2/GlcNAc) and *m/z* 464 (Gal/GlcNAc), together with the absence of *m/z* 260, indicated the occurrence of the structural element Gal → GlcNAc → Gal →.

Methylation analysis, including reduction with NaB^2H_4 of I.1, yielded the partially methylated alditol acetates presented in Table 2. Non-^2H-labeled 1,4,5,6-tetra-*O*-methyl-2,3-di-*O*-acetylgalactitol represented the original GalOH residue substituted at C-2 and C-3. Based on the finding of two terminal Gal*f* residues, of one Gal*f* residue substituted at C-2, and of one Glc*p*NAc residue substituted at C-3, together with the FAB-MS data (see above), the pentasaccharide alditol Gal*f*-(1 → 3)-Glc*p*NAc-(1 → 2)-Gal*f*-(1 → 2/3)-[Gal*f*-(1 → 3/2)-]GalOH can be proposed.

500-MHz ^1H-NMR spectroscopy was applied to verify and complete the monosaccharide sequence, including type

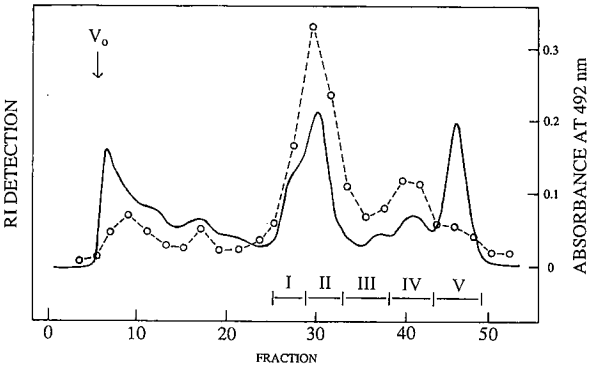


Fig. 2. Elution profile on Bio-Gel P-2 of alkaline-borohydride-treated cellulase complex of *B. cellulosolvens*. The column (95 cm × 1 cm) was eluted with double-distilled water. Fractions of 1.0 ml were collected at a flow rate of 6.0 ml/h. The eluate was monitored by refractive-index (RI) detection (——) and by hexose determination with phenol/sulfuric acid (○) [16]. Fractions were pooled as indicated. V_o, void volume.

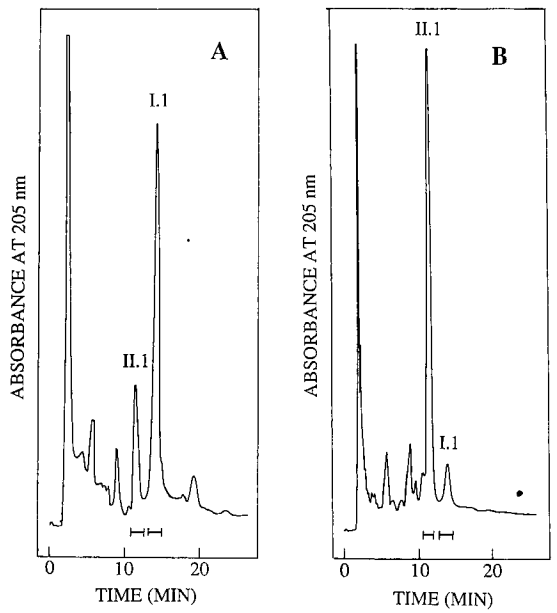


Fig. 3. HPLC elution profile on Lichrosorb-NH_2 of fraction I (A) and fraction II (B) derived from *B. cellulosolvens* cellulase complex. The column (250 mm × 4.6 mm) was run isocratically with a mixture of acetonitrile/water (75:25, by vol.) at a flow rate of 2.0 ml/min. The eluate was monitored at 205 nm. Fractions I.1 and II.1 were collected.

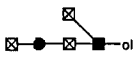
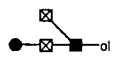
and configuration of the glycosidic linkages. The one-dimensional ^1H-NMR spectrum of I.1 is presented in Fig. 4A, and NMR data are listed in Table 3. The assignment of the signals is based on two-dimensional double-quantum-filtered ^1H-^1H correlation spectroscopy (2D DQF ^1H-^1H COSY; Fig. 5) and two-dimensional homonuclear Hartmann-Hahn spectroscopy (2D HOHAHA; Fig. 6) experiments. These data are essential for the interpretation of the two-dimensional rotating-frame nuclear Overhauser enhancement spectroscopy (2D ROESY) spectrum (Fig. 7). For the coding system of the different residues, see Fig. 4A.

The coupling constant of terminal Gal*f*-**e** H-1 at $\delta = 5.309$, being $J_{1,2} = 4.2$ Hz, indicated α configuration. This was confirmed by the observation of an intra-residual NOE from

Table 2. Methylation analysis of derived fractions from alkaline-borohydride-treated cellulase complex of *B. cellulosolvens*. GLC retention times (t_R) on SE-30 relative to 2,3,4,6-tetra-*O*-methyl-1,5-di-*O*-acetylgalactitol. The molar ratio was calculated from peak areas, not corrected by response factors

Partially methylated alditol acetate	t_R	Carbohydrate content			Structural features
		I.1	II.1	III.2	
		mol/mol			
1,3,4,5,6-Penta-*O*-methyl-2-mono-*O*-acetylgalactitol	0.69	–	–	1.0	→2)-GalOH
1,4,5,6-Tetra-*O*-methyl-2,3-di-*O*-acetylgalactitol	0.85	0.7	0.7	–	→2,3)-GalOH
2,3,5,6-Tetra-*O*-methyl-1,4-di-*O*-acetylgalactitol-(1-^{2}H)	0.97	2.0	1.0	–	Gal*f*-(1→
3,5,6-Tri-*O*-methyl-1,2,4-tri-*O*-acetylgalactitol-(1-^{2}H)	1.19	0.6	0.6	0.7	→2)-Gal*f*-(1→
2-(*N*-Methyl)acetamido-2-deoxy-3,4,6-tri-*O*-methyl-1,5-di-*O*-acetylglucitol-(1-^{2}H)	1.64	–	0.7	0.7	Glc*p*NAc-(1→
2-(*N*-Methyl)acetamido-2-deoxy-4,6-di-*O*-methyl-1,3,5-tri-*O*-acetylglucitol-(1-^{2}H)	1.91	0.6	–	–	→3)-Glc*p*NAc-(1→

Table 3. Relevant ^{1}H-NMR parameters of structural-reporter groups of constituent monosaccharides for oligosaccharide alditols from *B. cellulosolvens* cellulase complex. In the table heading, the structures are represented by (⊠) Gal*f*, (●) Glc*p*NAc and (■-ol) GalOH; n. d., not determined. Some coupling constants are given in brackets.

Residue	Reporter group	Chemical shift			
		1 	2 	3 	4 
		ppm (Hz)			
GalOH (**a**)	H-1a,b	≈3.76	≈3.75	≈3.79	≈3.79
	H-2	4.146	4.143	4.051	4.050
	H-3	3.990	3.991	n.d.	3.950
	H-4	3.908	3.902	n.d.	3.823
	H-5	4.043	4.043	3.975	4.145
	H-6a,b	3.696	3.696	3.696	3.69
Gal*f* (**b**)	H-1 ($J_{1,2}$)	5.378 (4.7)	5.374 (4.7)	5.400 (4.8)	–
	H-2 ($J_{2,3}$)	4.212 (8.7)	4.204 (8.7)	4.221 (8.9)	–
	H-3 ($J_{3,4}$)	4.350 (8.1)	4.334 (8.1)	4.350 (8.5)	–
	H-4	3.793	3.783	n.d.	–
Gal*f* (**c**)	H-1 ($J_{1,2}$)	5.127 (2.6)	5.132 (2.5)	–	5.120 (2.0)
	H-2	4.132	4.130	–	4.130
	H-3	4.092	4.091	–	4.060
	H-4	4.042	4.042	–	4.000
Glc*p*NAc (**d**)	H-1 ($J_{1,2}$)	5.092 (3.6)	5.068 (3.6)	5.026 (3.5)	–
	H-2	4.153	4.000	n.d.	–
	H-3	3.960	3.812	n.d.	–
	H-4	3.750	3.547	3.547	–
	H-5	3.840	3.848	n.d.	–
	NAc	2.118	2.104	2.069	–
Gal*f* (**e**)	H-1 ($J_{1,2}$)	5.309 (4.2)	–	–	–
	H-2	4.100	–	–	–
	H-3	4.061	–	–	–
	H-4	3.730	–	–	–

Gal*f*-**e** H-1 to Gal*f*-**e** H-2 in the ROESY spectrum. The inter-residual NOE from Gal*f*-**e** H-1 to Glc*p*NAc-**d** H-3, marked **A** in Fig. 7, demonstrated the (1→3) linkage between terminal α-D-Gal*f*-**e** and α-D-Glc*p*NAc-**d**. The Glc*p*NAc-**d** residue had α configuration, as could be deduced from the $J_{1,2}$ of 3.6 Hz of the H-1 signal at $\delta = 5.092$ [compare with methyl α-D-Glc*p*NAc (H-1, $\delta = 4.748$, $J_{1,2} = 3.6$ Hz)]. The observed intra-residual NOE from Glc*p*NAc-**d** H-1 to Glc*p*NAc-**d** H-2 confirmed this conclusion. The inter-residual NOE from Glc*p*NAc-**d** H-1 to Gal*f*-**b** H-2, marked **B** in Fig. 7, was in accordance with a (1→2) linkage between α-D-Glc*p*NAc-**d** and α-D-Gal*f*-**b**. The α configuration of Gal*f*-**b** was derived

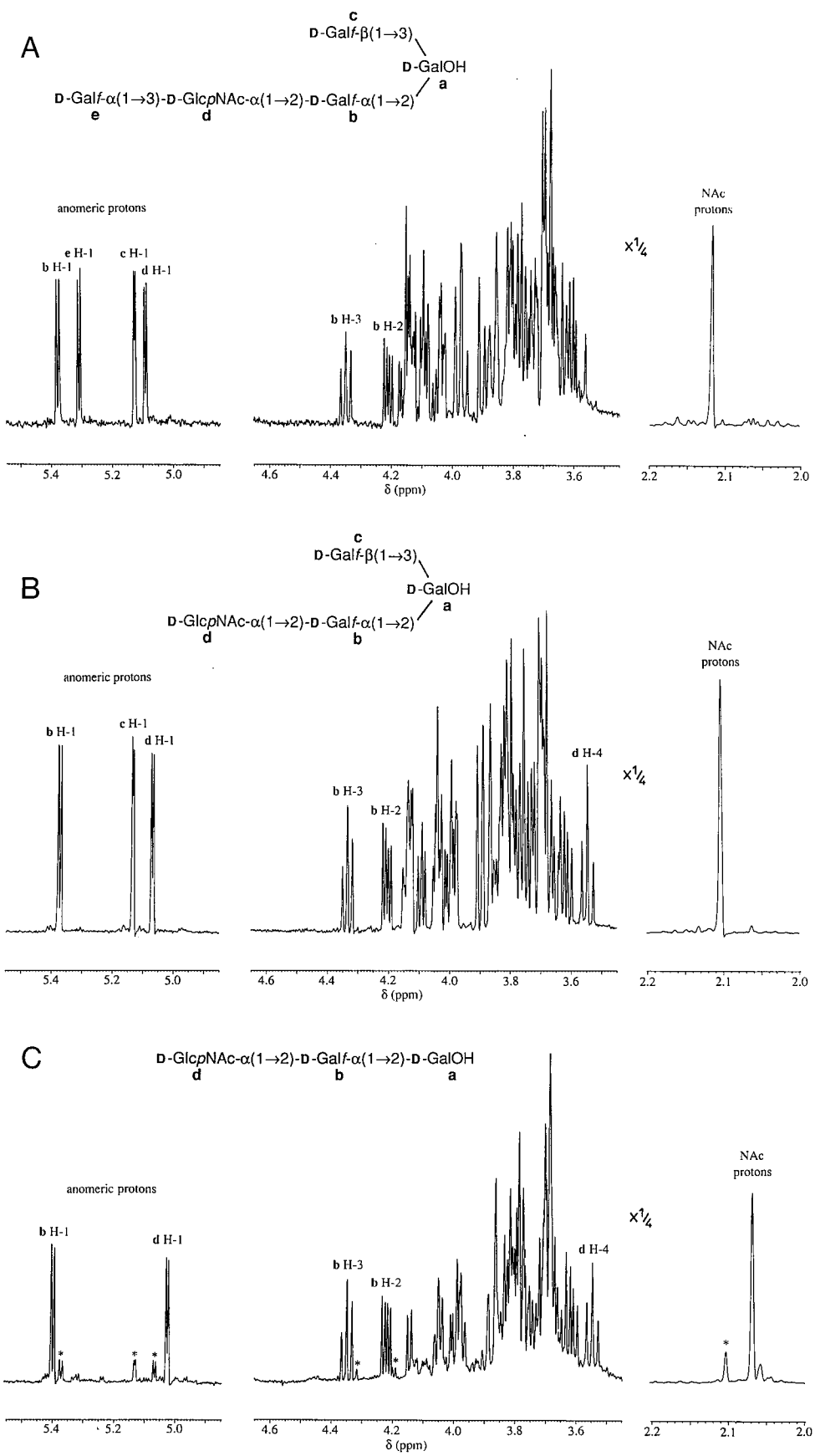


Fig. 4. 500-MHz ^{1}H-NMR spectrum of fraction I.1 (A), fraction II.1 (B) and fraction III.2 (C) derived from *B. cellulosolvens* cellulase complex recorded in ^2H_2O at 300 K. The assignment of the protons refer to the corresponding residues in the structure. The HO^{2}H signal (4.65–4.85 ppm) has been omitted. The relative intensity scale of the *N*-acetyl (NAc) region differs from that of the other part of the spectrum as indicated. (★) Stemming from B.

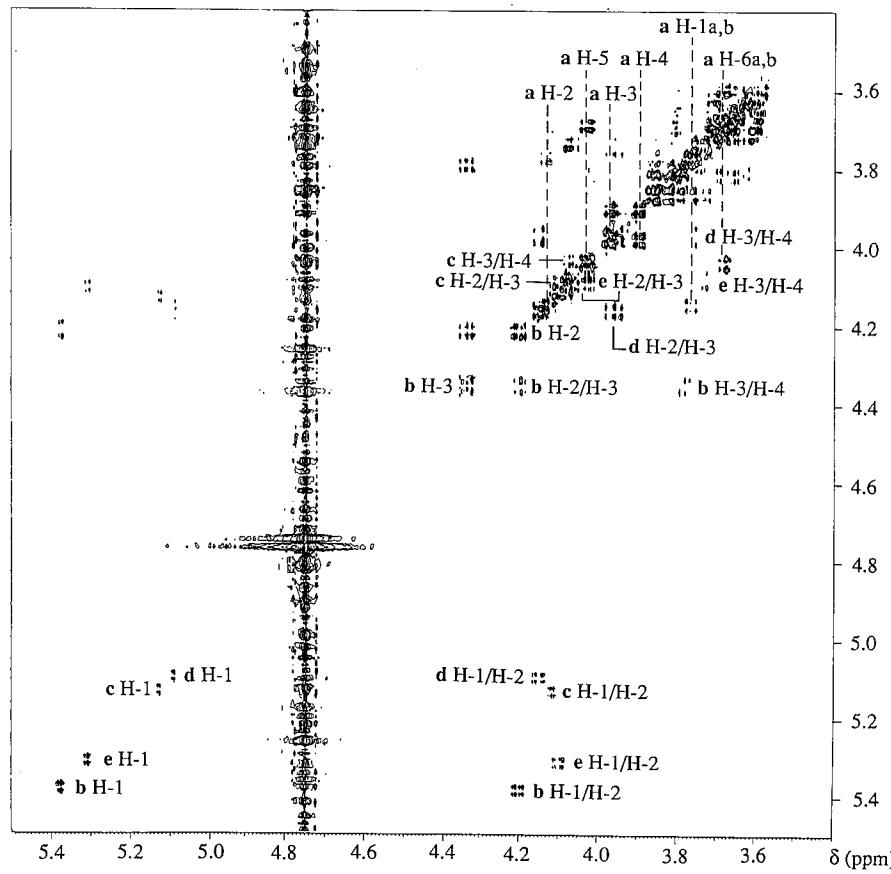


Fig. 5. 2D DQF $^1H-{}^1H$ COSY spectrum (region 3.4 – 5.5 ppm) of fraction I.1 derived from *B. cellulosolvens* cellulase complex recorded in ^2H_2O at 300 K. The assignment of the protons refer to the corresponding residues in the structure (Fig. 4A).

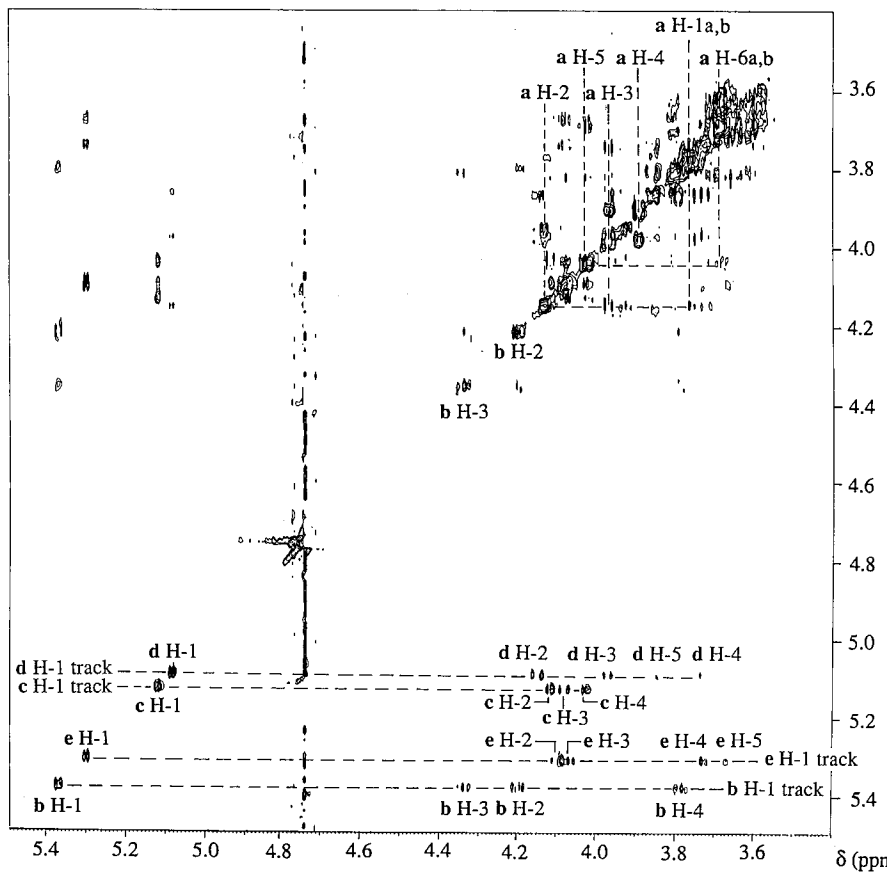


Fig. 6. 2D HOHAHA spectrum (region 3.4 – 5.5 ppm) of fraction I.1 derived from *B. cellulosolvens* cellulase complex recorded in ^2H_2O at 300 K. The assignment of the protons refer to the corresponding residues in the structure (Fig. 4A).

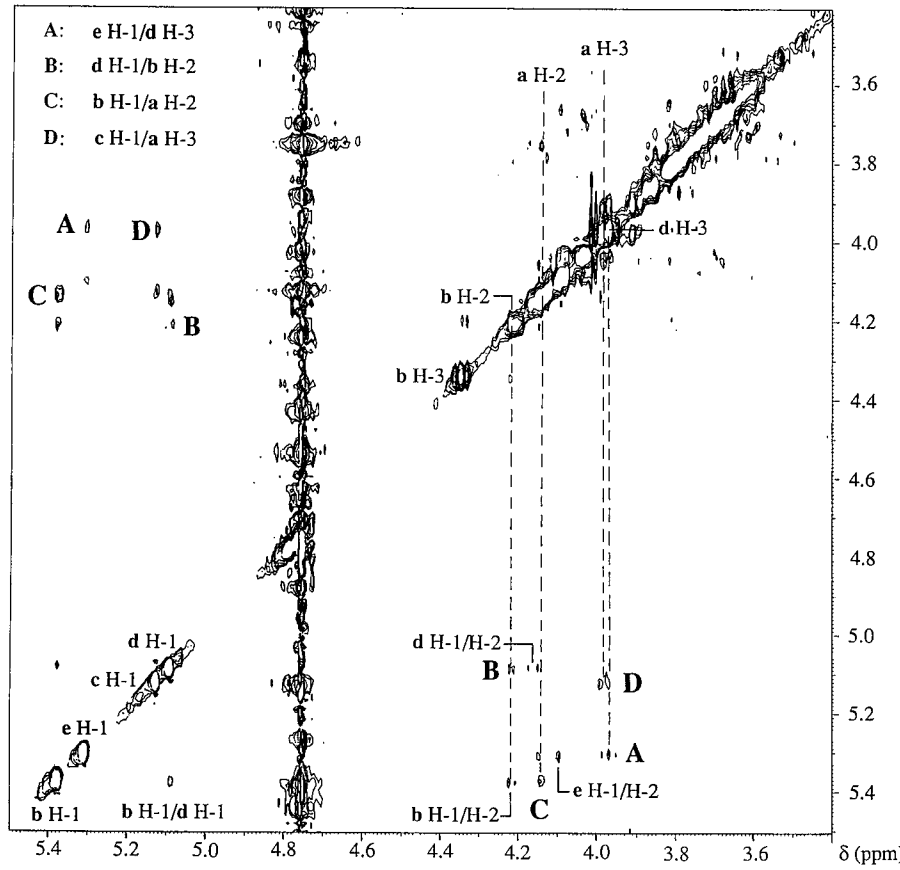


Fig. 7. 2D ROESY spectrum (region 3.4 – 5.5 ppm) of fraction I.1 derived from *B. cellulosolvens* cellulase complex recorded in ^2H_2O at 300 K. The assignments of the protons refer to the corresponding residues in the structure (Fig. 4A).

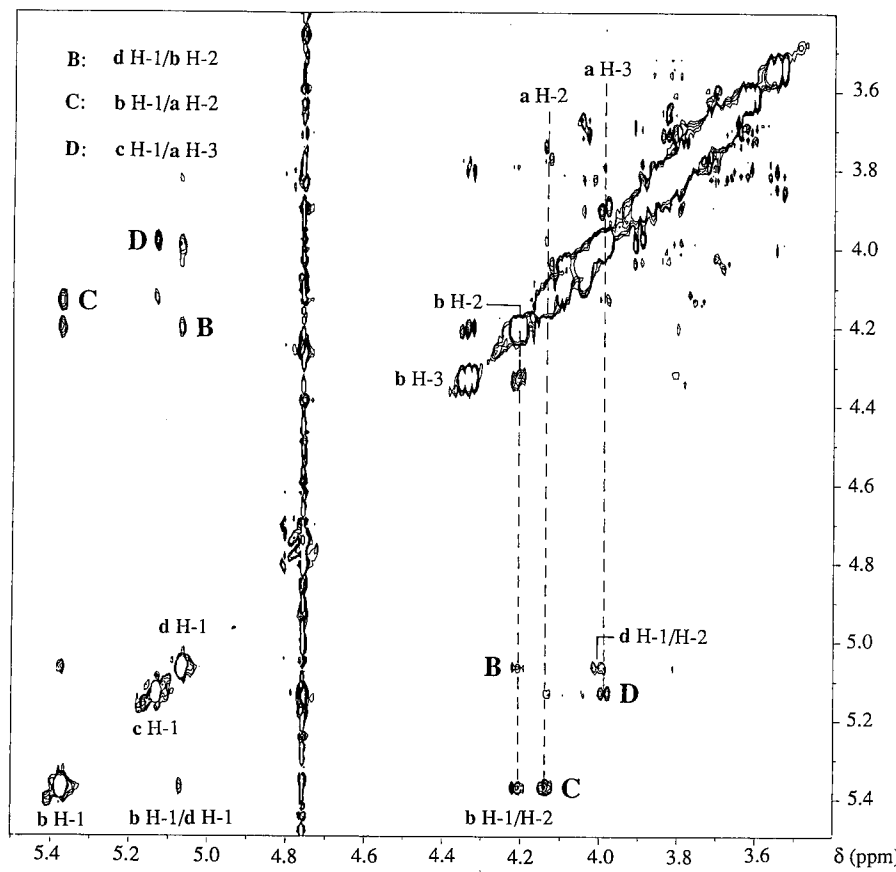
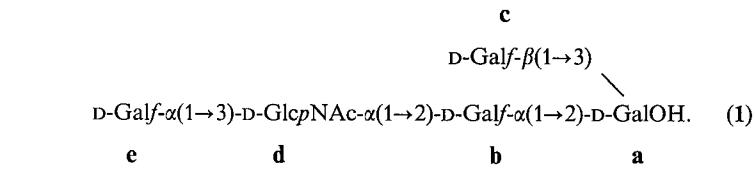


Fig. 8. 2D ROESY spectrum (region 3.4 – 5.5 ppm) of fraction II.1 derived from *B. cellulosolvens* cellulase complex recorded in ^2H_2O at 300 K. The assignments of the protons refer to the corresponding residues in the structure (Fig. 4B).

from a comparison of the coupling constant of the anomeric signal (H-1, $\delta = 5.378$, $J_{1,2} = 4.7$ Hz) with those of the anomeric signals of methyl α-D-Gal*f* (H-1, $\delta = 4.884$, $J_{1,2} = 4.0$ Hz) and methyl β-D-Gal*f* (H-1, $\delta = 4.908$, $J_{1,2} = 2.0$ Hz), supported by the intra-residual NOE from Gal*f*-**b** H-1 to Gal*f*-**b** H-2. The observed inter-residual NOE from Gal*f*-**b** H-1 to Glc*p*NAc-**d** H-1 is in agreement with Glc*p*NAc-**d** α-linked to Gal*f*-**b** at C-2. It is worthwhile to mention that the substitution at C-2 of α-D-Gal*f*-**b** gave rise to downfield positions outside the bulk signal for α-D-Gal*f*-**b** H-2 ($\delta = 4.212$) and H-3 ($\delta = 4.350$). The (1 → 2) linkage between α-D-Gal*f*-**b** and GalOH-**a** was indicated by an inter-residual NOE from Gal*f*-**b** H-1 to GalOH-**a** H-2, marked **C** in Fig. 7. The coupling constant of terminal Gal*f*-**c** H-1 at $\delta = 5.127$, being $J_{1,2} = 2.6$ Hz, proved β configuration. The β configuration of D-Gal*f*-**c** is confirmed by the absence of an intra-residual NOE from H-1 to H-2, in contrast to the presence of intense NOE from H-1 to H-2 for α-D-Gal*f*-**b** and α-D-Gal*f*-**e**. The observed inter-residual NOE from Gal*f*-**c** H-1 to GalOH-**a** H-3, marked **D** in Fig. 7, supported a (1 → 3) linkage between terminal β-D-Gal*f*-**c** and GalOH-**a**. The absence of inter-residual NOE from Gal*f*-**b** H-1 to GalOH-**a** H-3 and from Gal*f*-**c** H-1 to GalOH-**a** H-2 are in line with these findings. In conclusion, the data indicate the structure of the pentasaccharide alditol to be



Fraction II was subfractionated by HPLC on Lichrosorb-NH_2 (Fig. 3B), yielding one major peak denoted II.1. Monosaccharide analysis of II.1 suggested the presence of a tetrasaccharide alditol consisting of GalOH, GlcNAc and two Gal residues (Table 1). Positive FAB-MS of underivatized II.1 (spectrum not depicted) showed the presence of pseudo-molecular ions $(M + H)^+$ and $(M + Na)^+$ at *m/z* 710 and at *m/z* 732, respectively, in accordance with the assumed tetrasaccharide alditol with a molecular mass of 709 Da. The positive FAB mass spectrum of permethylated II.1 (spectrum not depicted) showed the pseudo-molecular ion $(M + H)^+$ at *m/z* 920 and fragment ions at *m/z* 464 and *m/z* 260, supporting the occurrence of the structural element GlcNAc → Gal →.

Methylation analysis, including reduction with NaB^2H_4 of II.1, yielded the partially methylated alditol acetates presented in Table 2. Non-^{2}H-labeled 1,4,5,6-tetra-*O*-methyl-2,3-di-*O*-acetylgalactitol represented the original GalOH residue substituted at C-2 and C-3. Based on the finding of terminal Gal*f* and Glc*p*NAc residues, and of a Gal*f* residue substituted at C-2, together with the FAB-MS data (see above), the tetrasaccharide alditol Glc*p*NAc-(1 → 2)-Gal*f*-(1 → 2/3)-[Gal*f*-(1 → 3/2)-]GalOH can be proposed.

The 1D ^{1}H-NMR spectrum of II.1 is presented in Fig. 4B, and NMR data are included in Table 3. For the assignment of essential resonances, 2D DQF ^{1}H-^{1}H COSY and D HOHAHA experiments were carried out (spectra not depicted), which enabled interpretation of the 2D ROESY spectrum (Fig. 8). For the coding system of the different residues, see Fig. 4B.

An inter-residual NOE from Glc*p*NAc-**d** H-1 to Gal*f*-**b** H-2, marked **B** in Fig. 8, could be detected, in accordance with a (1 → 2) linkage between terminal α-D-Glc*p*NAc-**d** (H-1, $\delta = 5.068$, $J_{1,2} = 3.6$ Hz) and α-D-Gal*f*-**b** (H-1, $\delta = 5.374$, $J_{1,2} = 4.7$ Hz). The (1 → 2) linkage between the α-D-Gal*f*-**b**

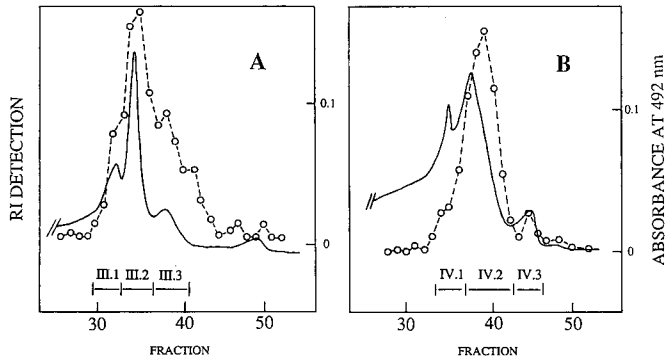
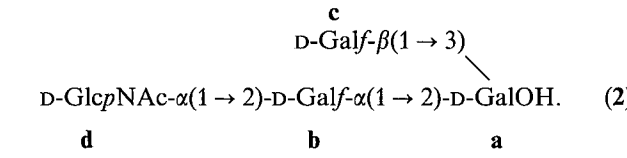


Fig. 9. Elution profile on Bio-Gel P-2 of fraction III (A) and fraction IV (B) derived from the cellulase complex of *B. cellulosolvens*. The column (95 cm × 1 cm) was eluted with double-distilled water. Fractions of 1.0 ml were collected at a flow rate of 6.0 ml/h. The eluate was monitored by refractive-index (RI) detection (——) and by hexose determination with phenol/sulfuric acid (○). Fractions were pooled as indicated.

and GalOH-**a** was demonstrated by an inter-residual NOE from Gal*f*-**b** H-1 to GalOH-**a** H-2, marked **C** in Fig. 8. The observed inter-residual NOE from Gal*f*-**c** H-1 to GalOH-**a** H-3, marked **D** in Fig. 8, pointed to a (1 → 3) linkage between terminal β-D-Gal*f*-**c** (H-1, $\delta = 5.132$, $J_{1,2} = 2.5$ Hz) and GalOH-**a**. The proximity of Gal*f*-**c** H-1 and GalOH-**a** H-2 causes a weak inter-residual NOE effect. Finally, the various intra-residual connectivities fit the proposed configuration of the individual monosaccharide residues (see fraction I.1). In conclusion, the data indicate the structure of the tetrasaccharide alditol to be



Fractions III from three identical, alkaline-borohydride-treated, cellulase complex samples were pooled and separated again on Bio-Gel P-2 (Fig. 9A), yielding three carbohydrate-containing fractions, denoted III.1, III.2 and III.3

^{1}H-NMR analysis showed that the compound in fraction III.1 was identical to that in fraction II.1. The material in fraction III.3 was too little for detailed structural analysis, but ^{1}H-NMR spectroscopy indicated the presence of the same compound as in fraction IV.1 (see below). Monosaccharide analysis of fraction III.2 showed the presence of GalOH, Gal and GlcNAc in a molar ratio of 1:1:1 (Table 1), suggesting the presence of a trisaccharide alditol. Positive FAB-MS of underivatized III.2 (spectrum not depicted) showed the presence of pseudo-molecular ions $(M + H)^+$ and $(M + Na)^+$ at *m/z* 548 and at *m/z* 570, respectively, in accordance with the assumed trisaccharide alditol (547 Da). The positve FAB mass spectrum of permethylated III.2 (spectrum not depicted) showed the pseudo-molecular ion $(M + H)^+$ at *m/z* 716, indicating a molecular mass of 715 Da, and sequence ions at *m/z* 464 and *m/z* 260, supporting the occurrence of the structural element GlcNAc → Gal →.

Methylation analysis, including reduction with NaB^2H_4 of III.2, yielded the partially methylated alditol acetates presented in Table 2. Non-^{2}H-labeled 1,3,4,5,6-penta-*O*-methyl-2-mono-*O*-acetylgalactitol represented the original GalOH

residue substituted at C-2. Furthermore, a Gal*f* residue substituted at C-2, and terminal Glc*p*NAc were detected.

The ^{1}H-NMR spectrum (Fig. 4C) showed two anomeric signals of equal intensity at $\delta = 5.400$ ($J_{1,2} = 4.8$ Hz) and $\delta = 5.026$ ($J_{1,2} = 3.5$ Hz), which can be assigned to α-D-Gal*f*-**b** H-1 and α-D-Glc*p*NAc-**d** H-1, respectively. Also the Gal*f*-**b** H-2 ($\delta = 4.221$, $J_{2,3} = 8.9$ Hz) and Gal*f*-**b** H-3 ($\delta = 4.350$, $J_{3,4} = 8.5$ Hz) signals are clearly observable outside the bulk signal. The singlet at $\delta = 2.069$ originates from the *N*-acetyl group of Glc*p*NAc. Based on these data, it can be concluded that the structure of III.2 is

D-Glc*p*NAc-α(1 → 2)-D-Gal*f*-α(1 → 2)-D-GalOH. (3)

d **b** **a**

Fractions IV from three identical, alkaline-borohydride-treated, cellulase complex samples were pooled and further separated on Bio-Gel P-2 (Fig. 9B), yielding three carbohydrate-containing fractions, denoted IV.1, IV.2 and IV.3. GLC-MS of permethylated IV.1 (Fig. 10), in conjunction with monosaccharide analysis (GalOH/Gal = 1:1, Table 1) and ^{1}H-NMR spectroscopy (β-Gal*f*-**c** H-1, $\delta = 5.120$, $J_{1,2} = 2.0$ Hz) proved the structure to be

D-Gal*f*-β(1 → 3)-D-GalOH. (4)

c **a**

The furanose form of Gal-**c** is supported by the ratio of the intensities of the mass spectrometric peaks (Fig. 10) at *m/z* 88 and *m/z* 101, being much less than one [4, 13, 17].

Monosaccharide analysis of IV.2 yielded also Gal and GalOH in a molar ratio of 1:1 (Table 1). Overlay of IV.2 and IV.1 could be expected from the elution pattern. The ^{1}H-NMR spectrum of IV.2 (not depicted) indeed showed a mixture of two compounds in equal amounts, being two disaccharide alditols, which could be identified as D-Gal*f*-β(1 → 3)-D-GalOH (see IV.1) and a compound of which the ^{1}H-NMR parameters completely matched [4] those of

D-Gal*f*-α(1 → 2)-D-GalOH. (5)

b **a**

Fraction IV.3 showed the presence of free GalOH, which was also present as component of fraction V (Fig. 2).

The relative amounts of the different compounds, released by alkaline-borohydride treatment, were calculated using hexose determination (phenol/sulfuric acid) and signal intensity ratios of anomeric and *N*-acetyl protons in ^{1}H-NMR spectra of the alditol mixtures. The major compound was the tetrasaccharide alditol (**2**), representing about 46% (by mass) of the total carbohydrate content. The pentasaccharide alditol (**1**), trisaccharide alditol (**3**), disaccharide alditols (**4, 5**) and galactitol (**6**) could be calculated to represent about 24%, 9%, 15% and 6% (by mass) of the total carbohydrate content, respectively.

Specificity of GSI-B$_4$

In view of the finding that galactose occurs only in the furanose form and not in the pyranose form in the terminal moieties of the oligosaccharide alditols, it was surprising that such a strong interaction occurred between GSI-B$_4$ and the cellulosome-like entity of *B. cellulosolvens*. GSI-B$_4$ is known to interact selectively with terminal α-Gal moieties [18, 19], but so far, studies have only been focused on galactopyranosides; the interaction of the lectin with terminal α- or β-galactofuranosyl moieties has not been reported. Therefore,

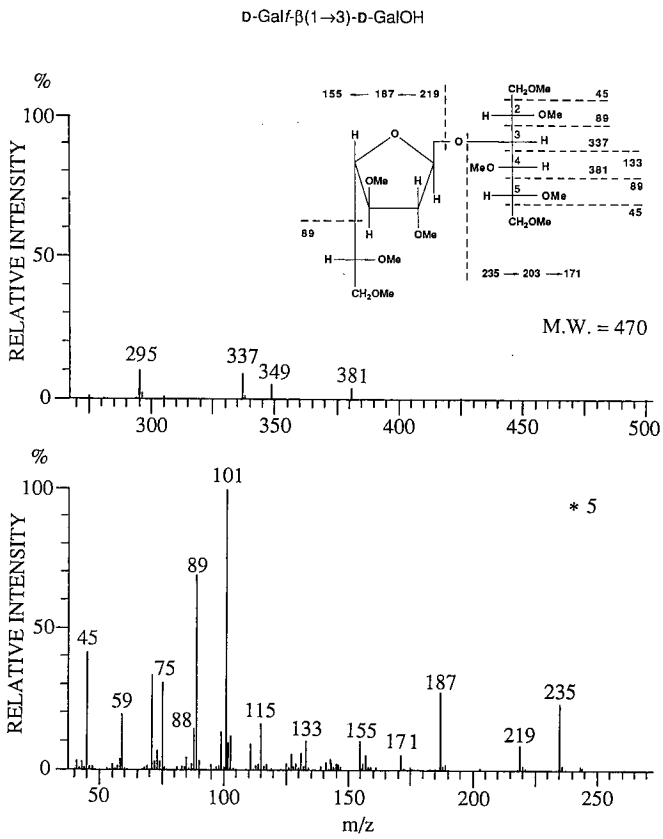


Fig. 10. 70-eV electron-impact mass spectrum of permethylated fraction IV.1 derived from *B. cellulosolvens* cellulase complex.

various synthetic methyl galactosides were prepared and examined for their capacity to inhibit the binding between GSI-B$_4$ and the cellulosome-like entity. Using a dot-blot assay, it was observed that methyl α-D-Gal*f* inhibited the interaction, in a manner similar to that of methyl α-D-Gal*p*. Under the same conditions, methyl β-D-Gal*f* did not inhibit the interaction.

DISCUSSION

In previous studies [5, 6, 20], we have shown that a remarkable similarity exists in the ultrastructural and biochemical properties of the cellulase complex of cellulolytic bacteria. A striking feature of these bacteria is the cross-reactivity of their distinctive protuberance-like surface structures with an α-Gal-binding lectin (GSI-B$_4$). In the cellulolytic anaerobic thermophile, *C. thermocellum*, it has been conclusively demonstrated that the primary lectin-binding glycoconjugate is a major component of the multifunctional cellulolytic complex of this bacterium, the cellulosome [3]. An extension of this work [4] revealed a series of partial structures which were proposed to represent intermediates in the biosynthetic pathway of the largest oligosaccharide structure of the cellulosome, namely, 3-*O*Me-Glc*p*NAc-α(1 → 2)-[Gal*p*-α(1 → 3)-]Gal*f*-α(1 → 2)-Gal.

We have recently isolated a cellulosome-like entity from the Gram-negative, cellulolytic mesophile, *B. cellulosolvens* [12]. Similar to *C. thermocellum*, a high-molecular-mass glycosylated component of this entity interacted strongly with GSI-B$_4$. Structural analysis of the carbohydrate chains as reported in the present study revealed a novel pentasaccharide alditol, namely, Gal*f*-α(1 → 3)-Glc*p*NAc-α(1 → 2)-Gal*f*-α(1 → 2)-[Gal*f*-β(1 → 3)-]GalOH, and a series of closely related partial

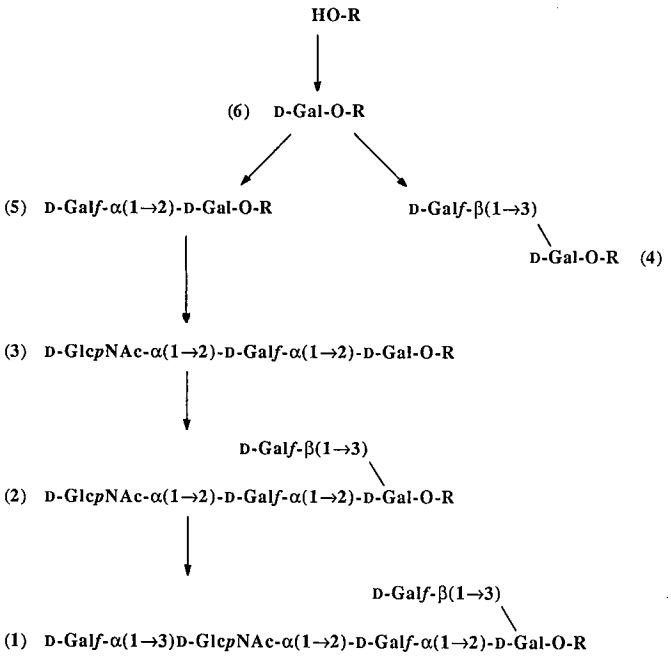


Fig. 11. Hypothetical biosynthetic pathway of the oligosaccharide chains occurring in the *B. cellulosolvens* cellulase complex. The numerals in brackets refer to the structures discussed in Results. R, protein core.

structures. By analogy to *C. thermocellum*, the observed heterogeneity possibly reflects different stages in the biosynthesis of the carbohydrate moiety in *B. cellulosolvens*. Based on the various structures found, it is tempting to propose a pathway for the biosynthesis of the pentasaccharide as depicted in Fig. 11.

It is interesting to note that the terminal α-Gal moiety was determined to have the furanose ring form. The strong interaction of GSI-B_4 with an α-D-Gal*f* derivative was observed here for the first time, representing a novel feature in terms of specificity of the lectin concerned. Based on the comparative structures, an explanation for this observation is not self-evident, and further investigations concerning the specificity of GSI-B_4 would be of interest.

Another very remarkable feature is the similarity of the structures of the oligosaccharides in the cellulase complex of both bacteria. The oligosaccharide alditols obtained from both cellulolytic bacteria are characterized by the presence of galactofuranose as the major component and *N*-acetylglucopyranosamine. Mainly α(1 → 2) and α(1 → 3) linkages are involved. The ring form and configuration of the peptide-bound galactose are still unknown.

Recent evidence obtained by us and by others has implicated cellulosome-like entities in various cellulolytic bacteria. On basis of their interaction with GSI-B_4, it is tempting to suggest that these entities express a common oligosaccharide sequence with species-dependent variations [6]. In addition, not all cellulosomal subunits are glycosylated to the same degree [3, 21] (this report). In cellulolytic bacteria, one or two specific components are heavily glycosylated (i.e. the 210-kDa S1 subunit in *C. thermocellum* and the 230-kDa subunit in *B. cellulosolvens*). Direct evidence for the role of the oligosaccharide moieties for the functioning of the cellulosome is still lacking. A role for the glycosylated component in cellulose binding was suggested earlier, and it is possible that the sugar moieties contribute to this property. Another function which has been proposed for the glycosylated subunits is that they could serve to bind the various cellulosome components into a complex. Alternatively, it has been suggested [22] that the glycosylated subunit(s) may comprise highly cohesive subcomponents which are not normally disassociated under standard SDS/PAGE conditions. The structural integrity of the complex could be dependent on the presence of oligosaccharides. In this respect, it is interesting to note that the glycosylated subunits of both *B. cellulosolvens* and *C. thermocellum* exhibit a pH-dependent and ionic-strength-dependent anomaly in their respective SDS/PAGE mobility patterns. Further research on the properties of the various cellulosomal components should clarify the involvement of the oligosaccharide substituents in the structure and/or function of the cellulosome complex.

We thank C. Versluis and A. C. van der Kerk-van Hoof (Bijvoet Center, Department of Mass Spectrometry, Utrecht University) for recording the mass spectra.

REFERENCES

1. Lamed, R. & Bayer, E. A. (1988) *Adv. Appl. Microbiol. 33*, 1–46.
2. Lamed, R. & Bayer, E. A. (1988) *Methods Enzymol. 160*, 472–482.
3. Gerwig, G. J., de Waard, P., Kamerling, J. P., Vliegenthart, J. F. G., Morgenstern, E., Lamed, R. & Bayer, E. A. (1989) *J. Biol. Chem. 264*, 1027–1035.
4. Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F. G., Morgenstern, E., Lamed, R. & Bayer, E. A. (1991) *Eur. J. Biochem. 196*, 115–122.
5. Lamed, R. & Bayer, E. A. (1988) in *Biochemistry and genetics of cellulose degradation* (Aubert, J. P., Beguin, P. & Millet, J., eds) pp. 101–116, Academic Press, Orlando, FL.
6. Lamed, R., Naimark, J., Morgenstern, E. & Bayer, E. A. (1987) *J. Bacteriol. 169*, 3792–3800.
7. Giuliano, C., Asther, M. & Khan, A. W. (1983) *Biotechnol. Lett. 5*, 395–398.
8. Murray, W. D., Sowden, L. C. & Colvin, J. R. (1984) *Int. J. Syst. Bacteriol. 34*, 185–187.
9. Giuliano, C. & Khan, A. W. (1984) *Appl. Environ. Microbiol. 48*, 446–448.
10. Murray, W. D., Sowden, L. C. & Colvin, J. R. (1986) *Lett. Appl. Microbiol. 3*, 69–72.
11. Murray, W. D. (1987) *Biotechnol. Bioeng. 29*, 1151–1154.
12. Lamed, R., Morag (Morgenstern), E., Mor-Yosef, O. & Bayer, E. A. (1991) *Curr. Microbiol. 22*, 27–33.
13. Kamerling, J. P. & Vliegenthart, J. F. G. (1989) in *Clinical biochemistry – principles, methods, applications* (Lawson, A. M., ed.) vol. 1, pp. 176–263, Walter de Gruyter, Berlin, New York.
14. Gerwig, G. J., Kamerling, J. P. & Vliegenthart, J. F. G. (1978) *Carbohydr. Res. 62*, 349–357.
15. Gerwig, G. J., Kamerlins, J. P. & Vliegenthart, J. F. G. (1979) *Carbohydr. Res. 77*, 1–7.
16. Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A. & Smith, F. (1956) *Anal. Chem. 28*, 350–356.
17. Kärkkäinen, J. (1970) *Carbohydr. Res. 14*, 27–33.
18. Wood, C., Kabat, E. A., Murphy, L. A. & Goldstein, I. J. (1979) *Arch. Biochem. Biophys. 198*, 1–11.
19. Liener, I. E., Sharon, N. & Goldstein, I. J., eds (1986) *The lectins – properties, functions and applications in biology and medicine*, Academic Press, London.
20. Lamed, R., Naimark, J., Morgenstern, E. & Bayer, E. A. (1987) *J. Microbiol. Methods 7*, 233–240.
21. Ljundahl, L. G., Coughlan, M. P., Mayer, F., Mori, Y., Hon-Nami, H. & Hon-Nami, K. (1988) *Methods Enzymol. 160*, 483–506.
22. Morag (Morgenstern), E., Bayer, E. A. & Lamed, R. (1991) *Appl. Biochem. Biotechnol.* in the press.
23. Gerwig, G. J., Kamerling, J. P. & Vliegenthart, J. F. G. (1984) *Carbohydr. Res. 129*, 149–157.