Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 338 (2003) 2653–2658
CARBOHYDRATE RESEARCH
www.elsevier.com/locate/carres

The structure of the glycopeptides from the fish pathogen *Flavobacterium columnare*

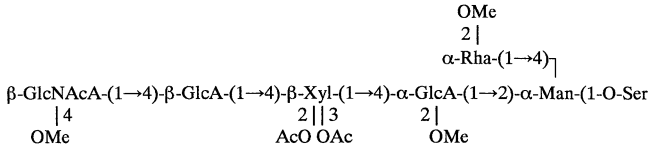
Evgeny Vinogradov,[a,*] Malcolm B. Perry,[a] William W. Kay[b]

[a] *Institute for Biological Sciences, National Research Council, 100 Sussex Dr., Ottawa, ON, Canada K1A 0R6*
[b] *Department of Bacteriology and Biochemistry, University of Victoria, Victoria, BC, Canada V8W 2T2*

Received 19 June 2003; accepted 19 August 2003

Abstract

Proteolytic digestion of the phenol–water extraction product of the fish pathogen *Flavobacterium columnare* afforded a mixture of glycopeptides in which the oligosaccharide moiety was an unusual hexasaccharide composed of 4-*O*-methyl-2-acetamido-2-deoxy-D-glucuronic acid (GlcNAcA), D-glucuronic acid (D-GlcA), 2,3-di-*O*-acetyl-D-xylose (D-Xyl), 2-*O*-methyl-D-glucuronic acid (D-GlcA), D-mannose (D-Man), and 2-*O*-methyl-L-rhamnose (L-Rha). By the application of high-resolution 1D and 2D NMR, mass spectrometry, and chemical analysis, the hexasaccharide structure was determined to be:



where all monosaccharides have the D-configuration except for 2-*O*-methyl-L-rhamnose; and were in the pyranose form. Only one carbohydrate structure was found. The peptide part was represented by tri- to hepta-peptides with a minimal common tripeptide fragment Asp-Ser-Ala, extended with Ala and Val.

Keywords: *Flavobacterium columnare*; Glycopeptide; NMR; MS

1. Introduction

Flavobacterium columnare (formerly *Flexibacter columnaris*) has been recognised as a worldwide pathogen of freshwater fish.[1,2] *F. columnare* is the aetiological agent of columnaris disease, which can be manifested as a chronic, ulcerative condition beginning as an infection of the body surface and gills. The disease often ends in death, leading to large economic losses in the fish breeding industry.

There is little information available concerning the factors determining virulence, although a clear correlation was made between virulence and the ability of the bacterium to adhere to the gill tissue. Immunization remains the most promising way to fight the disease, making analysis of the potential antigens important.

In the course of the analysis of *F. columnare* LPS and capsular polysaccharide we have found that phenol–water extraction of the cells with subsequent proteinase K treatment yields, together with LPS and capsular polysaccharide, a large quantity of glycopeptides. Here we present the results of the structural analysis of the latter compounds.

Abbreviations: LPS, lipopolysaccharide; GlcNAcA, 2-acetamido-2-deoxyglucuronic acid; FucNAm, 2-acetamido-2,6-dideoxygalactose; Q-TOF, hybrid quadrupole time-of-flight mass spectrometer.

* Corresponding author. Tel.: +1-613-990-0832; fax: +1-613-952-9092.
E-mail address: evguenii.vinogradov@nrc-cnrc.gc.ca (E. Vinogradov).

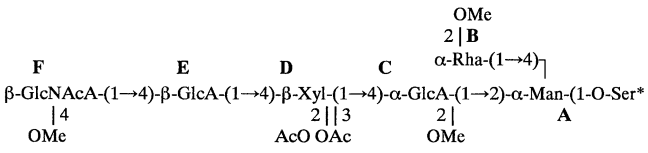
0008-6215/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.08.012

2. Results

Cells of *F. columnare* were extracted with a 45% phenol–water mixture according to the Westphal procedure.[3] The separated phenol layer was then dialysed, concentrated and treated with nucleases and proteinase K. LPS was precipitated by ultracentrifugation, and the supernatant separated by gel chromatography on a Sephadex G-50 column for the isolation of the capsular polysaccharide. Besides the polysaccharide fraction, eluted at the void volume, another large peak with lower molecular mass was collected. This product was separated by reverse phase chromatography on a C18 column to give a large number of fractions, which were analysed by NMR spectroscopy. The three most pure fractions were analysed in detail by this method (Fig. 1). All three had the same carbohydrate component. The analysis of product **1**, which contained the smallest peptide part, is described below.

A set of 2D NMR spectra of **1** (COSY, TOCSY, NOESY, HSQC, gHMBC) were recorded and completely interpreted (Table 1). The signals for six monosaccharides in the pyranose form were identified in the spectra of product **1**, as well as the signals of two Ala, Ser, and Asp residues, signals of three methyl groups, and signals of three acetyl groups. Monosaccharides were identified on the basis of vicinal proton coupling constants and ^{13}C NMR chemical shifts. Anomeric configurations were deduced from the $J_{1,2}$ coupling constants and chemical shifts of H-1, C-1, and C-5 signals. The connection between monosaccharides was determined on the basis of the interresidual NOE and HMBC correlations (A1Ser3, B1A4, C1A2, D1C4, E1D4, F1E4). The positions of *O*-methyl groups were also determined on the basis of NOE and HMBC data. NOE between Ser H-3 and Man A H-1 and corresponding HMBC correlation indicated that Man residue forms the linkage between the oligosaccharide and peptide parts via a serine hydroxyl group. In addition to the direct interresidual NOE correlations the following cross peaks, supporting the presented structure, were identified: C1A1, A1C5, B1-OMe at B2; C1-OMe at C2. The NOE A1C5 is only possible in the case of α-GlcA-(1-2)-α-Man (C-A) fragment where both monosaccharides have identical absolute configurations, as can be shown by molecular modeling.

The sequence of the amino acids was determined using HMBC correlations between H-2 of one amino acid and C-1 of the acylating amino acid. Thus correlations between Asp C-1 and Ser H-2, Ser C-1 and Ala[1] H-2, Ala[1] C-1 and Ala[2] H-2 were observed, leading to the sequence Asp-Ser-Ala-Ala.



1 Asp-Ser*-Ala[1]-Ala[2]

2 Val-Asp-Ser*-Ala[1]-Ala[2]

3 Val-Ala-Val-Asp-Ser*-Ala[1]-Ala[2]

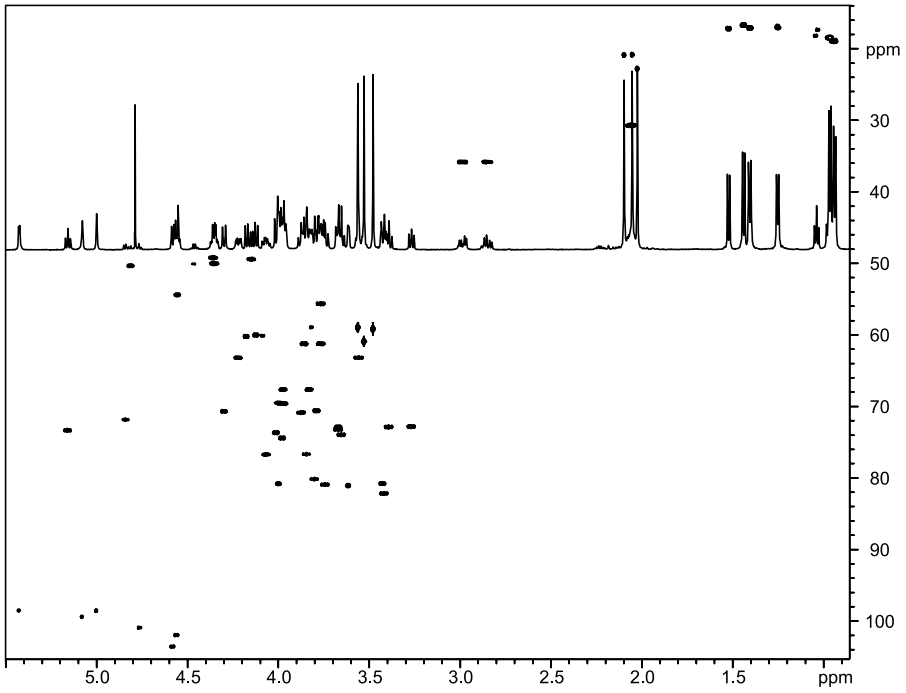


Fig. 1. $^{1}H-^{13}C$ HSQC correlation spectrum of compound **3**.

Table 1
NMR data for compound **1**

Unit, compound	Nucleus	1	2	3 (2b)	4 (3b)	5 (5a)	6 (6a, 5b)	6b (OMe)
Man A	^{1}H	5.11	4.01	3.98	3.84	3.70	3.78	3.87
	^{13}C	99.3	80.9	69.7	76.9	73.0	61.4	
Rha B	^{1}H	5.00	3.62	3.78	3.40	4.00	1.25	3.43
	^{13}C	98.6	81.1	70.6	72.9	69.7	17.1	59.2
GlcA C	^{1}H	5.42	3.43	3.87	3.79	4.28		3.56
	^{13}C	98.6	80.9	71.0	80.2	71.1	173.5	59.0
Xyl D	^{1}H	4.76	4.84	5.16	4.07	3.56	4.23	
	^{13}C	100.9	71.9	73.4	76.8	63.3		
GlcA E	^{1}H	4.58	3.27	3.65	3.74	3.97		
	^{13}C	103.6	72.9	74.1	81.1	74.7	171.7	
GlcNAcA F	^{1}H	4.56	3.77	3.67	3.41	3.99		3.53
	^{13}C	101.9	55.8	73.4	82.3	74.0	173.5	61.0
Asp	^{1}H		4.42	3.03	3.12			
	^{13}C	169.5	50.2	35.8	173.7			
Ser	^{1}H		4.64	3.86	4.00			
	^{13}C	170.9	54.5	67.8				
Ala[1]	^{1}H		4.36	1.41				
	^{13}C	174.8	50.2	17.3				
Ala[2]	^{1}H		4.36	1.45				
	^{13}C	177.2	49.4	16.9				
Asp in **3**	^{1}H		4.81	2.85	2.99			
	^{13}C	172.8*	50.3	35.9	174.4*			

Additional signals (δ^{1}H/δ^{13}C): OAc at D3: 2.06/21.0, 173.8; OAc at D2 2.10/21.0, 173.4; NAc at F2: 2.03/22.9, 175.4.

ESI Q-TOF mass spectrum of the product **3** (Fig. 2) contained the expected molecular ion at *m/z* 1768 together with a number of fragment ions, consistent with the proposed structure. Two series of ions were observed, one resulting from the loss of mono- or oligosaccharide fragments (B, F, FE, FED, and combinations of B with F, FE, FED) from the molecular ion, and another series of glycosyl-cation ions from the residues F, FE, or FED.

The amino acid sequence of longer peptides was not obvious from NMR data. For the sequence determination of the peptide part of the oligosaccharide **2** it was treated with anhydrous HF, peptide was isolated by reverse phase HPLC, and analysed by Q-TOF MS. Peaks of the b series (*m/z* 215.1, 302.1, 372.2) and the y series (*m/z* 363.1, 248.1, 161.1), together with an $[M+H]^{+}$ peak at *m/z* 462.2 were observed, leading to the sequence Val-Asp-Ser-Ala-Ala. The peptide moiety of compound **3** was analysed using a protein sequencer and was determined to be Val-Ala-Val-Asp-Ser-Ala-Ala.

Negative ion ESIMS analysis of the whole mixture of glycopeptides showed that all of them contained the same oligosaccharide part and that they differed only by the length of the peptide chain. The $[M^{-}]$ peaks (as well as the signals of the corresponding double charged ions) with the following amino acid composition were observed: *m/z* 1381.5 (Ala_2, Ser), 1496.7 (Ala_2, Ser, Asp), 1595.1 (Ala_2, Ser, Asp, Val), 1696.8 (Ala_2, Ser, Asp, Val_2), 1766.2 (Ala_3, Ser, Asp, Val_2), 1794.3 (Ala_2, Ser, Asp, Val_3).

Monosaccharide analysis (GC of alditol acetates) of the isolated products revealed the presence of xylose, mannose, glucose, 2-*O*-methylglucose, and 2-*O*-methylrhamnose. Glucose and 2-*O*-methylglucose were not present in the oligosaccharide structure however; their corresponding alditol acetates were derived as a result of the borohydride reduction of the lactones of the corresponding parent uronic acids. This was confirmed by the monosaccharide analysis with the reduction of the hydrolysate with deuterated borohydride, which led to the formation of 1,6,6-trideuteroglucitol (from D-GlcA) and 1,6,6-trideutero-2-*O*-methylglucitol. In another experiment, the sample was first subjected to methanolysis and $NaBH_4$ reduction to convert uronic acids to the corresponding hexoses. Subsequent monosaccharide analysis (GC of alditol acetates) revealed in addition, the acetate of 4-*O*-methyl-2-amino-2-deoxyglucitol, thus proving the presence of 4-*O*-methylglucosaminuronic acid in the glycopeptides.

For the determination of the absolute configuration of the monosaccharides, the sample was treated with 1 M HCl in (*S*)-2-butanol followed by acetylation. GC-MS analysis of the products with the corresponding reference standards, prepared from known monosaccharides and (*S*)- and (*R*)-2-BuOH determined that GlcA, Man, and Xyl have the D-configuration. For the determination of the absolute configuration of 2-*O*-

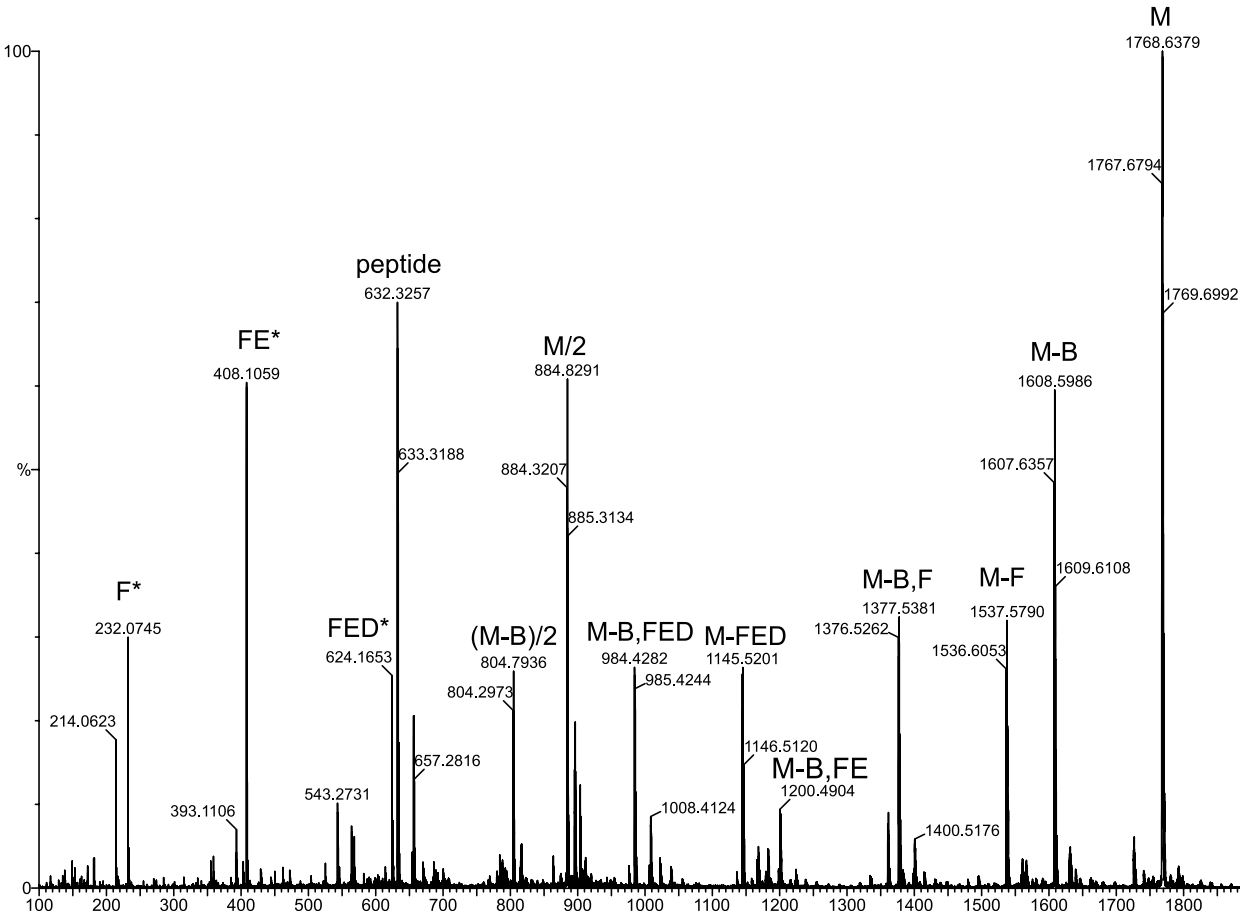


Fig. 2. Q-TOF mass spectrum of compound **3**. Peaks labelled with stars belong to glycosyl-cations; others originate from molecular ion with the loss of indicated residues.

methyl-Rha, the glycopeptide mixture was treated with MeOH–HCl, the products were separated by preparative paper chromatography, and the L-configuration of the methyl 2-*O*-methyl-α-rhamnopyranoside was established on the basis of optical rotation data.

For the determination of the absolute configuration of 4-*O*-Me-β-GlcNAcA (residue F), ^{13}C NMR chemical shifts (C-1 of the residue F and C-3, C-4, C-5 of the GlcA residue E) were analysed using database spectra.[4] *N*-Acetylglucosamine data were used for β-GlcNAcA residue, since the presence of a methyl group at O-4 and a carboxyl group at C-5 does not influence chemical shifts of interest. The results clearly show that 4-*O*-Me-β-GlcNAcA has the D-configuration.

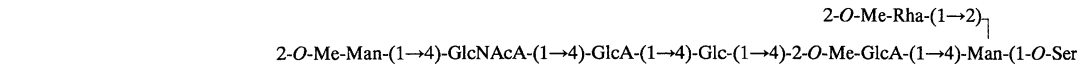
3. Discussion

A uniform peptide structure points to the origin of all products from a single protein, which remains to be identified. The search for Val-Ala-Val-Asp-Ser-Ala-Ala sequence in NCBI BLAST database gave exact matches for some *Mycobacterium* proteins but not for *Flavobacterium*. The antigenic properties of the characterized glycopeptides are under investigation.

A structure similar to that of *F. columnare* has been previously found for an oligosaccharide in *Flavobacterium meningosepticum* (Scheme 1).[5,6]

The *F. meningosepticum* oligosaccharide contained a GlcNAcA-(1→4)-GlcA fragment, found also in *F. columnare*, and was also linked to Ser residue through Man, substituted with 2-O-Me-GlcA and 2-O-Me-Rha. The positions of the attachment of GlcA and Rha to the Man residue, however, are presented in reversed order in the *F. meningosepticum* product. The suggested linkage positions of the GlcA and Rha to the Man residue in the latter oligosaccharide were based on MS data only, and the substitutions may be exchanged in the light of further study.

It is noteworthy that the capsular polysaccharide and the LPS O-polysaccharide antigen produced by *F. columnare* are linear unbranched polymers of a repeating trisaccharide unit having the structure in Scheme 2 (Sug = 2-acetamido-2,6-dideoxy-D-*xylo*-hexos-4-ulose),[7] apparently unrelated to the hexasaccharide structure of the investigated glycoprotein products.



Scheme 1.

The results of the current study show another example for the covalently linked carbohydrate modification of proteins in bacteria. Recent discoveries of the glycosylated bacterial proteins has dispelled the previously held view that prokaryotes were not able to synthesize glycoproteins.[8] In fact, the variety of different structures suggests that their variations exceed their analogous limited display in eukaryotes.[9]

The earliest examples of protein glycosylation in prokaryotes were discovered in *archaea*, which displayed glycosylated surface (S-layer) proteins.[10] Subsequently, non-S-layer eubacterial glycoproteins were found in pathogenic Gram-negative bacteria such as *Borrelia burgdorferi*,[11,12] *Campylobacter jejuni*,[13–15] *Chlamydia trachomatis*,[16] *Ehrlichia canis*,[17] *Neisseria meningitidis*,[18] *Neisseria gonorrhoea*,[19,20] *Pseudomonas aeruginosa*,[21,22] and *Escherichia coli*.[23,24] Although little is known about their biological functions, in several species gene clusters, suggested representing general protein glycosylation systems, have been identified. In some cases, genes encoding specific glycosyltransferases have been linked with virulence genes and give rise to the possible role of glycosylation in pathogenesis particular in the case of Gram-negative pathogens.

Most bacterial glycoproteins appear to be either associated with the surface of the organism as in pili, flagella or to be secreted into the environment and thus suggests that they may be exposed to host cells and play a role in adhesion, signalling events, or evasion of the immune system. The role of the analysed glycopeptides in fish disease still has to be clarified.

4. Experimental

4.1. Bacterial culture and glycopeptide isolation

F. columnare (ATCC 43622, NRCC 6160) was grown at 16 °C in a 52 L fermentor in a medium having the composition: Tryptone, 4 g; yeast extract, 0.4 g; $MgSO_4$, 0.5 g; $CaCl_2$, 0.5 g; NaOAc, 0.2 g and maltose, 10 g/L, adjusted to pH 7.0 with 0.1 M NaOH. A 2.5 L inoculum grown at 22 °C was used and stirring was maintained at 200 rpm. and dissolved oxygen at 20%. Cells were killed with 1% phenol (final concentration, 2 h at 4 °C) in late log phase 25 h growth (A_{600} 3.34) and, following acidification with AcOH to pH 4.0 at 0 °C to break the gel-like constitution, the suspended cells were harvested by centrifugation (yield ~300 g wet mass).

The *F. columnare* cells (300 g, wet mass) were extracted for 15 min at 65 °C with vigorously stirred 50% (w/v) aq phenol (1.2 L) and after cooling to 4 °C the separated phenol layer was collected and dialyzed against running water until free from phenol. The lyophilised dialyzed retentate dissolved in 0.02 M NaOAc (80 mL) was treated sequentially with RNase, DNase and proteinase K (37 °C, 2 h each). The digest was cleared by low speed centrifugation (3000 × g) and then subjected to ultracentrifugation (105000 × g, 4 °C, 12 h) and the centrifugate was collected from the precipitated lipopolysaccharide gel (1.7 g). The centrifugate on treatment with 6 vols of acetone gave a precipitate, which on Sephadex G-50 column chromatography afforded a high molecular mass fraction (K_{av} 0.02) of a capsular polysaccharide and a low molecular weight fraction (K_{av} 0.07, 180 mg) composed of oligosaccharide-linked peptides, the subject of this report.

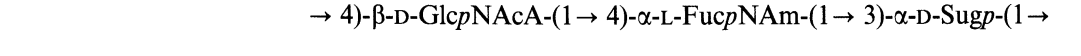
Glycopeptide mixture was separated on an Aqua C18 column (250 × 10 mm, Phenomenex) in 0.1% CF_3COOH (solvent A)-90% MeCN gradient (solvent B) from 5% B to 80% B, with the UV detection at 220 nm.

4.2. NMR spectroscopy and general methods

NMR spectra were recorded for solutions at 25 °C in D_2O on a Varian UNITY INOVA 600 instrument at using acetone as reference for (^{1}H 2.225 ppm, ^{13}C 34.5 ppm). Varian standard programs COSY, NOESY (mixing time of 200 ms), TOCSY (spin lock time 120 ms), HSQC, and gHMBC were used. Mass spectra were obtained with a Q-TOF2 hybrid quadrupole time-of-flight mass spectrometer (Micromass).

For monosaccharide analysis, a sample (1 mg) was hydrolysed with 4 M CF_3CO_2H (100 °C, 2 h), dried under a stream of nitrogen and conventionally reduced with $NaBH_4$. After adding conc AcOH and MeOH (2 × 1 mL), the sample was dried, acetylated with Ac_2O (0.5 mL, 100 °C, 20 min), dried and analysed by GCMS.

Another sample (1 mg) was treated with 1 M HCl in MeOH (85 °C, 3 h), dried, reduced with $NaBH_4$ in water (20 °C, 2 h), excess of $NaBH_4$ was destroyed with AcOH, boric acid removed by drying with MeOH twice,



Scheme 2.

the remaining material was hydrolysed and converted into alditol acetates as described above.

For the determination of the absolute configuration of the monosaccharides compound **2** (1 mg) was treated with 10:1 (*S*)-2-butanol–CH_3COCl (0.25 mL, 2 h, 85 °C), dried under a stream of air, acetylated and analysed by GC in comparison with authentic standards, prepared from respective commercial monosaccharides with (*S*)- and (*R*)-2-butanol.[25]

For the determination of the 2-*O*-methyl-rhamnose configuration the glycopeptide mixture (40 mg) was treated with 10:1 MeOH–CH_3COCl (4 mL, 2 h, 85 °C), evaporated, and the products separated by ascending paper chromatography on Whatman 1 paper in 3:10:3 pyridine–butanol–water. Paper was cut in strips, substances were eluted with water, dried and analysed by NMR spectroscopy. A fraction containing ~80% methyl 2-*O*-methyl-α-rhamnopyranoside had $[\alpha]_D$ of −18° (*c* 0.4, $CHCl_3$) [lit. −37° (*c* 1.3, $CHCl_3$)].[26]

Acknowledgements

This work was supported by the Canadian Bacterial Diseases Network. Authors thank Dr. J. Kelly, L. Nossova and D. Watson (NRC Canada) for the MS analysis, monosaccharide analysis and peptide sequencing.

References

1. Decostere, A. *Verh. K. Acad. Geneeskd. Belg.* **2002**, *64*, 421–430.
2. Decostere, A.; Haesebrouck, F.; Charlier, G.; Ducatelle, R. *Vet. Microbiol.* **1999**, *67*, 287–298.
3. Westphal, O.; Jann, K. *Methods Carbohydr. Chem.* **1965**, *5*, 83–91.
4. Lipkind, G. M.; Shashkov, A. S.; Knirel, Y. A.; Vinogradov, E. V.; Kochetkov, N. K. *Carbohydr. Res.* **1988**, *175*, 59–75.
5. Plummer, T. H., Jr.; Tarentino, A. L.; Hauer, C. R. *J. Biol. Chem.* **1995**, *270*, 13192–13196.
6. Reinhold, B. B.; Hauer, C. R.; Plummer, T. H.; Reinhold, V. N. *J. Biol. Chem.* **1995**, *270*, 13197–13203.
7. MacLean, L. L.; Perry, M. B.; Crump, E. M.; Kay, W. W. *Eur. J. Biochem.* **2003**, *270*, 3440–3446.
8. Benz, I.; Schmidt, M. A. *Mol. Microbiol.* **2002**, *45*, 267–276.
9. Schaffer, C.; Graninger, M.; Messner, P. *Proteomics* **2001**, *1*, 248–261.
10. Mescher, M. F.; Strominger, J. L. *FEBS Lett.* **1978**, *89*, 37–41.
11. Sambri, V.; Stefanelli, C.; Cevenini, R. *Arch. Microbiol.* **1992**, *157*, 205–208.
12. Sambri, V.; Massaria, F.; Ardizzoni, M.; Stefanelli, C.; Cevenini, R. *Zentralbl. Bakteriol.* **1993**, *279*, 330–335.
13. Young, N. M.; Brisson, J. R.; Kelly, J.; Watson, D. C.; Tessier, L.; Lanthier, P. H.; Jarrell, H. C.; Cadotte, N.; St Michael, F.; Aberg, E.; Szymanski, C. M. *J. Biol. Chem.* **2002**, *277*, 42530–42539.
14. Guerry, P.; Doig, P.; Alm, R. A.; Burr, D. H.; Kinsella, N.; Trust, T. J. *Mol. Microbiol.* **1996**, *19*, 369–378.
15. Thibault, P.; Logan, S. M.; Kelly, J. F.; Brisson, J. R.; Ewing, C. P.; Trust, T. J.; Guerry, P. *J. Biol. Chem.* **2001**, *276*, 34862–34870.
16. Swanson, A. F.; Kuo, C. C. *Infect. Immun.* **1991**, *59*, 2120–2125.
17. McBride, J. W.; Yu, X. J.; Walker, D. H. *Infect. Immun.* **2000**, *68*, 13–18.
18. Stimson, E.; Virji, M.; Makepeace, K.; Dell, A.; Morris, H. R.; Payne, G.; Saunders, J. R.; Jennings, M. P.; Barker, S.; Panico, M. *Mol. Microbiol.* **1995**, *17*, 1201–1214.
19. Parge, H. E.; Forest, K. T.; Hickey, M. J.; Christensen, D. A.; Getzoff, E. D.; Tainer, J. A. *Nature* **1995**, *378*, 32–38.
20. Marceau, M.; Forest, K.; Beretti, J. L.; Tainer, J.; Nassif, X. *Mol. Microbiol.* **1998**, *27*, 705–715.
21. Castric, P.; Cassels, F. J.; Carlson, R. W. *J. Biol. Chem.* **2001**, *276*, 26479–26485.
22. Arora, S. K.; Bangera, M.; Lory, S.; Ramphal, R. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9342–9347.
23. Lindenthal, C.; Elsinghorst, E. A. *Infect. Immun.* **2001**, *69*, 52–57.
24. Benz, I.; Schmidt, M. A. *Mol. Microbiol.* **2001**, *40*, 1403–1413.
25. Gerwig, G. J.; Kamerling, J. P.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1978**, *62*, 349–357.
26. Pozsgay, V.; Nanasi, P. *Carbohydr. Res.* **1980**, *81*, 184–186.