The structure of the Morganella morganii lipopolysaccharide core region and identification of its genomic loci

Vinogradov, Evgeny; Nash, John H. E.; Foote, Simon; Young, N. Martin

Publisher's version / Version de l'éditeur:
https://doi.org/10.1016/j.carres.2014.10.032
Carbohydrate Research, 402, pp. 232-235, 2014-11-13

The structure of the *Morganella morganii* lipopolysaccharide core region and identification of its genomic loci

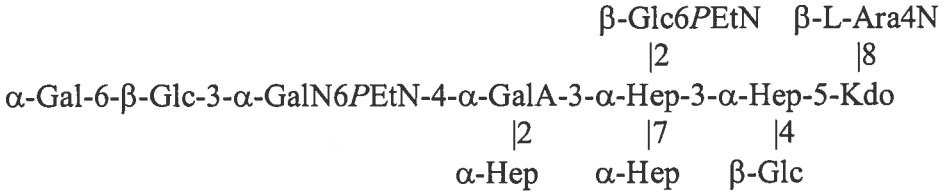
Evgeny Vinogradov[a,*], John H.E. Nash[b], Simon Foote[a] and N. Martin Young[a]

[a] *Human Health Therapeutics Portfolio, National Research Council of Canada, 100 Sussex Drive, Ottawa, ON, Canada K1A 0R6*
[b] *Public Health Agency of Canada, Toronto ON, Canada*

*Corresponding author. Tel.: 613 990-0832
E-mail address:Evguenii.Vinogradov@nrc-cnrc.gc.ca

ABSTRACT

The core region of the lipopolysaccharide of *Morganella morganii* serotype O:1ab was obtained by hydrolysis of the LPS and its structure was determined to be:



This structure is highly homologous to those from the two major members of the same Proteeaea tribe, *Proteus mirabilis* and *Providencia alcalifaciens*, and analysis of the *M. morganii* genome disclosed that the loci for its outer core, lipid A and 4-aminoAra moieties are similarly conserved. While a locus for the O-chain biosynthetic genes could be identified on the basis of two conserved flanking genes, the individual genes required for its synthesis could not be assigned.

Keywords: Morganella morganii O1, LPS, core, carbohydrate gene loci

1. Introduction

Morganella morganii is a human commensal organism which can be an opportunistic pathogen, causing urinary tract and post-surgical infections[1-4]. A facultative anaerobic Gram-negative organism, it is a member of the Proteeae tribe, and thus related to pathogens such as *Proteus mirabilis* and *Providencia alcalifaciens*[5]. The structure of the O-antigen of *M. morganii* serotype O:1ab has been determined[6] and it is unusual in having phosphocholine sidechains. It also has a second phosphate group and an amino group, and as a zwitterionic polysaccharide it is able to initiate T-cell immunological responses by binding to MHCII molecules[6]. The structure of the O-chain of a second strain of unknown serotype was found to be comprised of two unusual higher sugars[7].

In order to compare its lipopolysaccharide (LPS) to those of other Proteae, we have determined the structure of the core region of its LPS and acquired a preliminary genome sequence for this organism. This sequence allowed identification of several gene clusters responsible for the biosynthesis of the LPS.

2. Results and Discussion

2.1 Structure of the LPS core

To obtain core derivatives suitable for structural analysis LPS was treated with diluted acetic acid, which gave the core fraction after size-exclusion separation. NMR spectra showed that this material was heterogeneous, mostly due to the presence of different Kdo degradation products. The anomeric region of the ^{1}H NMR spectrum contained many signals of different intensity (data not shown). The core was purified to some degree by conventional anion-exchange separation and by high-pH anion-exchange chromatography (HPAEC) on a Dionex column, but none of the fractions obtained after these separations showed clean NMR spectra; in all cases the signals of most of the sugar residues were present in several variants. However, all constituent sugar residues and linkages between them were identified, and the results were in agreement with mass-spectrometry and methylation data expected from the structure 1 (Fig. 1).

Monosaccharide analysis of the core showed the presence of Glc, Gal, and LD-Hep. Methylation analysis (alditol acetates) showed t-Glc (residue K), t-Gal (Y), 6-substituted Glc (L), terminal (G,T), 3,4- (E) and 2,3,7- (F) substituted LD-Hep, all sugars being in pyranose form.

Interpretation of the 2D NMR spectra of compound 1 (Table 1) showed the presence of four heptoses, Kdo in various forms, three β-Glc, α-Gal, α-GalN, β-Ara4N, and α-GalA, identified by their characteristic signal patterns and chemical shifts. Kdo was present mostly in unidentified forms with H-3 signals between 2.5-2.7 ppm. The sequence of the monosaccharides was based on the following NOE correlations: Y1:L6; L1:M3,4; M1:H4; T1:H1,2; F1:E3; E1:C5,7; J1:F1,2; G1:F7; K1:E4,6; Z1:C8. Some of these correlations were confirmed by HMBC, although HMBC spectra of good quality were not possible to obtain and some

correlations were not visible. The substitution positions agree with downfield shift of the ^{13}C NMR signals of substituted carbons (Table 1).

The core was phosphorylated with *P*EtN at two positions, M6 and J6. The positions of phosphorylation were identified from ^{1}H-^{31}P HMQC and HMQC-TOCSY spectra.

NMR data indicated the presence of GalN (residue M) with a free amino-group, which was used for the deamination, and thus compound 2 was obtained. The anh-Tal derivative, formed after deamination of the GalN, lost most of the phosphate from O-6 after deamination. The rest of the core between Kdo and GalN M was also isolated from deamination products, but its spectra were even worse than the spectra of the whole core and the results of its analysis will not be discussed here.

Complete deacylation of the LPS produced very complex mixture of deacylated compounds, from which an oligosaccharide 3 was isolated in small amount by HPAEC. The 4-substituted galacturonic acid (residue H) was converted into a 4,5-ene derivative (α-ΔGalA) due to β-elimination of the O-4 substituent in the basic conditions of the deacylation. The NMR of compound 3 fully agreed with the structure deduced from the study of compound 1 (data not shown). Additionally, a second Kdo (residue D) was found in an usual arrangement, and a GlcN disaccharide originating from lipid A with only one phosphate at A1 was present.

ESI mass spectra of the core contained several peaks with the mass close to that expected for compound 1 (calculated exact mass 2350.6994, average 2351.9142), among them expected mass of 2351.8. Two other similarly abundant species had masses of 2324.0 and 2341.6, 28 and 10 amu lower than expected and probably corresponding to some degradation products of Kdo. Core fractions obtained by HPAEC separation contained these three species in various ratios. Also, fractions eluting close to the beginning of the HPAEC chromatogram were missing *P*EtN,

with the masses of 2200, 2218.8, and 2228.8 (expected as 2351.9-123). The mass spectrum of compound 3 had one peak at 2365.4 amu, in agreement with the proposed structure.

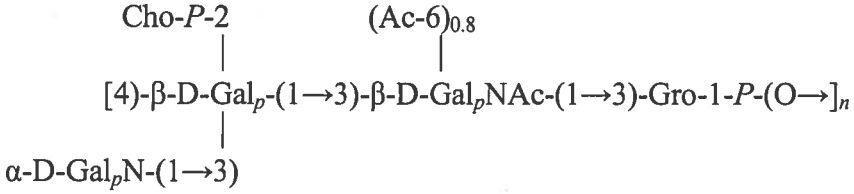
The structure of *M. morganii* core presented here has features in common with the structures of *Proteus*, *Serratia* and *Klebsiella* cores, which all contain α-HexN-(1-4)-α-GalA disaccharide linked to Hep_3-Kdo_2 inner part, always with a free aminogroup on the hexosamine (GlcN or GalN)[8-10]. GalA is often substituted at O-2 with a heptose or heptose disaccharide in these core oligosaccharides, as it was also observed in the *M. morganii* core. This reflects the taxonomic similarity of these microorganisms. Structural analysis of the *M. morganii* core was quite difficult because of the extremely messy NMR spectra of the acid-released core and unexplained degradation of the material during alkaline deacylation. Analysis of all fractions did not reveal the reason for the formation of similar but somehow different products.

2.2 Genome sequences of LPS-related loci

Sequencing of the genome of the *M. morganii* O1 type strain, ATCC 49993, disclosed several loci related to the biosynthesis of LPS. The main genes for the core (Table 1) occurred in a 13-gene cluster, which included several genes for heptose biosynthesis and transfer. It had more genes homologous to *Pt. mirabilis* than to *Pv. alcalifaciens,* and all had very high homology scores. The *Pt. mirabilis* genes were in identical order, but one of its genes did not have a homolog in *M. morganii*. Two other gene clusters (Table 2) included genes responsible for aspects of LPS export and Lipid A biosynthesis, all of which again showed high homology to those of the other Proteae bacteria. Candidates for a number of other genes involved in these aspects of LPS biosynthesis were found scattered throughout the genome, for example MmoY-00758 was annotated as the lipid A export ATP-binding/permease protein, *msbA* and MmoY_02113, as the phosphoheptose isomerase, *gmhA*. A locus for biosynthesis and transfer to

LPS of the antibiotic resistance factor L-Ara4N[11] was also found (Table 3) and it closely matched the loci in *Pt. mirabilis* and *Pv. alcalifaciens.*

In *Providencia* species, the O-chain locus can be defined by the two flanking genes *cpxA*, a sensor protein and *yibK*, a tRNA methyltransferase[12]. Genes that were highly homologous to the *cpxA* and *yibK* genes of *Pt. mirabilis* and *Pv. alcalifaciens* were found in the *M. morganii* genome (Table 4) and there were 14 intervening genes, which appeared related to carbohydrate synthesis. The O-chain structure[6] is



Disappointingly, the automated annotation of this locus produced only a few candidate genes for the enzymes needed for O-chain biosynthesis and assembly, which were chiefly related to synthesis of the UDP precursors of its glycerol-phosphate and phosphocholine components. Notably, no candidate gene for a phosphocholine transferase *licD* was found within this locus or elsewhere in the genome. As expected, there were no genes homologous to O-chain genes of *Pt. mirabilis* or *Pv. alcalifaciens* either. However, the genes between MmoY_02778 and MmoY-02785 were all homologous to genes in the bacterium *Arsenophonus nasoniea*[13] between CBA75649 and CBA75635, but no structural information is available for the O-chain from this bacterium.

One other carbohydrate gene locus was identified in the *M. morganii* genome, which is responsible for the biosynthesis of the enterobacterial common antigen (ECA). It was comprised of 13 genes between MmoY_00248 and MmoY_00260, and this set had very high homologies to the ECA loci in both *Pt. mirabilis* and *Pv. alcalifaciens* (data not shown).

3. Experimental

3.1. LPS isolation

Cells were precipitated from growth medium by low speed centrifugation, washed with water twice. Cells (200 g wet weight) were stirred with 400 mL of 45% phenol for 20 min at 70 °C, cooled to room temperature and dialyzed against running water for 3 days. Content of dialysis bag was mixed with 10 % (v/v) of acetic acid, stirred for 20 min and precipitate removed by centrifugation. Solution was dialyzed for another 3 days, concentrated to 200 mL and subjected to ultracentrifugation at 120 000g for 3 hours. Precipitate was dissolved in water and dried to give LPS.

3.2. NMR spectroscopy

NMR experiments were carried out on a Varian INOVA 500 MHz (^{1}H) spectrometer with 3 mm gradient probe at 25-50 °C with acetone internal reference (2.225 ppm for ^{1}H and 31.45 ppm for ^{13}C), using standard pulse sequences gCOSY (gradient COrrelation SpectroscopY), TOCSY (Total Correlation Spectroscopy) (mixing time 120 ms), ROESY (Rotating frame Nuclear Overhauser Effect Spectroscopy) (mixing time 300 ms), gHSQC (gradient Heteronuclear Single Quantum Coherence), and gHMBC (gradient Heteronuclear Multiple Bond Coherence) (100 ms long range transfer delay), HMQC for ^{1}H-^{31}P correlation, J_{HX} set to 10 Hz. AQ time was kept at 0.8-1 sec for H-H correlations and 0.25 sec for HSQC. 256 increments were acquired for t1 in all 2D spectra, except 512 for gCOSY.

3.3. Chromatography

Gel chromatography was performed on a Sephadex G-15 column (1.5x60 cm) or a Bio-gel P6 column (2.5x60 cm) in pyridine-acetic acid buffer (4 mL:10 mL:1 L water), and

monitored by refractive index detector (Gilson). Anion exchange chromatography was done on an Hitrap Q column (2x5 mL size, Amersham), with UV monitoring at 220 nm in a linear gradient of NaCl (0-1 M, 1 h) at the 3 mL/min. Fractions of 1 min were collected and additionally tested for carbohydrates, by spotting on an SiO_2 TLC plate, dipping them in 5% H_2SO_4 in EtOH and heating with a heat-gun. All fractions of interest were dried in a Savant drying centrifuge and ^1H spectra were recorded for each fraction without desalting. For 2D NMR, desalting was performed on a Sephadex G15 column.

3.4. Monosaccharide analysis

Samples with added inositol standard were hydrolyzed with 3 M TFA at 120 °C. Monosaccharides were converted to alditol acetates by conventional methods and identified by GC-MS on a Varian Saturn 2000 instrument on a DB17 capillary column (30 m x 0.25 mm ID x 0.25 μm film) with helium carrier gas, using a temperature gradient 170°C (3 min), 250 °C at 5 °C·min^{-1}.

3.5. Determination of absolute configurations of monosaccharides

To the polysaccharide sample (0.2 mg) (*R*)-2-BuOH (0.2 mL) and acetyl chloride (0.02 mL) were added at room temperature, heated at 90 °C for 2 h, dried by air stream, acetylated, analyzed by GC-MS as described above. Standards were prepared from monosaccharides of known configuration with (*R*)- and (*S*)-2-BuOH.

3.6. Methylation analysis

For the methylation analysis core sample (2 mg) was dephosphorylated with 50 □L of 48% HF for 20 h at +10 °C, diluted with 2 mL of ethanol, precipitate collected by centrifugation, washed wth 2 mL of ethanol, dried.

Methylation was performed by Ciucanu-Kerek procedure[14]. 0.5 mg of the sample was dissolved in 0.5 mL of dry DMSO with heating at 100 °C for 5-10 min until complete dissolution, powdered NaOH (about 50 mg) was added and mixture stirred for 30 min, then 0.2 mL of MeI was added, mixture stirred for 30 min, flushed with air to remove MeI, diluted to 10 mL with water, passed through C18 Seppak, cartidge washed with 10 mL of water and methylated compound eluted with 5 mL of methanol. The product was hydrolyzed with 3 M TFA (120 °C, 3h), dried, reduced with $NaBD_4$, reagent destroyed with 0.5 mL of 4 M HCl, solution dried under the stream of air, dried twice with addition of MeOH (1 mL), acetylated with 0.4 mL Ac_2O - 0.4 mL pyridine for 30 min at 100 °C, dried, analyzed by GC-MS.

3.7. Genome sequencing

Genomic DNA was obtained from

Acknowledgements

We thank David Watson for preparing the genomic DNA.

References

1. Nicholle, L. E. *J. Infect. Dis.* **2001**, *183 (Suppl.1)* S5-8.
2. Falagas, M. E.; Kavvadia, P. K.; Mantadakis, E.; Kofteridis, D. P.; Bliziotis, I. A.; Saloustros, E.; Maraki, S.; Samonis G. *Infection* **2006**, *34*, 315-321.
3. Kim, B. N.; Kim, N. J.; Kim, Y. S.; Woo, J. H.; Ryu, J. *Scand. J. Infect. Dis.* **2003**, *35*, 98-103.
4. Tsanaktsidis, G.; Agarwal, A. S.; Maloof, A. J.; Chandra, J.; Mitchell, P. *J. Cataract Refract. Surg,*, **2003**, *29*, 1011-1013.
5. O'Hara, C. M.; Brenner, F.; Miller, J. M. *Clin. Microb. Rev.* **2000**, *13*, 534-546.
6. Young, N. M.; Kreisman, L. S. C.; Stupak, J.; MacLean, L. L.; Cobb, B. A.; Richards, J. C. *Glycobiology* **2011**, *21*, 1266-1276.
7. Kilcoyne, M., Shashkov, A. S., Senchenkova, S. A., Knirel, Y. A., Vinogradov, E. V., Radziejewska-Lebrecht, J., Galimska-Stypa, R. & Savage, A. V. *Carbohydr. Res.* **2002**, *337,* 1697-1702.
8. Vinogradov, E., Sidorczyk, Z. & Knirel, Y. *Aust. J. Chem.* **2002,** *55*, 61-67.
9. Vinogradov, E. & Perry, M. B. *Carbohydr. Res.* **2001**, *335*, 291-296.
10. Vinogradov, E., Petersen, B. O., Duus, J. O. & Radziejewska-Lebrecht, J. *Carbohydr. Res.* **2003**, *338*, 2757-2761.
11. Raetz, C. R.; Reynolds, C. M.; Trent, M. S.; Bishop, R. E. *Annu. Rev. Biochem.* **2007**, *76*, 295-329.
12. Ovchinnikova, O. G.; Liu, B.; Guo, D.; Kocharova, N. A.; Shashkov, A. S.; Chen, M.; Feng, L.; Rozalski, A.; Knirel, Y. A.; Wang, L. *Microbiology* **2012**, *158*, 1024-1036.
13. Wilkes, T. E.; Darby, A. C.; Choi, J. –H.; Colbourne, J. K.; Werren, J. H.; Hurst, G. D. D. *Insect Mol. Biol.* **2010**, *19 (Suppl. 1),* 59-73.
14. Ciucanu, I. *Anal. Chim. Acta*, **2006**, *576*, 147-155.

Figure legend

Figure 1. Proposed structure of isolated compounds. * at anh-Tal means partial phosphate at O-6.

Table 1. ^{1}H and ^{13}C chemical shifts (δ, ppm) for compound **1** at 25 °C; ^{31}P on J6 at 1.8 ppm, and on M6 at 0.9 ppm.

Sugar residue	H-1	H-2	H-3	H-4	H-5	H-6 (6a,b)	H-7a,b
E	5.13	3.99	4.08	4.34	3.79	4.07	3.75; 3.75
		71.1	79.2	72.1	72.3	69.2	63.7
F	5.37	4.10	4.14	4.04	3.60	4.24	3.66; 3.76
		79.2	75.6	67.8	73.7	67.6	69.9
G	4.94	3.99	3.87	3.87	3.63	4.07	
	102.2	71.0	71.6	67.0	72.2	69.8	
T	5.13	3.98	3.87	3.87	3.87	4.04	3.73
	97.9	71.0	71.6	67.0	72.7	69.8	63.8
H	5.64	4.03	4.19	4.50	4.53		
	98.1	71.9	67.8	79.6	71.8		
J	4.47	3.23	3.50	3.60	3.50	4.19; 4.26	
	104.4	73.8	75.8	69.3	75.8	64.8	
K	4.54	3.37	3.53	3.37	3.44	3.65; 3.89	
	103.3	74.7	76.7	71.4	77.1	62.1	
L	4.74	3.41	3.56	3.56	3.70	3.78; 4.00	
	104.6	74.0	76.5	70.2	75.3	66.7	
M	5.28	3.73	4.22	4.44	4.70	3.96; 4.00	
	96.8	51.6	78.5	67.3	70.0	64.3	
Y	5.01	3.85	3.88	4.01	3.96	3.75; 3.75	
	99.2	69.3	70.4	70.3	71.9	62.2	
Z	5.02	3.76	4.18	3.72	3.80; 4.13		
	99.5	68.9	66.5	52.9	59.1		
EtN on J	3.31	4.19					
	41.1	63.1					
EtN on M	3.31	4.11					
	41.1	63.0					

Table 2

^{1}H chemical shifts (δ, ppm) for the oligosaccharide **2** without phosphate.

Sugar residue	H-1	H-2	H-3	H-4	H-5	H-6 (6a, 6b
Y	5.00	3.84	3.87	3.99	3.94	3.75; 3.75
	99.4	69.5	70.7	70.4	72.2	62.3
L	4.62	3.39	3.51	3.51	3.67	3.78; 3.96
	104.0	74.0	76.8	70.8	75.5	67.1
M	5.08	3.93	4.41	4.39	4.13	3.78; 3.83
	90.9	82.8	82.9	73.0	82.3	61.2

Table 3

Comparison of LPS gene loci [a]

M. morganii gene #	Gene	Annotation	aa	*Pv. alcalifaciens* gene #	*E* value	*Pt. mirabilis* gene #	*E* value
Outer core loci							
Mm0Y_02796	*hldD*	ADP-L-glycero-D-manno-heptose-6-epimerase	313	ZP_03320691	3e-153	ZP_03839186	7e-161
Mm0Y_02797	*waaF*	ADP-heptose-LPS heptosyltransferase 2 (*rfaF*)	354	ZP_03320692	1e-151	ZP_03839187	3e-158
Mm0Y_02798	*waaC*	Lipopolysaccharide heptosyltransferase 1 (*rfaC*)	326	ZP_03320693	9e-128	ZP_03839188	8e-130
Mm0Y_02799	*wabN*	GlcNAc deacetylase (*walW*)	326	no hit		ZP_03839190	2e-142
Mm0Y_02800	*wabH*	GlcNAc transferase	374	no hit		ZP_03839193	2e-138
Mm0Y_02801	*wabG*	GalA transferase (*rfaG*)	376	ZP_03320705	1e-22	ZP_03839194	7e-161
Mm0Y_02802	*waaQ*	Lipopolysaccharide core heptosyltransferase 3 (*rfaQ*)	358	ZP_03320706	3e-78	ZP_03839195	9e-123
Mm0Y_02803	*waaA*	3-deoxy-D-manno-octulosonic-acid transferase	426	ZP_03320695	0	ZP_03839196	0
Mm0Y_02804	*waaE*	Glucosyltransferase	260	no hit		ZP_03839197	2e-105
Mm0Y_02805	*coaD*	Phosphopantetheine adenylyltransferase	162	ZP_03320696	3e-66	ZP_03839198	4e-76
Mm0Y_02806	*mutM*	Formamidopyrimidine-DNA glycosylase (*gpg*)	270	ZP_03320697	4e-119	ZP_03839200	4e-64
Mm0Y_02807	*waaL*	O-antigen ligase (*rfaL*)	417	ZP_03320698	4e-21	no hit	
Mm0Y_02808		Putative glycosyl transferase	245	ZP_03320699	3e-12	no hit	
Lipid A loci							
MmoY_02446	*yrbG*	Inner membrane protein	319	no hit		ZP_03842622	1e-75
MmoY_02447	*kdsD*	Arabinose 5-phosphate isomerase	323	ZP_03319213	6e-157	ZP_03842623	2e-153
MmoY_02448	*kdsC*	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase	188	ZP_03319212	4e-86	ZP_03842624	2e-81
MmoY_02449	*lptC*	Lipopolysaccharide export system protein	194	ZP_03319211	6e-53	ZP_03842625	6e-74
MmoY_02450	*lptA*	Lipopolysaccharide export system protein precursor	179	ZP_03319210	7e-76	ZP_03842626	5e-77
MmoY_02451	*lptB*	Lipopolysaccharide export system ATP-binding protein	242	ZP_03319209	7e-123	ZP_03842627	1e-120

MmoY_01318	*lpxD*	UDP-3-O-(3-hydroxymyristoyl)glucosamine N-acyltransferase	343	ZP_03318515	3e-167	ZP_03841597	8e-163
MmoY_01319	*fabZ*	(3R)-hydroxymyristoyl-[acyl-carrier-protein] dehydratase	151	ZP_03318514	1e-77	ZP_03841596	7e-78
MmoY_01320	*lpxA*	Acyl-[acyl-carrier-protein]--UDP-N-acetylglucosamine O-acyltransferase	264	ZP_03318513	6e-120	ZP_03841595	3e-115
MmoY_01321	*lpxB*	Lipid-A-disaccharide synthase	385	ZP_03318512	4e-175	ZP_03841594	7e-176
4-AraN loci							
MmoY_01597	*arnB_2*	UDP-4-amino-4-deoxy-L-arabinose--oxoglutarate aminotransferase	382	ZP_03319576	0	ZP_03839725	0
MmoY_01598	*arnC*	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	330	ZP_03319577	4e-168	ZP_03839726	9e-165
MmoY_01599	*arnA*	Polymyxin resistance protein (*pmrI*)	662	ZP_03319578	0	ZP_03839727	0
MmoY_01600	*arnD*	4-deoxy-4-formamido-L-arabinose-phosphoundecaprenol deformylase	56	ZP_03319579	1e-12	ZP_03839728	1e-12
MmoY_01601	*arnD*	Putative 4-deoxy-4-formamido-L-arabinose-phosphoundecaprenol deformylase	254	ZP_03319579	7e-121	ZP_03839728	2e-121
MmoY_01602	*arnT_1*	Undecaprenyl phosphate-alpha-4-amino-4-deoxy-L-arabinose arabinosyl transferase	554	ZP_03319580	0	ZP_03839729	0
MmoY_01603	*arnE*	Undecaprenyl phosphate-aminoarabinose flippase subunit	117	ZP_03319581	2e-32	ZP_03839730	3e-34
MmoY_01604	*arnF*	4-amino-4-deoxy-L-arabinose-phosphoundecaprenol flippase subunit	131	ZP_03319582	1e-34	ZP_03839731	1e-41
MmoY_01605	*arnT-2*	Undecaprenyl phosphate-alpha-4-amino-4-deoxy-L-arabinose arabinosyl transferase	566	ZP_03319580	6e-151	ZP_03839729	1e-154

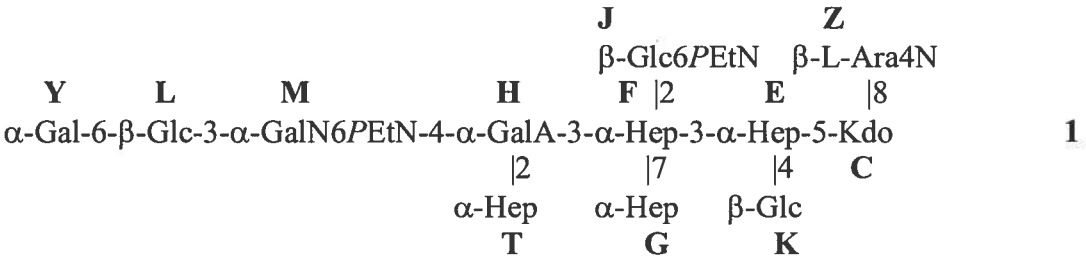
Table 4

Putative O-chain gene locus

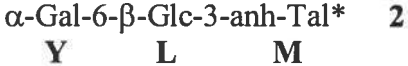
M. morganii gene #	Gene	Annotation	aa	*Pv. alcalifaciens* gene #	*E* value	*Pt. mirabilis* gene #	*E* value
MmoY_02773	*cpxA*	Sensor protein	457	ZP_03320660	0	ZP_03839160	0
MmoY_02774		CTP:phosphocholine cytidylyltransferase	299				
MmoY_02775		EamA-like transporter family protein	312				
MmoY_02776		Hypothetical protein	44				
MmoY_02777	*tagF*	CDP-glycerol:poly(glycerophosphate)glycerophospho-transferase	379				
MmoY_02778	*tagE*	Putative poly(glycerol-phosphate) α-glucosyltransferase	363				
MmoY_02779		Hypothetical protein	188				
MmoY_02780	*tagD*	Glycerol-3-phosphate cytidylyltransferase	178				
MmoY_02781		Hypothetical protein	162				
MmoY_02782		Putative glycosyl transferase	189				
MmoY_02783		Hypothetical protein	92				
MmoY_02784		Hypothetical protein	322				
MmoY_02785		Polysaccharide biosynthesis protein	259				
MmoY_02786		Hypothetical protein	50				
MmoY_02787		UDP-GlcNAc 4 epimerase	317	no hit		ZP_03839172	4e-178
MmoY_02788	*(trmL)*	tRNA (cytidine(34)-2'-O)-methyltransferase (*yibK*)	168	ZP_03320682	3e-84	ZP_03839173	2e-87

[a] Abbreviation used: aa, predicted length of the protein in amino acids. Alternate gene abbreviations are given in parentheses in the annotations.

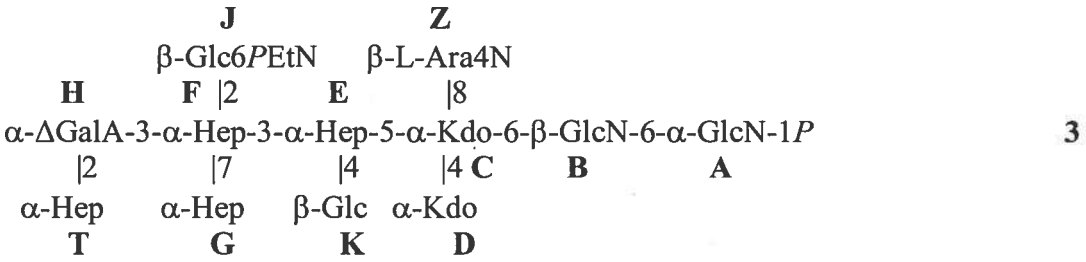
Figure 1


J Z
β-Glc6PEtN β-L-Ara4N
Y L M H F |2 E |8
α-Gal-6-β-Glc-3-α-GalN6PEtN-4-α-GalA-3-α-Hep-3-α-Hep-5-Kdo
|2 |7 |4 C
α-Hep α-Hep β-Glc
T G K

1


α-Gal-6-β-Glc-3-anh-Tal*
Y L M

2


J Z
β-Glc6PEtN β-L-Ara4N
H F |2 E |8
α-ΔGalA-3-α-Hep-3-α-Hep-5-α-Kdo-6-β-GlcN-6-α-GlcN-1P
|2 |7 |4 |4 C B A
α-Hep α-Hep β-Glc α-Kdo
T G K D

3