Pasteurella multocida Heddleston serovar 3 and 4 strains share a common lipopolysaccharide biosynthesis locus but display both inter- and intra-strain lipopolysaccharide heterogeneity.

Marina Harper1, Frank St. Michael2, Marietta John1, Evgeny Vinogradov2, Jennifer A. Steen1,5, Lieke van Dorsten2,6, Jason A. Steen1,5, Conny Turni3, Patrick J Blackall3, Ben Adler1, Andrew D. Cox2† and John D. Boyce1,4†*

1Australian Research Council Centre of Excellence in Structural and Functional Microbial Genomics, Department of Microbiology, Monash University, Melbourne, Australia, 3800.
2Institute for Biological Sciences, National Research Council, Ottawa, ON, Canada, K1A 0R6.
3Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Dutton Park, Brisbane, Australia 4102. 4Department of Microbiology, Monash University, Melbourne, Australia, 3800
†Joint senior author

Running title: Analyses of P. multocida serovar 3 and 4 LPS
Keywords: Core oligosaccharide/ Genetics/ LPS/ Pasteurella multocida/ Structure.

5Current address: The University of Queensland, St. Lucia, Queensland, Australia 4067
6Current address: Eppendorf Nederland B.V., Nijmegen, Netherlands

Correspondence:
Dr. John Boyce TEL: +61 3 9902 9179
Department of Microbiology FAX: +61 3 9902 9222
Building 76, Monash University, E-mail: john.boyce@monash.edu
Clayton, Victoria, Australia, 3800
ABSTRACT

*Pasteurella multocida* is a Gram-negative multi-species pathogen and the causative agent of fowl cholera, a serious disease of poultry which can present in both acute and chronic forms. The major outer membrane component lipopolysaccharide (LPS) is both an important virulence factor and a major immunogen. Our previous studies have determined the LPS structures expressed by different *P. multocida* strains and revealed that a number of strains belonging to different serovars contain the same LPS biosynthesis locus but express different LPS structures due to mutations within glycosyltransferase genes. In this study we report the full LPS structure of the serovar 4 type strain, P1662, and reveal that it shares the same LPS outer core biosynthesis locus, L3, with the serovar 3 strains, P1059 and Pm70. Using directed mutagenesis, the role of each glycosyltransferase gene in LPS outer core assembly has been determined. LPS structural analysis of 23 Australian field isolates that contain the L3 locus reveal that at least six different LPS outer core structures can be produced as a result of mutations within the LPS glycosyltransferase genes. Moreover, some field isolates produce multiple but related LPS glycoforms simultaneously and three LPS outer core structures are remarkably similar to the globo series of vertebrate glycosphingolipids. Our in-depth analysis showing the genetics and full range of *P. multocida* lipopolysaccharide structures will facilitate the improvement of typing systems and the prediction of the protective efficacy of vaccines.
INTRODUCTION

*Pasteurella multocida* is a Gram-negative pathogen that causes a range of diseases in wild and domestic animals (1). These diseases include fowl cholera in poultry (2), hemorrhagic septicemia and shipping fever in ruminants (3), and atrophic rhinitis and respiratory disease in pigs (4), all of which cause significant economic losses to primary industries worldwide. *P. multocida* can also cause human infections following animal bites or close contact with domestic animals (5).

Current classification of *P. multocida* strains combines Heddleston lipopolysaccharide (LPS) typing with Carter capsular typing, either by traditional serological methods (Heddleston and Carter) or multiplex PCR (Carter) (6-8). *P. multocida* strains are currently classified into 5 capsular serogroups (A, B, D, E and F) and 16 Heddleston LPS serovars (6, 8).

Protective immunity against *P. multocida* is generally considered to be humorally mediated, with protective antibodies directed primarily, but not exclusively, against the LPS antigen located on the surface of the *P. multocida* cell. Although both live and killed vaccines are available for protection against *P. multocida* infection, few afford good levels of protection against strains expressing different LPS (9, 10). Therefore, elucidating the full range of LPS structures expressed by *P. multocida*, including field isolates, and understanding the role of LPS in protective immunity is crucial for the formulation of effective, cross-protective vaccines.

The LPS structures produced by *P. multocida* type strains representing seven of the Heddleston serovars (1, 2, 5, 8, 9, 13, 14) have been determined, as well as the structure of the LPS expressed by the genome-sequenced serovar 3 strain, Pm70 (11-17). Some *P. multocida* strains can produce two LPS inner core glycoforms simultaneously, termed glycoform A and glycoform B. The inner core glycoform A is produced by all strains examined to date and contains a single phosphorylated 3-deoxy-D-manno-octulosonate (Kdo) residue that is substituted with a
phosphoethanolamine (PEtn) residue (16). This glycoform also contains a second glucose residue (Glc II) attached to the 6 position of the first heptose (Hep I). The inner core glycoform B contains two (Kdo) residues and does not have the additional Glc on Hep I. LPS structural analysis of a range of *P. multocida* strains representing seven serovars has revealed that the most variable part of the molecule is the outer core (structure beyond Glc I). Analysis of the corresponding LPS outer core biosynthesis locus in each strain revealed that although some shared a near identical locus, they expressed different LPS molecules due to point mutations or deletions within LPS biosynthesis genes (13). In this study we report that the LPS outer core biosynthesis locus, first identified in Pm70 (17, 18) and named L3 in this study, is found in the serovar 3 and serovar 4 type strains (P1059 and 1662) and in 23 Australian field isolates. Each of these strains express the same conserved inner core structure, but the strains display significant variability in the length of the LPS outer core and in the number of LPS glycoforms produced simultaneously. We also determined the role of each of the L3 LPS outer core glycosyltransferase genes in the assembly of the LPS outer core.
MATERIALS AND METHODS

Bacterial strains, plasmids, media and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1. Escherichia coli was grown routinely in Luria-Bertani broth. P. multocida was grown in nutrient broth, brain heart infusion (BHI) or heart infusion (HI) broth. Solid media were obtained by the addition of 1.5% agar (w/v). When required, the media were supplemented with kanamycin (50 μg/ml), spectinomycin (50 μg/ml) or tetracycline (2.5 μg/ml). To isolate LPS from the P. multocida serovar 4 type strain P1662, 2 L of an early log phase culture of P. multocida (grown in BHI at 37°C, shaking at 200 rpm) was used to inoculate 24 L of BHI in a 28 L NBS fermenter. The culture was grown for 18 h at 37°C, shaking at 200 rpm, with 24 L min⁻¹ aeration and 20% O₂ saturation. Following treatment with hyaluronidase for 1 h (1g, Sigma) to remove capsular material, phenol (2% final concentration) was added and the culture incubated at room temperature for 4 h to kill bacteria, which were then harvested with a Sharples continuous flow centrifuge. For compositional analysis of LPS produced by the Heddleston serovar 3 type strain P1059 and the Australian field isolates, small quantities of LPS were isolated from plate-grown cells as described previously (19).

DNA manipulations. Restriction digests and ligations were performed using enzymes and buffers obtained from NEB or Roche Diagnostics GmbH. Plasmid DNA and genomic DNA were prepared using the Plasmid Mini kit from QIAGEN and the Genomic DNA extraction kit from RBC respectively. PCR amplification of DNA was performed using Taq DNA polymerase or the Expand High Fidelity PCR System (Roche Diagnostics) and amplified fragments were purified using the Qiagen PCR Purification Kit. The oligonucleotides used in this study were synthesized by Sigma, Australia and are listed in Supplementary Table 1. Preparation of electro-competent P. multocida and electroporation conditions were as described previously (20).
sequencing reactions and cycle conditions using either plasmid DNA or PCR products as template were performed using the BigDye Terminator version 3.1 (Applied Biosystems) as per the manufacturer's instructions with the following modifications; 1.5 µl of 10 x Taq PCR buffer (Roche) and 1µl of BigDye Terminator version 3.1 were used in a 15 µl total volume. For direct sequencing from *P. multocida* genomic DNA, reaction and cycle conditions were performed as described previously (21) with the following modifications; reactions were performed in a 40 µl volume with 3 µl 10 x Taq PCR buffer (Roche), 13 pmol of oligonucleotide and 4 µl of BigDye Terminator version 3.1 (Applied Biosystems). All DNA sequences were determined on a capillary-platform Genetic Analyser (Applied Biosystems 3730) and analysed with Vector NTI Advance 11 (Invitrogen). Amino acid sequence alignments were conducted using BLAST and Clustal W2. Sequence data for the P1059 and P1662 outer core biosynthesis loci have been deposited in GenBank under accession numbers KF314825 and KF314826 respectively.

**Construction of *P. multocida* TargeTron® mutants.** For inactivation of genes in *P. multocida* strain P1059, we used the TargeTron® method (Sigma-Aldrich) of mutagenesis as described previously (20) but with the following modifications. To assist in the selection of *P. multocida* transformants containing an integrated intron, the plasmid pAL692 (20) was modified to introduce a kanamycin resistance gene, *aph3*, into the intron. Briefly, the *aph3* gene was amplified from pAL99 (Table 1) using the primers BAP6796 and BAP6797 then cloned into the MluI-digested pAL692, generating the plasmid pAL953 (Table 1). Retargeting of the intron within pAL953 was performed as per the TargeTron® user manual with oligonucleotides designed by the TargeTron® design site (Sigma-Aldrich) (Supplementary Table 1). Each plasmid containing an intron correctly targeted to the target gene (pAL1003, pAL1004 and pAL1006; Table 1) was used to transform the *P. multocida* serovar 3 type strain (P1059) by electroporation;
transformants containing the plasmid and/or integrated intron were selected on solid agar with kanamycin. Mutants were then cured of replicating plasmid by overnight growth in nutrient broth without antibiotic selection, and transformants containing the intron were identified by patching onto HI agar containing spectinomycin or kanamycin. Kanamycin resistant, spectinomycin sensitive transformants were then screened for the correct intron insertion by colony PCR using one oligonucleotide located within the target gene and one located within the intron (EBS universal). To confirm integration into the target gene, and to determine if there were any additional intron insertions elsewhere in the genome, direct sequencing from genomic DNA was performed using the intron-specific EBS universal primer (Table 1) together with DNA isolated from each TargeTron® mutant. Mutants that gave unambiguous sequencing data identical to the target gene sequence were selected for further study.

In trans complementation of mutants. For complementation of the kanamycin resistant TargeTron® mutants, the P. multocida expression plasmid pAL99 was modified to remove the kanamycin resistant gene, aph3, and replace it with the aadA or tetM gene, conferring spectinomycin (pAL99S) or tetracycline (pAL99T) resistance respectively (Table 1). The plasmid pAL99 was amplified using two outward firing primers (BAP5358 and BAP5359, Supplementary Table 1) flanked aph3, digested with EcoRV, then ligated to EcoRV-digested PCR-amplified fragment containing either the aadA gene (amplified from pUA826 using primers BAP5360 and BAP5361) or the tetM gene amplified from Tn916 (in strain AL523) using primers BAP7134 and BAP7135, to generate pAL99S and pAL99T respectively. For complementation of each of the P. multocida mutants, PCR amplifications were performed using P1059 genomic DNA and the appropriate pair of oligonucleotides that flanked the target gene (Supplementary Table 1). Each product was digested with the appropriate restriction
enzyme then ligated to the appropriately digested pAL99, pAL99S or pAL99T (Table 1), such that transcription of each gene would be driven by the constitutive *P. multocida tpiA* promoter. Each plasmid was then introduced into the corresponding *P. multocida* mutant, generating the complemented strain (Table 1). As a control, the appropriate empty vector was introduced separately into each mutant (Table 1).

**Isolation and purification of lipopolysaccharide.** LPS was isolated and purified as described previously (16). O-deacylated LPS (LPS-OH), core oligosaccharide (OS) and completely deacylated LPS were all isolated and purified from LPS as described previously (16).

**Analytical methods, mass spectrometry and nuclear magnetic resonance spectroscopy.** Sugars were determined as their alditol acetate derivatives and linkage analysis determined following methylation analysis by gas-liquid chromatography-mass spectrometry (GLC-MS) and electrospray-mass spectrometry (ES-MS) as described previously. To determine exact locations and linkage patterns of residues, nuclear magnetic resonance (NMR) techniques were performed as described previously (16). The assignment of $^1$H resonances was achieved using correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY). The assignment of $^{13}$C resonances was achieved using $^{13}$C-$^1$H heteronuclear single quantum coherence (HSQC) and $^{13}$C-$^1$H HSQC-TOCSY NMR. To identify the locations of PEtn substitutions, $^{31}$P-$^1$H heteronuclear multiple quantum coherence (HMQC) experiments were performed (16).
RESULTS

Analyses of LPS outer core biosynthesis loci. The region of the genome containing the LPS outer core biosynthesis locus in the serovar 3 and 4 type strains (P1059 and P1662 respectively) was sequenced and found to be 99% identical to each other and also to the equivalent locus in Pm70 (Fig. 1A)(18). However, the sequence of the locus in the serovar 4 type strain P1662 contained a nonsense mutation within natC due to a C291A nucleotide substitution. In keeping with our function-based nomenclature of P. multocida LPS glycosyltransferases (19), we have named the orthologues of the Pm70 glycosyltransferase genes in the serovar 3 and 4 type strains natC (pm1138), gatG (pm1139), natB (pm1140), gatF (pm1141), gctC (losA) and hptE (pm1144) (Fig. 1A). The LPS outer core biosynthesis locus common to all three strains, Pm70, P1059 and P1662, was named L3 (Fig. 1A).

Structural analyses of the P. multocida LPS derived from the serovar 3 and 4 type strains P1059 and P1662. Sugar analysis of the purified LPS isolated from the serovar 4 type strain P1662 revealed glucose (Glc), galactose (Gal) and L-glycero-D-manno-heptose (LD-Hep) in the approximate ratio of 2:1:3 respectively. A small amount of N-acetyl-glucosamine (GlcNAc) was also identified as is often observed from sugar analyses on LPS and presumably derives from the lipid A. In order to separate and detect closely related glycoforms, capillary electrophoresis coupled with electro-spray ionization mass spectrometry (CE-MS) was used to analyse the O-deacylated LPS (LPS-OH) (Table 2). This analysis revealed major triply charged ions at m/z 1025.4 and 1066.2 corresponding to a composition of 5Hex, 4Hep, Kdo-P, 2PEtn, O-deacylated lipid A (Lipid A-OH) for the smaller molecule and an additional phosphoethanolamine (PEtn) for the larger molecule. Glycoforms with one or no PEtn residues were also observed. Similarly, MS analysis of the non-fractionated core OS suggested a...
composition of 5Hex, 4Hep, 2PEtn, Kdo as the major glycoform. Interestingly, there was no
evidence of any glycoforms containing two Kdo residues (inner core glycoform B). Methylation
analysis of the fractionated core OS (Fraction 15) allowed determination of the linkage pattern of
the molecule which revealed the presence of terminal Glc, terminal Gal, 4-substituted Glc and
terminal LD-Hep in approximately equimolar amounts, with 6-substituted Glc, 2-substituted LD-
Hep, 4, 6-disubstituted LD-Hep and 3,4,6-trisubstituted LD-Hep identified in lower amounts.
In order to elucidate the exact locations and linkage patterns of the carbohydrate residues, NMR
studies were performed on the core oligosaccharide fraction that gave the most resolved and
homogeneous spectrum (Fraction 14). The inner core oligosaccharide \(^1\)H and \(^13\)C resonances
were assigned with reference to the published data for the structurally related oligosaccharides
from *P. multocida* strains belonging to serovars 1, 2, 3, and 5 and revealed that the serovar 4
strain elaborated the previously identified conserved inner core structure (Hep I-IV and Glc I-II)
(Table 3) (14-17).

The outer core residues of the oligosaccharide extension beyond Hep IV in strain P1662 were
characterized by COSY, TOCSY and NOESY experiments (Fig. 2). In addition to the residues of
the conserved inner core structure, a glucose residue (Glc III) was identified, based on a
characteristic spin-system in TOCSY experiments. Two spin systems were actually identified for
this residue with proton resonances at H-1 (4.76 ppm), H-2 (3.53 ppm), H-3 (4.13 ppm), H-4
(3.67 ppm), H-5 (3.73 ppm) and H-6ab (4.23, 4.10) in one system and at H-1 (4.75 ppm), H-2
(3.51 ppm), H-3 (4.11 ppm), H-4 (3.60 ppm), H-5 (3.60 ppm) and H-6ab (3.93, 3.75) in the other
system. Both exhibited inter-nuclear Overhauser effect (NOE) connectivities from the anomeric
protons to a resonance at 4.17 ppm (Fig. 2). This resonance was assigned as the proton at the 4-
position of the Hep IV residue by virtue of \(^13\)C-\(^1\)H HSQC and \(^13\)C-\(^1\)H HSQC-TOCSY
experiments (data not shown). The identification of two spin systems became apparent when $^{31}$P-$^1$H HMQC experiments were performed in conjunction with the homonuclear NMR data that revealed the presence of two PEtn residues linked to both the H-3 (4.13 ppm) and H-6ab (4.23, 4.10 ppm) resonances in one spin-system and to only the H-3 resonance (4.11 ppm) in the other system (Fig. 2). A further glucose residue (Glc IV) was identified, based on a characteristic spin-system in TOCSY experiments, with H-1 (4.72 ppm), H-2 (3.44 ppm), H-3 (3.70 ppm), H-4 (3.68 ppm), H-5 (3.66 ppm) and H-6ab (3.98, 3.81) resonances being identified. An inter-NOE connectivity from the anomeric proton of Glc IV to a resonance at 4.33 ppm was identified (Fig. 2). This resonance was assigned as the proton at the 6-position of the Hep IV residue by virtue of $^{13}$C-$^1$H HSQC and $^{13}$C-$^1$H HSQC-TOCSY experiments (data not shown). This assignment, coupled with the data for the Glc III residue, is consistent with the methylation analysis data which identified a 4,6-substituted LD-Hep residue. A terminal galactose residue (Gal I) was identified, based on a characteristic spin-system in TOCSY experiments, with H-1 (4.46 ppm), H-2 (3.56 ppm), H-3 (3.67 ppm) and H-4 (3.94 ppm) resonances being identified. An inter-NOE connectivity from the anomeric proton of Gal I to a resonance at 3.68 ppm was identified (Fig. 2). This resonance was assigned as the proton at the 4-position of the Glc IV residue by virtue of $^{13}$C-$^1$H HSQC and $^{13}$C-$^1$H HSQC-TOCSY experiments (data not shown). This assignment is also consistent with the methylation analysis data which identified a 4-linked Glc residue and a terminal Gal residue. The identification of two PEtn residues substituting the 3 and 6-positions of Glc III, along with evidence for a PEtn residue at the 3-position of Hep II in some glycoforms, combined with the identification of only 2 PEtn residues in total in the core OS, led us to examine closely the $^{31}$P-$^1$H heteronuclear and the homonuclear NMR data (Fig. 2). It became clear that the PEtn at the 6-position of the Glc III residue was present non-stoichiometrically;
however, the PEtn residue at the 3-position was found to be always present. Furthermore the
PEtn residue was identified at the 3-position of the Hep II residue (in the inner core) approximately 20% of the time. However, the simultaneous expression of all three PEtn residues in the core OS was not identified, perhaps suggesting a mutual exclusivity. Taken together, these data show that the serovar 4 type strain P1662 produces LPS that is a truncated derivative of the Pm70 LPS (17) with the exception of one or two additional PEtn residues on the Glc III of the P1662 outer core (Fig. 1B).

As the full structure of the LPS core oligosaccharide from the serovar 3 strain Pm70 has been determined previously (17) only CE-MS analysis of LPS-OH was conducted on the serovar 3 type strain P1059. Analysis of the P1059 LPS revealed that all LPS glycoforms lacked the terminal HexNAc; CE-MS analysis of LPS-OH revealed a major triply charged ion at \( m/z \) 1147.3 corresponding to a composition of HexNAc, 6Hex, 4Hep, 2PEtn, Kdo, Lipid A-OH (Table 2). To confirm this finding, we undertook CE-MS analysis of LPS-OH isolated from two independent freeze-dried cultures of strain P1059 that had been sourced from different collections and found that the LPS composition consisted mainly of HexNAc, 6Hex, 4Hep, 2PEtn, Kdo, Lipid A-OH but trace amounts of the 2HexNAc, 6Hex, 4Hep, 2PEtn, Kdo, Lipid A-OH glycoform were detected in both samples (data not shown). A trace amount of the inner core glycoform B (HexNAc, 5Hex, 4Hep, 2Kdo, PEtn, Lipid A-OH) was also detected in one of the samples (data not shown).

**Characterization of the glycosyltransferases involved in biosynthesis of the L3 LPS outer core.** The functions of two glycosyltransferases (HptE and NatC) encoded within the L3 LPS outer core biosynthesis locus have been previously determined: HptE is required for the addition of the first outer core sugar, Hep IV (19) and NatC (annotated as PM1138 in strain
Pm70) is an N-acetyl-galactosyltransferase required for the addition of the last outer core sugar GalNAc II (23) (Fig. 1). Previous bioinformatic analysis of the equivalent region in strain Pm70 indicated that pm1141 (glycosyltransferase family 25) and pm1139 (glycosyltransferase family 8) likely encode 1,4 galactosyltransferases that add Gal I and Gal II respectively, and pm1140 (glycosyltransferase family A) encodes an N-acetyl-galactosyltransferase that adds GalNAc I (17).

To confirm the predicted function of these enzymes, the pm1141, pm1139 and pm1140 orthologues (gatF, gatG, and natB respectively), were separately inactivated in the serovar 3 type strain P1059 and the effect of each mutation on LPS structure determined (Table 2). To confirm that any LPS structural changes were due specifically to inactivation of the mutated glycosyltransferase, each mutant was complemented with a functional copy of the appropriate transferase provided in trans. Structural analysis of the mutants revealed that the gatF mutant (AL2116) was unable to assemble LPS beyond the Glc IV and that the gatG mutant (AL2155) was unable to assemble LPS beyond the Gal I (Fig. 1, Table 2), confirming bioinformatic predictions that GatF is the galactosyltransferase which adds Gal I to the 4th position of the Glc IV and GatG is a galactosyltransferase which adds Gal II to the 4th position of Gal I (Fig. 1).

Similarly, analysis of the LPS produced by the P1059 natB mutant (AL2117) confirmed that NatB is the 1, 3 N-acetyl galactosyltransferase responsible for the addition of GalNAc I to Gal II (Fig. 1, Table 2). Importantly, in each case, complementation experiments showed that mutants provided with a functional copy of the appropriate transferase in trans produced full-length LPS (Table 2) while mutants transformed with vector only did not (data not shown).

To determine the role of gctC in LPS outer core assembly we used an LPS mutant, AL539, constructed in the serovar 1 strain VP161, which expresses only the highly conserved inner core...
LPS; including the GlcI which is the acceptor molecule for the HptE heptosyltransferase (19). This strain was transformed with the expression plasmid pAL446 containing a functional copy of the L3 glycosyltransferase genes hptE and getC (PCR amplified from Pm70) (Table 1). MS (Table 2) and NMR structural analyses (data not shown) of the LPS produced by the complemented strain (AL806) revealed that it elaborated an LPS outer core of two terminal Glc residues linked to the 4 and 6 position of the Hep IV, confirming the role of the L3 HptE as the 1,6 heptosyltransferase and identifying that GtcC is a bifunctional glucosyltransferase that adds glucose to both the 4 and 6 position of D-Hep IV.

**Analysis of genetics underlying P1662 LPS truncation.** The serovar 4 type strain, P1662, produced a truncated LPS structure terminating at Gal I, suggesting that the 1,4 galactosyltransferase, GatG, was non-functional in this strain. Analysis of the P1662 gatG sequence indicated that the translated product contained eight amino acid changes (R85K, V240A, A253D, N259K, R272K, E277K, V279M, T284N) when compared to the functional GatG glycosyltransferases expressed by Pm70 and P1059. The *Neisseria meningitidis* galactosyltransferase LgtC is 39% identical (57% conserved) to GatG and the crystal structure of LgtC has been determined both in complex with its donor and acceptor molecules (24). Alignment of LgtC with other family 8 glycosyltransferases, including GatG (PM1139), identified a number of highly conserved amino acids critical for the interaction with the donor molecule, the acceptor molecule, or the divalent cation Mn$^{2+}$ (24) (Fig. 3). Only one of the eight GatG substitutions identified in strain P1662 (R85K) was in a highly conserved position (R86 in LgtC) and this residue is critical for the interaction of LgtC with the galactose portion of the donor. Therefore, we suggest that the R85K is responsible for the loss of GatG function in P1662.
Structural and genetic analyses of LPS biosynthesis in *P. multocida* field isolates belonging to genotype L3. To determine the diversity of LPS structures produced by *P. multocida* field isolates belonging to the LPS genotype L3, we identified 23 *P. multocida* isolates from Australian poultry farms that contained L3 LPS outer core biosynthesis loci. Proposed compositions determined from CE-MS analysis of LPS-OH purified from each of the isolates (Supplementary Table 2) indicated that a number of different LPS structures were produced, ranging from inner core only structures (terminating at Glc I) to full-length structures terminating at the GalNAc II (Fig. 4; Table 4). Interestingly, 14 of the isolates simultaneously expressed multiple LPS glycoforms containing outer core sugar extensions ranging from 3 to 7 residues (Supplementary Table 2) and none of the field isolates analysed elaborated an LPS structure terminating with the Hep IV (Fig. 4, supplementary Table 2).

To determine if the truncated LPS structures observed in the field isolates were the result of mutations within one of the glycosyltransferase genes in the L3 LPS outer core biosynthesis locus, the glycosyltransferase gene predicted to be responsible for the arrested assembly in each case was amplified by PCR and the nucleotide sequence determined using the same oligonucleotides used for complementation experiments (Table 1). The translated amino acid sequences were then compared with those of the functional orthologs in P1059 and Pm70. In eight of the field isolates there were clear mutations resulting in an immediate stop codon or a reading frame shift leading to a stop codon (PM48, PM1075, PM1120, PM1103, PM1205, PM1258, PM1434, PM1439 - Table 4). In five field isolates (PM1, PM3, PM8, PM1369, PM1474) there were no changes detected in the gene examined although one isolate (PM1) had nucleotide changes and deletions in the predicted promoter region. However, seven of the field isolates contained one or more missense mutations in the genes encoding the predicted...
inactivated transferase; two strains (PM18 and PM146) contained mutations within $gatF$, two
had changes in $natC$ (PM1099 and PM1441) and three (PM1268, PM1320, PM1098) contained
mutations in $gatG$ (Table 4).

As noted previously, the structure of the $N. meningitidis$ LgtC (a GatG ortholog) is known and
alignments of GatG from PM1098, PM1268 and PM1320 (each containing missense mutations)
with LgtC revealed that there were amino acid changes located in close proximity to residues
that are known to be critical in LgtC. These residues included those required for LgtC to interact
with the UDP portion of the sugar donor (I104, G247 and K250), the divalent cation Mn$^{2+}$
(D103, D105 and H244) or the acceptor molecule (D130 and F132) (24) (Fig. 3). Therefore, it is
highly likely that the missense mutations in GatG abrogate transferase function in these field
isolates. There are no data available on the structure of GatF and NatC homologs in other
bacteria so it is not possible to predict if the missense mutations identified in $gatF$ and $natC$ were
responsible for the altered LPS production observed in those strains.

DISCUSSION
The strains belonging to the L3 LPS genotype include the first genome sequenced $P. multocida$
strain, Pm70, and the Heddleston serovar 3 and 4 type strains (P1059 and P1662 respectively).
Using LPS structural analysis of directed mutants followed by in trans complementation, the role
of each of the uncharacterized transferases in the L3 locus (NatB, GatG, GatF and GctC) was
defined and previous functional predictions were confirmed (17). Interestingly, P1059 elaborates
only trace quantities of LPS glycoforms containing GalNAc II. However, when the P1059 $gatF$,
$gatG$ and $natB$ mutants were complemented with the appropriate wild-type genes all produced
significant levels of the full length glycoform containing the terminal GalNAc II (Table 2). This
finding strongly suggests that the efficiency of addition of GalNAc II by NatC is dependent on
the rate of production of the appropriate LPS acceptor molecule (extended to GalNAc I), which
is in-turn dependent on the efficiency of the transferases involved in the earlier steps of LPS
assembly.

LPS structural analysis of the strains belonging to the LPS genotype L3, which includes the
serovar 3 and 4 type strains, Pm70 and the 23 field isolates analyzed here revealed that all
produced LPS containing the inner core glycoform A (containing Kdo-P) that is the most
common LPS inner core structure produced by *P. multocida*. However, the inner core glycoform
B (containing Kdo-Kdo), previously reported by us to be expressed in significant quantities in
the serovar 1 strains VP161 and X73 (15, 16), was not detected in strain P1662 or in any of the
L3 Australian field isolates; although glycoform B was detected in trace amounts in strain P1059
and was previously found in LPS isolated from Pm70 (17). Although all strains share the same
LPS outer core biosynthesis locus, L3, the length of the LPS outer core structure varied
considerably; Pm70 produces significant quantities of the fully extended LPS structure, whereas
nearly all LPS glycoforms in strain P1059 lacked the terminal GalNAc II, while the longest of
the LPS glycoforms produced by P1662 terminated at Gal I. Whilst some of this variation is due
to clear mutations inactivating genes within the L3 locus, some field isolates produced a fully
extended LPS glycoform as well as up to four shorter, but related, outer core structures. This
intra-strain variation in LPS structure as the result of truncations in the outer core region has
been previously reported in *Haemophilus influenzae* (25), but has not been observed in strains
belonging to any of the other *P. multocida* serovars examined to date. There is no genetic
evidence of any phase variation in *P. multocida*, but it is possible that the expression of multiple
glycoforms from the same locus could be controlled post-transcriptionally via an unknown
mechanism. Alternatively, MsbA, the ABC transporter responsible for "flipping" LPS across the inner membrane (26), may transport an unusually high number of LPS glycoforms before the outer core region of the LPS has fully extended.

The LPS produced by strains belonging to the L3 LPS genotype contains outer core structures remarkably similar to the oligosaccharide component of the P, P^K and Forssman antigens found on the surface of many cell types in vertebrates, including birds (27). Forssman antigens have been identified in birds in the vascular endothelial tissue, perivascular connective tissue, and hematopoietic organs including the spleen (28). Forssman antigen is also present on the surface of chicken erythrocytes (29). Notably, P and P^K antigen mimicry has been reported in other pathogens that express O-antigen-deficient LPS, including *Campylobacter jejuni* (30), *Haemophilus influenzae* (31) and *Neisseria* spp. (32).

In summary, we have shown that fowl cholera strains containing the LPS outer core biosynthesis locus L3 display a remarkable level of inter- and intra-strain LPS structural heterogeneity. This group includes the serovar 3 and 4 type strains (P1059 and P1662) as well as the first genome-sequenced strain Pm70. This level of LPS diversity, particularly within a single strain and derived from a single LPS outer core biosynthesis locus, has not been observed in any other *P. multocida* strains examined to date. The role that these multiple LPS glycoforms play in disease progression and immunity is not known, but three of the outer core structures expressed are identical to the carbohydrate structure in the globo series of vertebrate glycosphingolipids P, P^K and Forssman antigen. LPS is considered a major immunogen of Gram-negative bacteria, but as the vertebrate immune system is tolerant of self antigens there may be a failure to respond appropriately to strains expressing these antigens. Thus, it is likely that the expression of L3 LPS
molecules may aid the in vivo survival/persistence of the bacteria through avoidance of the host immune system.
ACKNOWLEDGEMENTS

We thank Mark Edmunds for the PCR analysis of the Australian field isolates and our NRC-IBS colleagues Perry Fleming and the Bacterial Culture Facility for large scale biomass production and Jacek Stupak for mass spectrometry. This work was funded in part by the Australian Research Council, Canberra, Australia and the Poultry Cooperative Research Centre, Armidale, Australia.
REFERENCES


Figure 1.

Panel A: Schematic representation of the genetic organization of the *P. multocida* L3 LPS outer core biosynthesis locus, common to serovars 3 and 4. Genes involved in LPS outer core assembly are shown in white. Conserved genes unrelated to LPS biosynthesis are shown in gray. Original gene annotation numbers for Pm70 are shown below.

Panel B: Schematic representation of the LPS structure expressed by *P. multocida* serovar 3 and 4 type strains, P1059 and P1662 respectively. Only inner core glycoform A is shown. The *P. multocida* glycosyltransferase genes predicted to be required for the assembly of the outer core are shown below each linkage on the structure expressed by P1059. The rare addition of GalNAcII onto the P1059 LPS is shown with a dotted line and in gray. The specific number and position of phosphoethanolamine (PEtn) residues attached to each LPS structure are strain dependent; non-stoichiometric additions of PEtn are shown with a dotted line. Regions of the L3 LPS that are identical to the oligosaccharide component of the vertebrate glycosphingolipids, Forssman, P and Pk, are shown boxed in the Pm70 structure. Residues are GalNAc, N-acetyl-galactosamine; Gal, galactose; Glc, glucose; Hep, heptose; PEtn, phosphoethanolamine; KDO, 3-deoxy-D-manno-octulosonate; P, phosphate.

Figure 2.

NMR analyses of core OS Fraction 14 from *P. multocida* Hedleston strain serovar 4. Regions of overlapped COSY (green), TOCSY (red) and NOESY (purple) spectra (two lower panels) and \(^1\)H-\(^{31}\)P HMQC (green / purple) spectrum (upper panel) of the core OS. Residues are labeled as indicated in Table 3.
Amino acid sequence alignment of GatG (PM1139) from Pasteurella multocida strain Pm70 (Swiss-Prot: Q9CLS0) aligned with the galactosyltransferase LgtC from Neisseria meningitidis (Swiss-Prot: P96945). Amino acids that are highly conserved within LgtC-related glycosyltransferases are highlighted in grey (24). LgtC amino acids involved in binding to Mg$^{2+}$ are marked with square. Amino acids known in LgtC to interact with the UDP or galactose region of the nucleotide sugar UDP-galactose are marked with open and filled circles respectively while those involved in binding to the acceptor molecule are marked with a triangle (24).

Schematic representation showing the range of LPS outer core structures expressed by the 23 Australian field isolates containing the L3 outer core biosynthesis locus. Only the longest LPS outer core structure is shown for strains expressing multiple LPS glycoforms. The specific number and position of phosphoethanolamine (PEtn) residues attached to each LPS structure are strain dependent and are not shown. Residues are GalNAc, N-acetyl-galactosamine; Gal, galactose; Glc, glucose; Hep, heptose.
Table 1. Bacterial strains and plasmids used in this study.

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant description</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli DH5α</td>
<td>deoR, endA1, ggyA96, hsdR17(k- m-), recA1, relA1, supE44, thi-1, (lacZYA-argFV169), Φ80lacZ ΔM15, F-</td>
<td>Bethesda Research Laboratories</td>
</tr>
<tr>
<td><strong>P. multocida</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL523</td>
<td>P. multocida VP161 strain containing Tn916 encoding tetM.</td>
<td>This study</td>
</tr>
<tr>
<td>AL539</td>
<td>VP161 insertional gatA mutant generated using pAL290. No expression of hptE due to polar effects of the gatA insertion.</td>
<td>This study</td>
</tr>
<tr>
<td>AL806</td>
<td>VP161 gatA mutant containing pAL446</td>
<td></td>
</tr>
<tr>
<td>AL2116</td>
<td>P1059 gatF (pm1141) TargeTron mutant of P1059</td>
<td>This study</td>
</tr>
<tr>
<td>AL2117</td>
<td>P1059 natB (pm1140) TargeTron mutant of P1059</td>
<td>This study</td>
</tr>
<tr>
<td>AL2155</td>
<td>P1059 gatG (pm1139) TargeTron mutant of P1059</td>
<td>This study</td>
</tr>
<tr>
<td>AL2192</td>
<td>P1059 gatG mutant containing pAL1022</td>
<td>This study</td>
</tr>
<tr>
<td>AL2297</td>
<td>P1059 gatF mutant containing pAL1046</td>
<td>This study</td>
</tr>
<tr>
<td>AL2299</td>
<td>P1059 natB mutant containing pAL1045</td>
<td>This study</td>
</tr>
<tr>
<td>AL2354</td>
<td>P1059 gatF mutant containing vector pAL99T</td>
<td>This study</td>
</tr>
<tr>
<td>AL2355</td>
<td>P1059 gatF mutant containing vector pAL99T</td>
<td>This study</td>
</tr>
<tr>
<td>AL2395</td>
<td>P1059 gatG mutant containing vector pAL99S</td>
<td>This study</td>
</tr>
<tr>
<td>P1059</td>
<td>Heddleston serovar 3 type strain.</td>
<td>(33)</td>
</tr>
<tr>
<td>P1662</td>
<td>Heddleston serovar 4 type strain</td>
<td>(33)</td>
</tr>
<tr>
<td>PM1</td>
<td>P. multocida turkey isolate</td>
<td>(34)</td>
</tr>
<tr>
<td>PM3</td>
<td>P. multocida turkey isolate</td>
<td>(34)</td>
</tr>
<tr>
<td>PM8</td>
<td>P. multocida turkey isolate</td>
<td>(34)</td>
</tr>
<tr>
<td>PM18</td>
<td>P. multocida chicken isolate</td>
<td>(34)</td>
</tr>
<tr>
<td>PM48</td>
<td>P. multocida chicken isolate</td>
<td>(34)</td>
</tr>
<tr>
<td>PM72</td>
<td>P. multocida chicken isolate</td>
<td>(34)</td>
</tr>
<tr>
<td>PM146</td>
<td>P. multocida chicken isolate</td>
<td>(35)</td>
</tr>
<tr>
<td>PM1075</td>
<td>P. multocida field isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1098</td>
<td>P. multocida field isolate</td>
<td>MRG culture collection</td>
</tr>
<tr>
<td>PM1099</td>
<td>P. multocida field isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1103</td>
<td>P. multocida field isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1120</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1153</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1205</td>
<td>P. multocida emu isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1258</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1268</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1320</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1369</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1434</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1439</td>
<td>P. multocida chicken isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1441</td>
<td>P. multocida turkey isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1470</td>
<td>P. multocida turkey isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td>PM1474</td>
<td>P. multocida duck isolate</td>
<td>MRG* culture collection</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pAL99</td>
<td>P. multocida expression plasmid (KanR) which contains the P. multocida constitutive Ptpi promoter.</td>
<td>(36)</td>
</tr>
<tr>
<td>pAL99S</td>
<td>P. multocida expression plasmid (SpecR); derivative of pAL99</td>
<td>This study</td>
</tr>
<tr>
<td>pAL99T</td>
<td>P. multocida expression plasmid (TetR); derivative of pAL99</td>
<td>This study</td>
</tr>
<tr>
<td>pAL290</td>
<td>pUA826 derivative containing an internal fragment of gatA</td>
<td>This study</td>
</tr>
<tr>
<td>pAL446</td>
<td>Complete getC and hptE genes from Pm70 cloned into pAL99</td>
<td>This study</td>
</tr>
<tr>
<td>pAL692</td>
<td>P. multocida plasmid (SpecR) containing TargeTron Group II intron.</td>
<td>(20)</td>
</tr>
<tr>
<td>Vector</td>
<td>Description</td>
<td>Reference</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td>pAL953</td>
<td>pAL692 with <em>aph</em>&lt;sub&gt;3&lt;/sub&gt; (Kan&lt;sup&gt;R&lt;/sup&gt;) inserted into TargeTron Group II intron.</td>
<td>This study</td>
</tr>
<tr>
<td>pAL1003</td>
<td>pAL953 retargeted to <em>natB</em> using primers BAP6851/6852/6853</td>
<td>This study</td>
</tr>
<tr>
<td>pAL1004</td>
<td>pAL953 retargeted to <em>gatG</em> using primers BAP6854/6855/6856</td>
<td>This study</td>
</tr>
<tr>
<td>pAL1006</td>
<td>pAL953 retargeted to <em>gatF</em> using primers BAP6864/6865/6866</td>
<td>This study</td>
</tr>
<tr>
<td>pAL1022</td>
<td>Complete <em>gatG</em> gene cloned into pAL99S</td>
<td>This study</td>
</tr>
<tr>
<td>pAL1045</td>
<td>Complete <em>natB</em> gene cloned into pAL99T</td>
<td>This study</td>
</tr>
<tr>
<td>pAL1046</td>
<td>Complete <em>gatF</em> gene cloned into pAL99T</td>
<td>This study</td>
</tr>
</tbody>
</table>
| pUA826  | *Mob*<sup>+</sup>, R6K replicon, single-crossover mutagenesis vector Ap<sup>R</sup>, Spc<sup>R</sup> (37) | (37)      

*Microbiology Research Group, Queensland Alliance for Agriculture and Food Innovation, University of Queensland, Australia.*
Table 2: Negative ion CE-ES-MS data and proposed compositions of O-deacylated LPS (LPS-OH) and core OS from the Heddleston serovar 3 and 4 type strains (P1059 and P16612), and proposed compositions of LPS-OH from the P1059 Targetron® mutants.

<table>
<thead>
<tr>
<th>Serovar study</th>
<th>(Description)</th>
<th>Observed ions (m/z)</th>
<th>Mol. Mass (Da)</th>
<th>Relative Intensity</th>
<th>Proposed Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(M-H)</td>
<td>(M-4H)</td>
<td>Observed</td>
<td>Calculated</td>
</tr>
<tr>
<td>Serovar 3 (P1059)</td>
<td>LPS-OH</td>
<td>-</td>
<td>1106.5</td>
<td>3322</td>
<td>3320.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>1147.3</td>
<td>3444</td>
<td>3443.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>889.8</td>
<td>1188.2</td>
<td>3568</td>
<td>3566.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serovar 4 (P1662)</td>
<td>LPS-OH</td>
<td>-</td>
<td>707.4</td>
<td>2833.6</td>
<td>2831.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>738.0</td>
<td>2956.0</td>
<td>2954.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>768.6</td>
<td>3079.0</td>
<td>3077.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>799.5</td>
<td>3202.2</td>
<td>3200.7</td>
</tr>
<tr>
<td>Core OS</td>
<td></td>
<td>-</td>
<td>-</td>
<td>899.1</td>
<td>1800.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>960.3</td>
<td>1922.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>1022.1</td>
<td>2046.2</td>
</tr>
<tr>
<td>GetC study</td>
<td>AL539 LPS-OH</td>
<td>-</td>
<td>709.7</td>
<td>1064.9</td>
<td>2131.9</td>
</tr>
<tr>
<td></td>
<td>(gat4 polar mutant)</td>
<td>-</td>
<td>717.0</td>
<td>1073.9</td>
<td>2151.9</td>
</tr>
<tr>
<td></td>
<td>AL806 LPS-OH</td>
<td>-</td>
<td>758.0</td>
<td>1137.4</td>
<td>2276.9</td>
</tr>
<tr>
<td></td>
<td>(gatC, lptF)</td>
<td>-</td>
<td>773.5</td>
<td>1160.9</td>
<td>2323.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>780.5</td>
<td>1171.5</td>
<td>2344.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-</td>
<td>822.1</td>
<td>1233.2</td>
<td>2469.2</td>
</tr>
<tr>
<td></td>
<td>AL2116 LPS-OH</td>
<td>-</td>
<td>691.8</td>
<td>923.2</td>
<td>2772.1</td>
</tr>
<tr>
<td>(gatF mutant)</td>
<td></td>
<td>-</td>
<td>728.5</td>
<td>971.4</td>
<td>2917.6</td>
</tr>
<tr>
<td></td>
<td>AL2297 LPS-OH</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Serovar (Description)</td>
<td>Observed Ions (m/z) (M-H)⁻ (M-4H)⁴⁻ (M-3H)³⁻ (M-2H)²⁻</td>
<td>Mol. Mass (Da) a</td>
<td>Relative Intensity</td>
<td>Proposed Composition</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td>--------------------------------------------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td><strong>GatG study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL2155 LPS-OH</td>
<td>- 768.9 - 1025.3 - 3079.2 - 3077.7</td>
<td>1.0</td>
<td>5Hex, 4Hep, Kdo-P, 2PEtn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(gatG mutant)</td>
<td>- 799.6 - 1066.2 - 3202.0 - 3200.8</td>
<td>1.0</td>
<td>5Hex, 4Hep, Kdo-P, 3PEtn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL2192 LPS-OH</td>
<td>- - 1106.5 - 3322.5 - 3320.0</td>
<td>1.0</td>
<td>HexNAc, 6Hex, 4Hep, Kdo-P, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(gatG mutant + gatG)</td>
<td>- - 1147.5 - 3445.5 - 3443.1</td>
<td>0.9</td>
<td>HexNAc, 6Hex, 4Hep, Kdo-P, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- - 1174.0 - 3525.0 - 3523.2</td>
<td>0.3</td>
<td>2HexNAc, 6Hex, 4Hep, Kdo-P, PEn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- - 1215.0 - 3648.0 - 3646.3</td>
<td>0.4</td>
<td>2HexNAc, 6Hex, 4Hep, Kdo-P, 2PEn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NatB study</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL2117 LPS-OH</td>
<td>- 768.7 - 1025.4 - 3079.0 - 3077.7</td>
<td>0.8</td>
<td>5Hex, 4Hep, Kdo-P, 2PEn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(natB mutant)</td>
<td>- 809.4 - 1079.5 - 3241.5 - 3239.9</td>
<td>1.0</td>
<td>6Hex, 4Hep, Kdo-P, 2PEn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL2299 LPS-OH</td>
<td>- - 1174.0 - 3525.0 - 3523.2</td>
<td>0.4</td>
<td>2HexNAc, 6Hex, 4Hep, Kdo-P, PEtn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(natB mutant + natB)</td>
<td>- - 1215.0 - 3648.0 - 3646.3</td>
<td>1.0</td>
<td>2HexNAc, 6Hex, 4Hep, Kdo-P, 2PEn, Lipid A-OH</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Average mass units were used for calculation of molecular weight based on proposed composition as follows: LipidA-OH, 952.00; Hex, 162.15; Hep, 192.17; Kdo, 220.18; anhydro (a) Kdo, 202.18; PEtn, 123.05; P, 79.95.

aKdo represents the hydrolysis products of Kdo formed during core hydrolysis.
Table 3: $^1$H- and $^{13}$C-NMR chemical shifts\(^a\) for the core oligosaccharide (OS) from the Pasteurella multocida serovar 4 type strain, P1662.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Chemical shift</th>
<th>Inter NOE's</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H-1 (C-1)</td>
<td>H-2 (C-2)</td>
</tr>
<tr>
<td>Hep I</td>
<td>5.03 (97.7)</td>
<td>3.97 (71.4)</td>
</tr>
<tr>
<td>E</td>
<td>5.82 (99.7)</td>
<td>4.28 (80.1)</td>
</tr>
<tr>
<td>Hep II(^+) + 3-P</td>
<td>5.65 (100.6)</td>
<td>4.18 (80.8)</td>
</tr>
<tr>
<td>F(^b)</td>
<td>5.65 (100.3)</td>
<td>4.15 (70.6)</td>
</tr>
<tr>
<td>Hep III</td>
<td>5.11 (102.8)</td>
<td>4.02 (71.7)</td>
</tr>
<tr>
<td>G</td>
<td>4.95 (100.3)</td>
<td>4.15 (70.6)</td>
</tr>
<tr>
<td>β-Glc (Glc I)</td>
<td>4.63 (104.4)</td>
<td>3.51 (74.4)</td>
</tr>
<tr>
<td>J</td>
<td>4.76 (103.4)</td>
<td>3.53 (74.0)</td>
</tr>
<tr>
<td>α-Glc (Glc II)</td>
<td>5.21 (102.5)</td>
<td>3.59 (73.1)</td>
</tr>
<tr>
<td>H</td>
<td>4.76 (103.4)</td>
<td>3.53 (74.0)</td>
</tr>
<tr>
<td>β-Glc (Glc III)(^1) + 3,6-P</td>
<td>4.76 (103.4)</td>
<td>3.53 (74.0)</td>
</tr>
<tr>
<td>M(^b)</td>
<td>4.75 (103.4)</td>
<td>3.51 (74.0)</td>
</tr>
<tr>
<td>β-Glc (Glc IV)</td>
<td>4.72 (104.5)</td>
<td>3.44 (74.2)</td>
</tr>
<tr>
<td>L</td>
<td>4.76 (103.2)</td>
<td>3.56 (71.3)</td>
</tr>
</tbody>
</table>

\(^a\) Chemical shifts were recorded at 25°C; referenced to internal acetone at 2.225 ppm ($^1$H)/ 31.07 ppm ($^{13}$C).

\(^b\) Signals for ethanolamine of phosphoethanolamine at F-3, M-3 and M-6 were observed at 4.19/63.1, 3.31/41.3.
Table 4. Detected mutations in LPS outer core biosynthesis loci of the *P. multocida* field isolates

<table>
<thead>
<tr>
<th>Longest LPS outer core structure detected</th>
<th>Predicted gene disrupted</th>
<th>Field isolates</th>
<th>Detected mutations$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>no detectable outer core</td>
<td><em>hptE</em></td>
<td>PM1075, PM1120, PM1258</td>
<td>Q229*, Q194* Single nucleotide deletion at position 61</td>
</tr>
<tr>
<td>Hep [Glc] Glc</td>
<td><em>gatF</em></td>
<td>PM3, PM8, PM18, PM146</td>
<td>none detected S41P, V138M</td>
</tr>
<tr>
<td>Hep [Glc] Glc-Gal</td>
<td><em>gatG</em></td>
<td>PM1098, PM1205, PM1268, PM1320, PM1439</td>
<td>Insertion of 42 nucleotides at position 353 14 nucleotides deleted, position 496-509</td>
</tr>
<tr>
<td>Hep [Glc] Glc-Gal-Gal</td>
<td><em>natB</em></td>
<td>PM1369, PM1474</td>
<td>None detected None detected</td>
</tr>
<tr>
<td>Hep [Glc] Glc-Gal-Gal-GalNAc</td>
<td><em>natC</em></td>
<td>PM1, PM48, PM1099, PM1103, PM1441, PM1434</td>
<td>Nucleotide changes/deletions in promoter region 14 nucleotides deleted, position 496-509</td>
</tr>
</tbody>
</table>

$^a$ Glc residue in square brackets is attached to the main chain residues via a 1,4 linkage to Hep.

$^b$ Compared to nucleotide and amino acid sequence of the equivalent functional orthologues in Pm70 and/or P1059. * = Stop codon. $^c$ GatG residue changes shown in bold are in close proximity to amino acids predicted to interact with the donor or acceptor molecule, or the divalent cation Mg$^{2+}$. 
<table>
<thead>
<tr>
<th>LPS outer core structure</th>
<th>No. field isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-Glc III</td>
<td>Inner core</td>
</tr>
<tr>
<td>1-GalNAc; 1,3; 3-GalNAc; 1,6; 6-Gal I II III; 2-Gal I II III; 3-Gal I II III</td>
<td>n = 3</td>
</tr>
<tr>
<td>p-Glc III</td>
<td>Inner core</td>
</tr>
<tr>
<td>1-GalNAc; 1,3; 3-GalNAc; 1,6; 6-Gal I II III; 2-Gal I II III; 3-Gal I II III</td>
<td>n = 6</td>
</tr>
<tr>
<td>p-Glc III</td>
<td>Inner core</td>
</tr>
<tr>
<td>1-GalNAc; 1,3; 3-GalNAc; 1,6; 6-Gal I II III; 2-Gal I II III; 3-Gal I II III</td>
<td>n = 2</td>
</tr>
<tr>
<td>p-Glc III</td>
<td>Inner core</td>
</tr>
<tr>
<td>1-GalNAc; 1,3; 3-GalNAc; 1,6; 6-Gal I II III; 2-Gal I II III; 3-Gal I II III</td>
<td>n = 5</td>
</tr>
<tr>
<td>p-Glc III</td>
<td>Inner core</td>
</tr>
<tr>
<td>1-GalNAc; 1,3; 3-GalNAc; 1,6; 6-Gal I II III; 2-Gal I II III; 3-Gal I II III</td>
<td>n = 4</td>
</tr>
<tr>
<td>p-Glc III</td>
<td>Inner core</td>
</tr>
<tr>
<td>1-GalNAc; 1,3; 3-GalNAc; 1,6; 6-Gal I II III; 2-Gal I II III; 3-Gal I II III</td>
<td>n = 0</td>
</tr>
<tr>
<td>L-Lys; Hep IV IV III</td>
<td>Inner core</td>
</tr>
<tr>
<td></td>
<td>n = 0</td>
</tr>
<tr>
<td>L-Lys; Hep IV IV III</td>
<td>Inner core</td>
</tr>
<tr>
<td></td>
<td>n = 3</td>
</tr>
</tbody>
</table>