Identification and characterisation of a biosynthetic locus for *Moraxella bovis* lipooligosaccharide

Isabelle Faglin¹ [◇], I. Darren Grice^{1,2} [◇], S.R.A.M. Eranda Ratnayake¹, Terese-Marie Daal¹,
Sanjesh Singh¹, Jennifer C. Wilson^{2,3}, Ian R. Peak^{1,2}

¹ Institute for Glycomics, ² School of Medical Science, ³ Menzies Health Institute, Molecular Basis of Disease, Griffith University, Gold Coast Campus, Southport, QLD 4222, Australia.

For correspondence: i.peak@griffith.edu.au, jennifer.wilson@griffith.edu.au

[◇] Contributed equally to this study

Abstract:

Moraxella bovis is a Gram-negative gammaproteobacterium and is one of the causative agents of infectious bovine keratoconjunctivitis. The structure of lipooligosaccharide (LOS) from strain Epp63 was recently elucidated. In the present study a genetic locus of seven genes encoding genes with high similarity to glycosyltransferases has been identified. Mutation of these putative glycosyltransferase genes resulted in *M. bovis* mutant bacteria that expressed truncated LOS structures. The structures of the oligosaccharide (OS) expressed by the mutant strains were elucidated and demonstrated the role of the glycosyltransferase enzymes in the LOS biosynthesis of *M. bovis*. The glycosyltransferase genes designated *lgt1*, *lgt3*, and *lgt6* are highly similar to the genes in the related bacterium *M. catarrhalis*. In addition, there are syntenic similarities with the corresponding LOS biosynthesis locus in *M. catarrhalis* and other members of Moraxellaceae.

Introduction:

The Gram-negative bacterium *Moraxella bovis* is one of the primary etiologic agents of Infectious Bovine Keratoconjunctivitis (IBK, also known as Pink Eye). Transmission of *M. bovis* through respiratory secretions, and possibly by arthropods, can rapidly lead to spread of the infection throughout a herd. Strain-specific vaccines based on whole killed organisms can be effective, but may provide limited protection [1]. Moreover, attempts at subunit vaccination are not broadly protective [2].

Economic impacts due to *M. bovis*-induced Pink Eye may be significant, particularly with regards to loss of milk production, and also significant weight reduction and loss of condition in affected animals [3]. To date, the bacterium is relatively poorly studied and in particular, little is known regarding the cell surface carbohydrate structures that are possible virulence factors for this microorganism. Previously we have elucidated the structure of the oligosaccharide (OS) derived from lipo-oligosaccharide (LOS) produced by *M. bovis* strain Epp63 [4]. *M. bovis* Epp63 LOS has several unusual features relative to other LOS structures, including a lack of heptose residues in its inner core OS that link via Kdo to the lipid A, and an unusual terminal open chain acetal-linked *N*-acetylgalactosamine, [(1*S*)-GalaNAc residue]→4,6-linked to a sub-terminal galactose residue [4]. Here we describe the identification and characterisation of a genetic locus encoding the glycosyltransferase genes responsible for synthesis of the OS of *M. bovis* LOS, and compare it to those identified in the human pathogen *M. catarrhalis*.

Results and Discussion

Identification and *in silico* analysis of *M. bovis* LOS biosynthesis locus:

We identified several contigs from the incomplete genome sequence project of *M. bovis* Epp63 (Baylor College, Medicine). As LOS biosynthesis genes are frequently located in a single locus, we hypothesised that these contigs may form a single genomic region in the chromosome of *M. bovis* that is responsible for the biosynthesis of LOS. To assess this, oligonucleotide primers were designed. The PCR results and subsequent sequencing identified a gap of approximately 500 bp between contigs 389 and 396, and approximately 400 bp between the available sequence of contigs 396 and 288. The sequencing allowed assembly into a single contig of approximately 9 kbp encompassing all seven putative LPS glycosyltransferase genes, preceded by a putative lipid A acyltransferase (Figure 1, Table 1, accession number KP795984). Four of the encoded proteins show >75% similarity to LOS glycosyltransferases from the human mucosal pathogen *M. catarrhalis* [5-9]. Two distinct genes were similar to *M. catarrhalis* *lgt2*, but sequence comparisons showed that the *M. catarrhalis* alleles are more similar to each other than to the *M. bovis* genes (Tables 1 and 2). One other gene did not have a *M. catarrhalis* homologue, with its closest match being to an uncharacterised putative glycosyltransferase found in *Acinetobacter radioresistens*, (another member of the Moraxellaceae). This gene is referred to as *lgt7*.

Given the similarity between the structures of *M. catarrhalis* and *M. bovis* LOS, we hypothesised that the *M. bovis* glycosyltransferase homologues have similar functions to those of *M. catarrhalis*. To address this, we mutated several of the genes in the locus as described below.

Generation of mutants

Genes were amplified and cloned, and a kanamycin resistance determinant was inserted into conveniently located restriction sites to inactivate the particular glycosyltransferase gene under investigation (see Figure 2). *M. bovis* strains Epp63 and Mb25 were transformed with either linearised plasmid or PCR products encompassing the mutant allele, resulting in a double cross-over allelic replacement of the wild-type genes with the disrupted genes. Kanamycin-resistant clones were analysed for LOS expression. *M. bovis* strains were successfully generated from both parental strains with mutations in *lgt1* (strains 63,1Δ and 25,1Δ), *lgt6* (strain 63,6Δ and 25,6Δ), and *lgt2A1* (strain 63,2Δ and 25,2Δ). Despite multiple attempts using different conditions (such as varying DNA concentrations; broth, or plate transformations; electroporation; extended incubation time before selection; reduced kanamycin concentrations; each time with pMblgt2A2:k-lgt5 as a positive control) we were unable to generate mutants in the remaining genes. The *lgt1* and *lgt2A1* genes are convergently located, so it is unlikely that mutation of either of these genes will alter transcription of respective downstream genes (Figure 1).

OS structure of mutants:

Crude LOS preparations were initially examined using silver stained Tricine-SDS-PAGE and confirmed for each of the mutant strains that the LOS was truncated relative to wild-type, and that the equivalent mutants from each of the parental strains (Epp63 and Mb25) had the same migration profile (not shown). The LOS was subsequently isolated from broth-grown *M. bovis* Epp63 mutants and purified as previously described [4]. After removal of lipid A from the LOS by acid hydrolysis, the OS structures were investigated by NMR. ¹H and ¹³C NMR chemical shift data for the OS are shown in Table 3.

M. bovis 63,6Δ

Examination of ^1H and 2D ^1H , ^{13}C -HSQC NMR spectra of *M. bovis* 63,6 Δ OS revealed Kdo. Chemical shift assignments are provided in Table 3. The confirmed structure of *M. bovis* 63,6 Δ OS as a Kdo residue is consistent with the gene encoded by *lgt6* being responsible for the addition of the α -glucose residue **D** (see Figure 3). Most reported LPS structures have heptose as the first sugar linking OS core to the Kdo. Both *M. catarrhalis* and *M. bovis* are unusual in having a single glucose residue at this position [4, 10]. The *M. catarrhalis* *lgt6* gene has previously been demonstrated to encode an α -(1-5)-glucosyltransferase that adds the primary glucose to the Kdo₂-lipid A-OH in LOS biosynthesis [7]. The protein encoded by the *M. bovis* *lgt6* gene is 78% identical to Lgt6 of *M. catarrhalis*. Unlike in *M. bovis*, the *lgt6* *M. catarrhalis* gene is not co-located with other LOS biosynthesis genes. However, *lgt6* of both *M. bovis* and *M. catarrhalis* are located downstream of lipid A acyltransferases, involved in synthesis of lipid A (Figure 4). The encoded *M. bovis* protein upstream of *lgt6* is 63% identical to LpxX, the decanoyl acyltransferase that is located upstream of *lgt6* in *M. catarrhalis* [7, 11]. Given the syntenic and sequence similarity between *lgt6* in these closely related species, and the effect of mutation on LOS structure, *M. bovis* *lgt6* is proposed to encode an α -(1-5)-glucosyltransferase.

M. bovis 63,1 Δ

The structure of *M. bovis* 63,1 Δ is shown in Figure 3. It comprises three glucose residues linked to a Kdo residue. In accordance with this structure the 1D ^1H NMR spectrum of *M. bovis* 63,1 Δ revealed the presence of 3 anomeric protons, with coupling constants J_{H1H2} characteristic of one α -(5.13 ppm, 3.9 Hz, residue **D**) and two β -configured sugars (4.58 ppm, 7.8 Hz and 4.49 ppm, 7.8 Hz residues **G** and **H**, respectively). Signals consistent with the H3 protons of Kdo (residue **K**) were also immediately obvious from this spectrum. This labelling

scheme is consistent with that of the recently published structure of *M. bovis* WT OS [4].

Determination of the ^1H proton and ^{13}C NMR resonances for *M. bovis* 63,1Δ was achieved by combining the information from all the NMR spectra recorded for the sample (see structure in Figure 3). Further confirmation of the structure of 63,1Δ OS was obtained by comparison to the *M. bovis* Epp63 wildtype spectra [4].

Residue **D** is an α -glucose $J_{\text{H1H2}} = 3.9$ Hz and $J_{\text{CH}} = 161.4$ Hz. Assignment of the specific sugar resonances was achieved by selectively irradiating the anomeric resonance at 5.13 ppm and acquiring a selective TOCSY experiment with a 180 ms spin lock time, providing a well dispersed spectrum that was readily assigned. The H2 and H3 resonances were confirmed by inspection of a COSY spectrum. The location of the H6 correlations was easily identified from this spectrum and the C6 correlations confirmed from an edited HSQC experiment. The chemical shift location of the C6 methylene ^{13}C signals at 68.0 ppm reflected that it was glycosylated. Likewise, a pronounced glycosylation effect was observed for the C4 at 78.8 ppm. ^{13}C chemical shifts were obtained from examining a ^1H - ^{13}C HSQC-TOCSY spectrum. These assignments were in accordance with residue **D** in *M. bovis* wildtype OS.

In *M. bovis* wildtype OS, the α -glucose residues **G** and **H** are both substituted at the $\rightarrow 2)$ position, however, in *M. bovis* 63,1Δ these residues are in a terminal position. Accordingly, residues **G** and **H** chemical shifts were consistent to those expected for terminal α -glucose residues and lacked a glycosylation effect at C2 that was observed in the *M. bovis* wildtype OS spectra. Selective TOCSY experiments overlaid with a ^1H - ^{13}C HSQC-TOCSY spectra confirmed the assignments that are provided in Table 3. Linkage information was obtained from selective NOESY spectra: Irradiation of H1 of **H** at 4.49 ppm gave NOE correlations to the H6 protons of **D** and the H1 of **G** at 4.58 ppm gave NOE correlations to the H4 proton of **D** at 3.77 ppm. Resonances due to Kdo (residue **K**) are provided in Table 3 and were

identified from overlay of edited ^1H - ^{13}C HSQC and ^1H - ^{13}C HSQC-TOCSY spectra and by comparison to spectra for *M. bovis* wildtype OS. Mutation of the *lgt1* gene in *M. bovis* Epp63 (strain 63,1 Δ) resulted in OS containing Kdo and only three glucose residues. In the wild-type OS, residues **G** and **H** have a glucose added in an α -(1-2) linkage. Our prior work, and that of others, showed that *M. catarrhalis* Lgt1 (81% similar to Lgt1_{*Mbovis*}) is an α -(1-2)-glucosyltransferase, adding glucose to a β -glucose acceptor molecule [5, 8]. In addition, Lgt1 of *M. catarrhalis* catalyses addition of glucose onto more than one acceptor molecule. In *M. bovis* wild-type OS, residue **C** is linked to an α -glucose residue. We propose that *M. bovis* Lgt1 is an α -(1-2)-glucosyltransferase responsible for catalysing addition of residues **A**, **B**, and **C**. However, given the different acceptor residue for residue **C**, it is possible that an (as yet) unidentified glycosyltransferase catalyses this linkage.

M. bovis 63,2 Δ

M. bovis 63,2 Δ contains a mutation of the *lgt2A₁* gene. The structure of *M. bovis* 63,2 Δ is shown in Figure 3. It comprises six glucose residues linked to a Kdo residue and as expected the 1D ^1H spectrum of *M. bovis* 63,2 Δ revealed the presence of six anomeric protons, with coupling constants J_{H1H2} characteristic of four α - and two β -configured sugars in addition to Kdo. The additional sugars compared to *M. bovis* 63,1 Δ were all identified on the basis of coupling constants (J_{H1H2}) to be in the α -anomeric configuration. Furthermore, these residues were deemed to be glucose residues on examination of their spin systems in the ^1H - ^{13}C HSQC-TOCSY and selective TOCSY spectra. In *M. bovis* wildtype OS, the α -glucose residues **G** and **H** are both substituted at the $\rightarrow$ 2) position, however in 63,1 Δ these residues are in a terminal position. The $^1\text{H}/^{13}\text{C}$ chemical shifts for the G-2 and H-2 position in *M. bovis* 63,2 Δ OS are 3.57/78.9 and 3.57/75.9 ppm respectively, which represent significant glycosylation shifts compared to the same residues in *M. bovis* 63,1 Δ (3.26/74.1 and 3.23/74.1

ppm), confirming the 2-substitution of these residues. Residues **G** and **H** were confirmed to be β -glucose residues (**G**: $J_{\text{H1H2}} = 6.9$ Hz and $J_{\text{CH}} = 160.5$ Hz) and (**H**: $J_{\text{H1H2}} = 7.7$ Hz and $J_{\text{CH}} = 164.3$ Hz).

Linked in a terminal position to residue **G** is the α -glucose residue **B** with an anomeric resonance at 5.31 ppm ($J_{\text{H1H2}} = 3.8$ Hz and $J_{\text{CH}} = 173.1$ Hz). Chemical shift analysis of the ring protons of residue **B** confirmed its terminal location. Its connectivity to residue **G** was confirmed by the presence of a NOE from B1-G2 in a 1D-selective ^1H NOESY spectrum irradiated at the anomeric proton B1: 5.31 ppm to G2 at 3.57 ppm and also the presence of a B1-G2 correlation in the $^1\text{H}^{13}\text{C}$ HMBC spectrum.

Identification of the carbohydrate residue linked directly to **H** was provided by the presence of a NOE from A1-H2 in a 1D selective ^1H NOESY spectrum irradiated at the anomeric proton A1 (5.80 ppm) which showed a correlation at 3.57 ppm consistent with the A1-H2 NOE and also the presence of a A1-H2 correlation in the $^1\text{H}^{13}\text{C}$ HMBC spectrum. Residue **A** is an α -glucose with its anomeric resonance at 5.80 ppm ($J_{\text{H1H2}} = 3.7$ Hz and $J_{\text{CH}} = 173.7$ Hz). The chemical shift position of A2 (^1H : 3.71 ppm and ^{13}C : 75.4 ppm) was consistent with this residue being glycosylated at the $\rightarrow 2$) position. Confirmation that residue **C** was linked at the $\rightarrow 2$) position of residue **A** was provided by the presence of a C1-A2 NOE in a 1D selective ^1H NOESY spectrum irradiated at the anomeric proton C1 (5.21 ppm) at 3.71 ppm and also the presence of a strong C1-A2 correlation in the $^1\text{H}^{13}\text{C}$ HMBC spectrum. The terminal location of residue **C** was reflected in the chemical shift of its ring protons and its anomeric configuration from examination of ($J_{\text{H1H2}} = 3.8$ Hz and $J_{\text{CH}} = 169.7$ Hz).

Taken together the confirmed structure of *M. bovis* 63,2 Δ is consistent with Lgt2A being responsible for the addition of a β -glucose residue to the $\rightarrow 4$) position of residue **A** and

therefore confirming its β -glucosyltransferase function in the biosynthesis of *M. bovis* Epp63 LOS.

Having assigned function to genes that we had mutated, we next proposed functional assignments for the remaining genes in the locus on the basis of sequence similarity to related glycosyltransferases.

Lgt3/5/7/2A₂

Most reported LPS structures have a heptose residue as the first sugar linking the OS core to the Kdo. In both *M. catarrhalis* and *M. bovis* this first glucose has multiple glucose residues added. In *M. catarrhalis* this is catalysed by a single multi-functional gene product, Lgt3 [12, 13]. The *M. bovis* Lgt3 homologue is 79% identical to Lgt3_{*M. catarrhalis*}. The structural and sequence similarities leads us to assign the function of Lgt3 as indicated in Figure 4, adding glucose residues (**G** and **H**) to the central core glucose (residue **D**).

The unusual terminal sugar in the LOS of *M. bovis* is an open chain acetal-linked *N*-acetylgalactosamine, (1*S*)-GalaNAc residue (residue **F**, Figure 3) [4]. This residue is also present in *Actinobacillus pleuropneumoniae*, *Shewenella oneidensis*, and *Pasteurella multocida* [14-16], and NatA of *Pasteurella multocida* has been inferred to add this unusual sugar. The *M. bovis* *lgt7* gene product has homologues in these three species (protein sequence identity of 65%, 58%, and 57% respectively to homologues from *Acinetobacter*, *Pasteurella multocida*, and *Shewenella oneidensis*). From the sequence similarities, we propose that *M. bovis* Lgt7 encodes the glycosyltransferase responsible for addition of (1*S*)-GalaNAc, but confirmation of this awaits mutation of this gene or other biochemical evidence. The mechanism of action for such a glycosyltransferase has not been characterized, and as yet, direct evidence from other species for the activity of this class of glycosyltransferase has not yet been reported.

M. bovis Lgt5 is highly similar to a homologue in *M. catarrhalis* (with sequence 78% identity). Lgt5_{*M. catarrhalis*} is an α -(1-4)-galactosyltransferase, adding a galactose to a β -galactose residue [5, 6, 9]. This linkage and acceptor is identical to residue **E** of Epp63 OS (Figure 3) [4]. Based on the sequence similarity between the *Moraxella* Lgt5 homologues, we propose that *M. bovis* Lgt5 is also an α -(1-4)-galactosyltransferase, adding galactose to a β -galactose residue.

The remaining unassigned glycosyltransferase activity is a β -(1-4)-galactosyltransferase, adding residue **J** (Figure 3). The encoded *M. bovis* Lgt2_{A2} is 43% identical to a *M. catarrhalis* Lgt2A which is also a β -(1-4)-galactosyltransferase [5, 9].

Although it is clear that sequence similarity is not sufficient to define enzyme function, the functional assignment we propose provides a basis for experimental investigation of each of the glycosyltransferases for *M. bovis*. The gene arrangement of the LOS biosynthetic genes is similar between the LOS glycosyltransferase of *M. bovis* and *M. catarrhalis*, (Figure 4). This synteny between species within the same genus is perhaps unsurprising, given the structural and proposed functions of these genes. At this time, little is known regarding the LOS variability within a strain. We found no evidence of genetic sequences that would suggest *M. bovis* LOS biosynthesis is phase variable via tandem repeat slippage. Similarly, little is known of inter-strain LOS differences, although one study showed LOS of at least three differing electrophoretic mobilities [17]. The similarity in electrophoretic mobility of mutants between Mb25 and Epp63 in the present study suggests that at least for these genes, the associated structures are conserved in these strains. Another related bacterium, *M. bovoculi*, has relatively recently been recognised as another causative agent of IBK (see [18] for review). To date, nothing is known of the LOS structure of *M. bovoculi*, but the recent draft genome

sequence may facilitate structural predictions and genetic manipulation [19] for this bacterium.

This study provides insight into the biosynthesis of the major glycolipid of an economically important bovine pathogen, and provides tools to assess the role of the LOS cell surface glycan residues in the colonisation and pathogenesis of *M. bovis*.

Experimental

Strains and Growth: *M. bovis* strain Epp63 was obtained from Assoc. Prof. S. K. Highlander, Baylor College of Medicine, Texas, USA. *M. bovis* strain Mb25 was obtained from the collection of Dr P. Blackall, Department of Primary Industries, Queensland, Australia. The lyophilized bacteria (stored in 1957) were imported into Australia from Omaha, Nebraska from Dr J. Baldwin, Corn States Serum Company, Harney Street, Omaha, Nebraska. Other strains and plasmids are described in Supplementary Table 1. *M. bovis* was grown in liquid BHI medium supplemented with 10% of a clarified 50% (v/v) BHI/defibrinated horse blood (DHB) (Biomerieux), with 1.5% agar added for solid media where appropriate. For selection of mutants, kanamycin (15 µg/mL) was added. *Escherichia coli* were grown in LB medium with agar for solid media, and addition of kanamycin (50 g/mL) or ampicillin (50 µg/mL) when required.

Sequence analysis: All oligonucleotide primers used are listed in the Supplementary Table 2. Preliminary sequence data was obtained from Baylor College of Medicine Human Genome Sequencing Center website at <http://www.hgsc.bcm.edu>. Sanger dideoxy sequencing was performed using BigDye termination mix, with primers based on three identified contigs containing putative LOS biosynthesis genes (Figure 1, Supplementary Table 2). Sequence analysis and assembly used MacVector, and similarity searches were performed using the BLAST suite hosted at www.ncbi.nlm.nih.gov. The sequence assembly is deposited as Accession number KP795984.

Cloning and Mutagenesis: *M. bovis* *lgt2_{A1}* and *lgt1* were amplified, using Phusion polymerase with primers Mb396-32FR and Mb396-1598 and the 1.6 kbp amplicon was gel purified, sequenced and ligated into *Sma*I digested pSK to generate plasmid pMblgt2_{A1}-lgt1 (Figure 2, Table 2). This plasmid was separately digested with *Aat*II and *Sty*I for subsequent kanamycin insertion, in order to inactivate *lgt2_{A1}* and *lgt1*, respectively. *M. bovis* *lgt2_{A2}* and

lgt5 were amplified using Taq polymerase with primers 288F_34 and 288R_2737. The amplicon was gel purified, sequenced and blunt-ended with Klenow before ligating into *Sma*I digested pSK, generating pMblgt2_{A2}-lgt5. This plasmid was separately digested with *Bln*II and *Nde*I for subsequent kanamycin insertion, in order to inactivate *lgt2_{A2}* and *lgt5*, respectively. Other genes were amplified individually: *lgt6* (primers 380F_1532 and 380R_2355), *lgt3* (380FF_2 and 380R_1848), and *lgt7* (288F_1650 and 288R_2737) and cloned into pGemT-easy to generate plasmids pMblgt6, pMblgt3 and pMblgt7, respectively. These had *Hinc*II kanamycin gene inserted into conveniently located blunted *Hind*III sites to inactivate the genes.

Transformation: Approximately 50 colonies of strains (Epp63 or Mb25) were resuspended in 500 µL sterile 1 x PBS. Bacterial suspension (10 µL) was pipetted onto a Blood/BHI plate to form the inoculum, left to dry for 10-15 min and incubated at 37°C for 1-3 hr to allow the cells to grow. Plasmid inserts, excised from 1-5 µg of plasmid by restriction endonuclease digest, and purified from agarose after electrophoretic separation, were added to the bacterial inoculum in PBS and incubated for a further 1-5 hr at 37°C to allow DNA uptake. Cells were plated to selective media and resulting antibiotic resistant clones analysed for allelic replacement by mutant alleles.

LOS isolation: For small-scale crude LOS preparation, one colony of *M. bovis* was suspended in PBS to give an absorbance reading of 5 at 280 nm. This suspension was Proteinase-K treated at 37°C for 1 hr. Tricine-SDS-PAGE gel of undigested *M. bovis* versus proteinase-K digested *M. bovis* was run for 2.5 hr at 200 mA followed by Coomassie blue and silver-staining.

After overnight growth on solid media, sequential subcultures were prepared in 10 mL, then 600 mL, then 6 L of supplemented media. After growth, bacteria were killed by addition of

0.5% (v/v) phenol. After centrifugation at 5,000 *g* for 20 min, the wet cell mass (~40 g) was dehydrated by sequential washing in: ethanol (1 x 2 volumes), acetone (2 x 2 volumes), diethyl ether (2 x 2 volumes) followed by centrifugation at 5,000 *g* for 10 min. The cell mass was dried under house vacuum overnight.

The dried cell mass was solubilized in Milli-Q water and enzymatically digested to remove RNA, DNA, and protein (RNase and DNase, 40°C, 3 hr with gentle shaking; Protease Type XIV, 40°C, 16 hr with gentle shaking). Enzymes were inactivated by heat treatment at 90°C for 20 min. The sample was dialyzed against Milli-Q water and lyophilized. LOS was isolated from the dry material using the hot phenol/water extraction method. To *O*-deacylate the LOS (~180 mg, aqueous-phase) was hydrolysed in 1% acetic acid (100°C, 2 hr) with lipid A precipitate separated by centrifugation (8,000 *g*, 30 min). The aqueous supernatant was separated by size exclusion chromatography on a column (80 x 20 mm) of P2-BioGel (Biorad) with Milli-Q water as eluent to give purified OS.

4.4 NMR Spectroscopy

OS (1 mg) dissolved in D₂O (CIL 99.998%) was lyophilized in triplicate to remove exchangeable protons. The sample was dissolved in 500 µL of D₂O under nitrogen and ¹H and ¹³C spectra were recorded at 298 K on a Bruker Avance spectrometer with a TXI cryoprobe operating at 600 MHz and 150 MHz, respectively. Chemical shifts are reported in ppm referenced to internal acetone (¹H 2.225 ppm, ¹³C 31.45 ppm). Spectral assignment was aided by the recording of gradient COSY, DQF-COSY, gradient TOCSY (60 and 120 ms mixing time), ¹³C attached proton test (APT), ¹H¹³C-HSQC, ¹H¹³C-HSQC with no decoupling during acquisition to determine ¹³C-¹H coupling constants, ¹H¹³C-HSQC-TOCSY (15, 45, 60 and 120 ms mixing times), ¹H¹³C-HMBC and 1D selective COSY, TOCSY (with mixing times ranging from 15 ms to 180 ms) and NOESY spectra (250 ms). All spectra were

acquired using unmodified pulse sequences from the Bruker pulse sequence library. The most informative are discussed.

Acknowledgements

M. bovis (Epp63) and (Mb25) were obtained from Assoc. Prof. S. K. Highlander, Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston, Texas, USA and Dr. P. Blackall, Department of Primary Industries, Queensland, Australia, respectively.

The Institute for Glycomics, Menzies Health Institute Queensland - Griffith University and the School of Medical Science Griffith University are thanked for financial support. IF was supported by a Griffith University Postgraduate Research Scholarship (GUPRS). This project was undertaken as part of the Queensland NMR Network (QNN).

The DNA sequence of *Moraxella bovis* was supported by USDA/NRICGP 2004-04989 to S. K. Highlander and G. M. Weinstock at the Baylor College of Medicine, and BCM-HGSC.

Figure legends:

Figure 1: *M. bovis* glycosyltransferase gene locus: ordering of the three contigs by PCR (not to scale). Contigs **380, 396, 288** were identified in preliminary genome sequence of *M. bovis* Epp63 (Highlander et al, Baylor College of Medicine). Primers were used to amplify and sequence, allowing assembly into a single contig: Primers used: 1: Mb380R110; 2: Mb396F1543; 3: Mb396R180; 5: Mb288R105.

Figure 2. Overview of the *M. bovis* mutagenesis constructs. Grey arrows symbolize the kanamycin cassette; Restriction enzyme used for Kan cassette insertion shown in parentheses.

Figure 3. Structure of OS from wildtype Epp63 (top panel), 63,1Δ (middle panel), and 63,2Δ (bottom panel).

Figure 4: Genetic basis for LOS biosynthesis in *M. bovis* and structural and gene syntenic similarity with *M. catarrhalis* **4A**. Assignment of *M. bovis* Epp63 LOS biosynthesis gene function, compared with LOS structure of *M. catarrhalis* serotype B. Functions for *lgt1,3, 6* are derived from mutation studies, remaining functions inferred from sequence similarity. **4B**. Comparison of the gene sequence between LOS biosynthesis loci of *M. bovis* Epp63 and serotype B strain *M. catarrhalis*. Lines between genes indicate homologues, with % protein sequence identity indicated. Bold double line indicates *lpxX* and *lgt6* are at a different chromosomal location than the other glycosyltransferases in *M. catarrhalis*.

References

1. O'Connor, A.M., S. Brace, S. Gould, R. Dewell, and T. Engelken, *A randomized clinical trial evaluating a farm-of-origin autogenous Moraxella bovis vaccine to control infectious bovine keratoconjunctivitis (pinkeye) in beef cattle*. J Vet Intern Med, 2011. **25**(6): p. 1447-53.
2. Angelos, J.A., K.G. Gohary, L.M. Ball, and J.F. Hess, *Randomized controlled field trial to assess efficacy of a Moraxella bovis pilin-cytotoxin-Moraxella bovoculi cytotoxin subunit vaccine to prevent naturally occurring infectious bovine keratoconjunctivitis*. Am J Vet Res, 2012. **73**(10): p. 1670-5.
3. Funk, L.D., J.M. Reecy, C. Wang, R.G. Tait, Jr., and A.M. O'Connor, *Associations between infectious bovine keratoconjunctivitis at weaning and ultrasonographically measured body composition traits in yearling cattle*. J Am Vet Med Assoc, 2014. **244**(1): p. 100-6.
4. De Castro, C., I.D. Grice, T.M. Daal, I.R. Peak, A. Molinaro, and J.C. Wilson, *Elucidation of the structure of the oligosaccharide from wild type Moraxella bovis Epp63 lipooligosaccharide*. Carbohydr Res, 2014. **388**: p. 81-6.
5. Peak, I.R., I.D. Grice, I. Faglin, Z. Klipic, P.M. Collins, L. van Schendel, P.G. Hitchen, H.R. Morris, A. Dell, and J.C. Wilson, *Towards understanding the functional role of the glycosyltransferases involved in the biosynthesis of Moraxella catarrhalis lipooligosaccharide*. FEBS J, 2007. **274**(8): p. 2024-37.
6. Wilson, J.C., P.M. Collins, Z. Klipic, I.D. Grice, and I.R. Peak, *Identification of a novel glycosyltransferase involved in LOS biosynthesis of Moraxella catarrhalis*. Carbohydr Res, 2006. **341**(15): p. 2600-6.
7. Schwingel, J.M., F. St Michael, A.D. Cox, H. Masoud, J.C. Richards, and A.A. Campagnari, *A unique glycosyltransferase involved in the initial assembly of Moraxella catarrhalis lipooligosaccharides*. Glycobiology, 2008. **18**(6): p. 447-55.
8. Edwards, K.J., S. Allen, B.W. Gibson, and A.A. Campagnari, *Characterization of a cluster of three glycosyltransferase enzymes essential for Moraxella catarrhalis lipooligosaccharide assembly*. J Bacteriol, 2005. **187**(9): p. 2939-47.
9. Edwards, K.J., J.M. Schwingel, A.K. Datta, and A.A. Campagnari, *Multiplex PCR assay that identifies the major lipooligosaccharide serotype expressed by Moraxella catarrhalis clinical isolates*. J Clin Microbiol, 2005. **43**(12): p. 6139-43.
10. Edebrink, P., P.E. Jansson, M.M. Rahman, G. Widmalm, T. Holme, M. Rahman, and A. Weintraub, *Structural studies of the O-polysaccharide from the lipopolysaccharide of Moraxella (Branhamella) catarrhalis serotype A (strain ATCC 25238)*. Carbohydr Res, 1994. **257**(2): p. 269-84.
11. Gao, S., D. Peng, W. Zhang, A. Muszynski, R.W. Carlson, and X.X. Gu, *Identification of two late acyltransferase genes responsible for lipid A biosynthesis in Moraxella catarrhalis*. FEBS J, 2008. **275**(20): p. 5201-14.
12. Faglin, I., J. Tiralongo, J.C. Wilson, P.M. Collins, and I.R. Peak, *Biochemical analysis of Lgt3, a glycosyltransferase of the bacterium Moraxella catarrhalis*. Biochem Biophys Res Commun, 2010. **393**(4): p. 609-13.
13. Luke-Marshall, N.R., K.J. Edwards, S. Sauberan, F. St Michael, E.V. Vinogradov, A.D. Cox, and A.A. Campagnari, *Characterization of a trifunctional glucosyltransferase essential for Moraxella catarrhalis lipooligosaccharide assembly*. Glycobiology, 2013. **23**(8): p. 1013-21.

14. Michael, F.S., J.R. Brisson, S. Larocque, M. Monteiro, J. Li, M. Jacques, M.B. Perry, and A.D. Cox, *Structural analysis of the lipopolysaccharide derived core oligosaccharides of Actinobacillus pleuropneumoniae serotypes 1, 2, 5a and the genome strain 5b*. Carbohydr Res, 2004. **339**(11): p. 1973-84.
15. Vinogradov, E., A. Korenevsky, and T.J. Beveridge, *The structure of the rough-type lipopolysaccharide from Shewanella oneidensis MR-1, containing 8-amino-8-deoxy-Kdo and an open-chain form of 2-acetamido-2-deoxy-D-galactose*. Carbohydr Res, 2003. **338**(19): p. 1991-7.
16. Harper, M., F. St Michael, E. Vinogradov, M. John, J.A. Steen, L. van Dorsten, J.D. Boyce, B. Adler, and A.D. Cox, *Structure and biosynthetic locus of the lipopolysaccharide outer core produced by Pasteurella multocida serovars 8 and 13 and the identification of a novel phospho-glycero moiety*. Glycobiology, 2013. **23**(3): p. 286-94.
17. Prieto, C.I., O.M. Aguilar, and O.M. Yantorno, *Analyses of lipopolysaccharides, outer membrane proteins and DNA fingerprints reveal intraspecies diversity in Moraxella bovis isolated in Argentina*. Vet Microbiol, 1999. **70**(3-4): p. 213-23.
18. Angelos, J.A., *Infectious Bovine Keratoconjunctivitis (Pinkeye)*. Vet Clin North Am Food Anim Pract, 2015.
19. Calcutt, M.J., M.F. Foecking, N.T. Martin, T. Mhlanga-Mutangadura, and T.J. Reilly, *Draft Genome Sequence of Moraxella bovoculi Strain 237T (ATCC BAA-1259T) Isolated from a Calf with Infectious Bovine Keratoconjunctivitis*. Genome Announc, 2014. **2**(3).

Table 1: Genes in the *M. bovis* LOS biosynthesis locus

<i>M. bovis</i> gene	Protein homologue (accession number)	Sequence identity (%) of encoded protein	Function of homologue	Function of <i>M. bovis</i> glycosyltransferase. E=experimental evidence from mutagenesis P= proposed function from similarity/syntenly
<i>lpxX</i>	LpxX (ABY64521.1)	63%	<i>M. catarrhalis</i> , decanoyl lipid A acyltransferase	lipid A acyltransferase P
<i>lgt2_{A1}</i>	Lgt2 _A (AAZ29047)	43%	<i>M. catarrhalis</i> , β -(1→4)-galactosyltransferase	
<i>lgt5</i>	Lgt5 (ABM92438)	78%	<i>M. catarrhalis</i> , α -(1→4)-galactosyltransferase	α -(1→4)-galactosyltransferase P
<i>lgt7</i>	Putative glycosyltransferase (ZP06073353)	66%	<i>Acinetobacter radioresistens</i> , putative glycosyltransferase	(1S)-GalaNAc transferase P
<i>lgt6</i>	Lgt6 (ABX82336)	78%	<i>M. catarrhalis</i> , α -(1→5)-glucosyltransferase	α -(1→5)-glucosyltransferase E
<i>lgt3</i>	Lgt3 (ABB16338)	79%	<i>M. catarrhalis</i> , glucosyltransferase	Dual functional glucosyltransferase E
<i>lgt2_{A2}</i>	Lgt2 _A (ABA02572)	41%	<i>M. catarrhalis</i> , β -(1→4)-galactosyltransferase	
<i>lgt1</i>	Lgt1 (AAZ29050)	81%	<i>M. catarrhalis</i> , α -(1→2)-glucosyltransferase	α -(1→2)-glucosyltransferase E

Table 2: Percent Identity Matrix of Lgt2 homologues.

	McLgt2 _A	McLgt2 _{BC}	MbLgt2 _{A1}	Mblgt2 _{A2}
1: McLgt2 _A	100			
2: McLgt2 _{BC}	63	100		
3: MbLgt2 _{A1}	43	39	100	
4: MbLgt2 _{A2}	41	38	43	100

Table 3. (600 MHz, D₂O, 298 K) Proton (plain) and carbon (italic) chemical shift data (ppm)

for OS: A) 63,6Δ; B) 63,1Δ; and C) 63,2Δ. $^3J_{\text{H1H2}}$ (Hz) are given in brackets and spectra are calibrated on internal acetone, $^1\text{H} = 2.225$ ppm, $^{13}\text{C} = 31.45$ ppm).

A) *M. bovis* 63,6Δ

	3	4	5	6	7	8
K	2.21; 1.81	4.01	3.98	3.51	3.76	3.84; 3.61
α-Kdo	<i>40.3</i>	<i>68.2</i>	<i>71.2</i>	<i>69.6</i>	<i>71.4</i>	<i>64.2</i>

B) *M. bovis* 63,1Δ

	1	2	3	4	5	6
D	5.13 (3.9)	3.58	3.89	3.77	4.36	3.98, 4.13
α-4,6-Glc	<i>100.4</i>	<i>72.7</i>	<i>72.6</i>	<i>78.8</i>	<i>70.4</i>	<i>68.0</i>
G	4.58 (7.8)	3.26	3.50	3.39	3.45	3.93; 3.74
β-Glc-t	<i>103.2</i>	<i>74.1</i>	<i>76.6</i>	<i>70.5</i>	<i>76.8</i>	<i>61.7</i>
H	4.49 (7.8)	3.23	3.49	3.39	3.44	3.92; 3.72
β-Glc-t	<i>103.2</i>	<i>74.1</i>	<i>76.8</i>	<i>70.6</i>	<i>76.8</i>	<i>61.7</i>
	3	4	5	6	7	8
K	2.19; 1.81	4.13	4.06	3.81	4.01	3.82; 3.59
α-5-Kdo	<i>40.2</i>	<i>67.1</i>	<i>76.5</i>	<i>72.6</i>	<i>68.5</i>	<i>63.5</i>

C) *M. bovis* 63,2Δ

	1	2	3	4	5	6
A	5.80 (3.7)	3.71	3.85	3.50	4.03	3.77 x 2
α-2-Glc	<i>94.4</i>	<i>75.4</i>	<i>70.8</i>	<i>72.9</i>	<i>72.4</i>	<i>61.9</i>
B	5.31 (3.8)	3.51	3.72	3.44	4.03	3.91, 3.77
α-t-Glc	<i>99.6</i>	<i>73.0</i>	<i>74.5</i>	<i>70.8</i>	<i>73.0</i>	<i>61.6</i>
C	5.21 (3.8)	3.55	3.78	3.43	3.94	3.91, 3.86
α-t-Glc	<i>96.3</i>	<i>73.0</i>	<i>74.5</i>	<i>70.8</i>	<i>73.0</i>	<i>62.3</i>
D	5.14 (3.9)	3.57	4.10	3.78	4.56	4.10
α-4,6-Glc	<i>100.9</i>	<i>73.7</i>	<i>73.7</i>	<i>78.0</i>	<i>70.6</i>	<i>68.8</i>
G**	4.76 (6.9)**	3.57	3.57	3.43	3.41	3.84, 3.71
β-2-Glc	<i>100.6</i>	<i>78.9</i>	<i>76.2</i>	<i>71.0</i>	<i>77.1</i>	<i>62.3</i>
H	4.64 (7.7)	3.57	3.42	3.67	3.42	3.90, 3.72
β-2-Glc	<i>103.5</i>	<i>75.9</i>	<i>76.9</i>	<i>70.9</i>	<i>76.9</i>	<i>62.3</i>
	3	4	5	6	7	8
K	2.02; 1.85	4.13	4.06	3.81	4.01	3.78, 3.61
α-5-Kdo	<i>40.6</i>	<i>67.4</i>	<i>76.5</i>	<i>72.6</i>	<i>69.7</i>	<i>64.3</i>

Supplementary Table 1. Strains and plasmids used in this study.

Strain	Genotype, relevant phenotype, or selection marker	Source/ Reference
<i>Moraxella bovis</i>		
Epp63	Wild type	Dr S.K. Highlander
Mb25	Wild type	Dr P. Blackall
63,1Δ	Epp63, <i>lgt1</i> mutated	This study
63,2Δ	Epp63, <i>lgt2_{A1}</i> mutated	This study
63,6Δ	Epp63, <i>lgt6</i> mutated	This study
25,1Δ	Mb25, <i>lgt1</i> mutated	This study
25,2Δ	Mb25, <i>lgt2_{A1}</i> mutated	This study
25,6Δ	Mb25, <i>lgt6</i> mutated	This study
Vector/construct	Genotype, relevant phenotype, or selection marker	Source/ Reference
pBluescriptSK	F1(+) ColE1 ori <i>lacZα</i> Amp ^r	Stratagene
pGemT-Easy	F1 ori <i>lacZ</i> Amp ^r	Promega
pUC4kan	PBR322 ori <i>lacZ</i> Amp ^r Kan ^r	Pharmacia
pMblgt _{A1} -lgt1	<i>lgt2_{A1}</i> + <i>lgt1</i> in pSK	This study-Figure 2
pMblgt _{A1} :k-lgt1	<i>lgt2_{A1}::kan</i> + <i>lgt1</i> in pSK	This study-Figure 2
pMblgt _{A1} -lgt1:k	<i>lgt2_{A1}</i> + <i>lgt1::kan</i> in pSK	This study-Figure 2
pMblgt _{A2} -lgt5	<i>lgt2_{A2}</i> + <i>lgt5</i> in pSK	This study-Figure 2
pMblgt _{A2} -lgt5:k	<i>lgt2_{A2}</i> + <i>lgt5::kan</i> in pSK	This study-Figure 2
pMblgt _{A2} :k-lgt5	<i>lgt2_{A2}::kan</i> + <i>lgt5</i> in pSK	This study-Figure 2
pMblgt6	<i>lgt6</i> in pGEMTEasy	This study-Figure 2
pMblgt6:k	<i>lgt6::kan</i> in pGEMTEasy	This study-Figure 2
pMblgt7	<i>lgt7</i> in pGEMTEasy	This study-Figure 2
pMblgt7:k	<i>lgt7::kan</i> in pGEMTEasy	This study-Figure 2
pMblgt3	<i>lgt3</i> in pGEMTEasy	This study-Figure 2
pMblgt3:k	<i>lgt3::kan</i> in pGEMTEasy	This study-Figure 2

Supplementary Table 2. Oligonucleotide primers used in this study.

Oligonucleotide name	Sequence (5'-3')	Location
380F_62	AAAGGTATAGCTTGACGAAGC	US of <i>lgt3</i>
380R_1848	TTGTCCAATTATCTGCGTGAAC	DS of <i>lgt3</i>
380F_1532	TAAGGCAGGTAAAGAGATTGC	US of <i>lgt6</i>
380R_2355	AAATCCATGACCGACAACCGC	DS of <i>lgt6</i>
Psy U2520	CGTGTATCAAGGGCAAGGATG	US of <i>lgt6</i>
Psy D3513	TGAATAACTCATCGCCAATACTCC	DS of <i>lgt6</i>
288F_34	TTGATGCAGTTACACCCAACG	US of <i>lgt2_{A2}</i>
288R_1669	GTGCGAGTTACTTTTCCCAG	DS of <i>lgt5</i>
288F-1650	TGGGAAAAAGTAACTCGCACC	US of <i>lgt7</i>
288R_2737	TATTGTGTTGCCTGACATAGC	DS of <i>lgt7</i>
Mb396F1543	CCAAATCACACGCTCCGTACC	Ends of contig 396, pointing outwards
Mb396R180	GTCTTAGGGTCAGATAAATGCACG	
Mb380R110	GTTATTCGTGGCGGTATGTTAGGC	<i>lgt3</i> end of 380 contig, outwards
Mb288F2637	AGGTCTGCAAGACTGATTGTGAC	Ends of contig 288, pointing outwards
Mb288R105	CAAGCCAATTCACCCTTGGTAAG	
Mb396-32FR	GGCAAAAGGCGATTGATGATGAT	US of <i>lgt2_{A1}</i>
Mb396-1598	TAACACGTCTTGCCATTCCGC	US of <i>lgt1</i>
Mblgt2R	ATCAGGCGTGTTACGAAAAAGTCA	DS of <i>lgt2_{A1}</i>
Mblgt1F	TGCCCTTTTAATCATCATAACAGC	DS of <i>lgt1</i>

Figure 1

Figure 2

M. bovis Epp63 OS

Figure 3 Structure of OS from parental Epp63 (top panel), 63,1Δ (middle panel), and 63,2Δ.

Figure 4A *M. bovis* Epp63

M. catarrhalis serotype B

Figure 4 B *M. bovis* Epp63

M. catarrhalis serotype B