The *Mycoplasma gallisepticum* Virulence Factor Lipoprotein MslA Is a Novel Polynucleotide Binding Protein

Yumiko Masukagami, Kelly A. Tivendale, Karim Mardani, Idan Ben-Barak, Philip F. Markham, Glenn F. Browning

Asia-Pacific Centre for Animal Health, Faculty of Veterinary Science, The University of Melbourne, Parkville, Victoria, Australia

Although lipoproteins of mycoplasmas are thought to play a crucial role in interactions with their hosts, very few have had their biochemical function defined. The gene encoding the lipoprotein MslA in *Mycoplasma gallisepticum* has recently been shown to be required for virulence, but the biochemical function of this gene is not known. Although this gene has no significant sequence similarity to any gene of known function, it is located within an operon in *M. gallisepticum* that contains a homolog of a gene previously shown to be a nonspecific exonuclease. We mutagenized both genes to facilitate expression in *Escherichia coli* and then examined the functions of the recombinant proteins. The capacity of MslA to bind polynucleotides was examined, and we found that the protein bound single- and double-stranded DNA, as well as single-stranded RNA, with a predicted binding site of greater than 1 nucleotide but less than or equal to 5 nucleotides in length. Recombinant MslA cleaved into two fragments *in vitro*, both of which were able to bind oligonucleotides. These findings suggest that the role of MslA may be to act in concert with the lipoprotein nuclease to generate nucleotides for transport into the mycoplasma cell, as the remaining genes in the operon are predicted to encode an ABC transporter.

Mycoplasmas are among the simplest self-replicating organisms, with genomes ranging in length from 500 to 1,500 kbp. However, their small genome belies the complexity of their relationship with their hosts and their significance as pathogens. Indeed, mycoplasmas are among the most economically significant bacterial pathogens of cattle, goats, pigs, and poultry. They cause chronic diseases in a wide range of animals and plants and can persist in their hosts for very long periods, even though they appear to be predominantly extracellular parasites (1).

Mycoplasmas lack much of the biosynthetic capacity of their more complex phylogenetic neighbors, the high-G+C-content Gram-positive bacteria, and thus must scavenge most of their nutritional requirements from their host (2, 3). They are highly adapted to their hosts and contain relatively few pseudogenes (4), unlike pathogens that have undergone relatively recent reductive evolution and still contain genomic remnants of genes that have ceased to play a functional role, as seen, for example, in the rickettsias (5). They have lost the capacity to produce a rigid cell wall and thus are bounded by a single cytoplasmic membrane, in which lipoproteins are overrepresented compared to other bacteria (1). The lipoproteins are hypothesized to play a crucial role in interactions with the host because they are tethered to the cytoplasmic membrane and, in most cases, exposed to the extracellular milieu. However, the biochemical function has not been defined for most mycoplasma lipoproteins. A number of lipoprotein genes are located immediately 5' to an operon encoding a putative ATP-binding cassette (ABC) transporter, leading to the suggestion that the lipoproteins encoded by these genes may act as the substrate binding proteins, increasing the efficiency of import of specific nutrients into the mycoplasma cell. However, this has been established experimentally only for the OppA lipoprotein (6).

In previous studies in our laboratory, we have focused on defining the biochemical role of genes of unknown function in mycoplasmas, focusing on those predicted to encode lipoproteins (7, 8). One of these lipoprotein genes, Mhp379, in the porcine pathogen *Mycoplasma hyopneumoniae* was found to have some structural identity with members of a family of nucleases. After site-

directed mutagenesis to convert the mycoplasma TGA tryptophan codons to TGG, the gene was expressed in *Escherichia coli* and the purified recombinant protein was examined for nuclease activity. We established that the protein is an exonuclease that can digest RNA and DNA (8). Mhp379 is part of a five-gene operon that can be found in a number of mycoplasma species (8). The three genes following Mhp379 in the operon (Mhp380, Mhp381, and Mhp382) have sequence motifs typical of ABC transporters, while the gene preceding it in the operon (Mhp378) is predicted to be a lipoprotein (9). Although it has been suggested that the ortholog of Mhp378 in the avian pathogen *M. gallisepticum* (MGA0674) is not likely to be a member of an operon (10), the same gene order seen in *M. hyopneumoniae* is also found in *M. gallisepticum* and a number of other pathogenic mycoplasmas (8). In the human pathogen *M. pneumoniae*, while the Mhp379, Mhp380, Mhp381, and Mhp382 homologs (Mpn133, Mpn134, Mpn135, and Mpn136, respectively) are adjacent, there are multiple paralogs of Mhp378, forming a gene family with 12 members (3). The two homologs of Mhp378 found in *M. genitalium* (MG185, MG260) have been included in the set of essential genes for this organism (11), but the homolog in *M. pulmonis* (MYP0240) was not found to be essential in the most recent global transposon mutagenesis study in this species (12).

Recent transposon mutagenesis studies in *M. gallisepticum* have conclusively demonstrated that the ortholog of Mhp378 in this pathogen (MGA0674) is a virulence gene, with mutants car-

Received 25 March 2013 Returned for modification 21 April 2013

Accepted 8 June 2013

Published ahead of print 24 June 2013

Editor: S. M. Payne

Address correspondence to Glenn F. Browning, glenfb@unimelb.edu.au.

Y.M. and K.A.T. should be regarded as joint first authors.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.

doi:10.1128/IAI.00365-13

TABLE 1 Oligonucleotides used for PCR

Oligonucleotide	Sequence (5'–3') ^a
001For1BamMGA0674	AGTGgATCtTTATTgGTTGCAAGTTG
001 amended	AGTGgATCcTTATTgGTTGCAAGTaGTACA
002For2/3MGA0674	CAAAAGTTATGgGCTGAATGgAAAATGAAAG
003For4 MGA0674	TAAAAAATCTTgGgTAATTACACATCAG
004For5 MGA0674	TAACtACTCTCAAACtGgGGTAAATACC
005For6 MGA0674	GTAGCCGTAAACGTTGgGTAATCTATAAG
006For7 MGA0674	CGCTTTATTAACtGgTTAATTAAGGAA
007Rev2/3 MGA0674	CTTTCATTTTcCATTcAGCcCATACTTTTG
008Rev4 MGA0674	CTGATGTGTAAATTACCcCAAAGTTTTTA
009Rev5 MGA0674	GGTATTTACCcCAGTTTGAAGAGTAGTTA
010Rev6 MGA0674	CTTATAGATTAcCAACGTTTACGGCTAC
011Rev7 MGA0674	TTCCTTTAATTAAcCAGTTAATAAAGCG
012RevSalMGA0674	TTATGTCGACTTAGTTTGTGTTAACGATCGTAGG

^a Lowercase letters indicate differences in nucleotide sequences from published sequences. These changes were introduced to create a restriction endonuclease cleavage site, to mutate a TGA tryptophan codon to TGG, or to mutate the TGT cysteine codon to TGG.

rying transposon insertions in this gene having a reduced capacity to colonize the trachea of chickens and being less able to induce tracheal pathology (10). The lipoprotein encoded by the virulence gene *mslA* produces two cleavage products (P22 and P57), both of which are immunogenic (10), but the biochemical function of the MslA protein or its cleavage products has not yet been determined.

The aim of the studies described here was to confirm the function of MGA0676 as an exonuclease and to define the biochemical function of MslA. In view of the positional association of *mslA* with a presumed exonuclease and two genes predicted to encode an ABC transporter, we hypothesized that its role was likely to be in capture of either the substrates or the products of the exonuclease.

MATERIALS AND METHODS

Bacterial strains and media. Plasmid-transformed *E. coli* JM109 cells were grown at 37°C in Luria-Bertani (LB) broth or on agar containing 50 µg of ampicillin ml⁻¹.

DNA cloning, sequencing, and expression of the MGA0674 and MGA0676 genes. The DNA sequences of the MGA0674 (*mslA*) and MGA0676 genes from *M. gallisepticum* strain R_{low} (PubMed identification number 20123709) were used to design gene sequences with a BamHI site directly preceding the codon following the cysteine of the predicted lipoprotein acylation signal sequence (LiopP) to the termination codon. The primers used are provided in Table 1. The mycoplasma TGA tryptophan codons were altered to TGG to enable the expression of the full-length gene in *E. coli*. The mature protein-coding sequences were followed by two additional termination codons and a SalI site at the 3' end. The gene sequences were synthesized and cloned separately into the plasmid pUC57, and each sequence was supplied as 4 µg of lyophilized DNA (GenScript). DNA was reconstituted to 200 ng µl⁻¹ and digested with SalI and BamHI according to the manufacturer's instructions. Digested DNA was separated by agarose gel electrophoresis, and the 2,180-bp (MGA0674) and 779-bp (MGA0676) products were excised and gel purified using a Qiaex II kit according to the manufacturer's instructions (Qiagen). These products were ligated into the expression vector pGEX-4T-1 (Amersham Pharmacia Biotech) according to the manufacturer's instructions. *E. coli* JM109 cells were transformed with the ligation mixture, and clones containing the synthesized MGA06074 and MGA0676 genes were selected. The DNA sequences of the synthesized recombinant MGA0674 and MGA0676 gene constructs in *E. coli* were confirmed using a BigDye Terminator cycle sequencing reaction kit (Applied Biosystems) and the oligonucleotides pGEXfwd, pGEXrev, and 0674fwd.

The expression of recombinant glutathione S-transferase (GST)–

MslA and GST-MGA0676 fusion proteins in *E. coli* was induced by the addition of isopropyl-beta-D-thiogalactopyranoside to a final concentration of 2 mM. The *E. coli* cells were lysed by incubation in 2 mg lysozyme ml⁻¹ for 2 h in lysis buffer (1× phosphate-buffered saline, 1 µM EDTA, 1 µM phenylmethylsulfonyl fluoride) and then sonication, followed by incubation for 30 min in 1% Triton X-100. The recombinant proteins were purified from the soluble total protein fraction by affinity chromatography with a glutathione-Sepharose column (Amersham Pharmacia Biotech), and the elution buffer was replaced with Tris-HCl, pH 8.0, using a centrifugal filter column (Millipore). The purified protein products were analyzed by sodium dodecyl sulfate–12.5% polyacrylamide gel electrophoresis (SDS-PAGE) (see Fig. 1). GST, used as a negative control, was expressed from the pGEX-4T-1 vector in *E. coli* JM109 cells and purified in the same manner as the recombinant proteins.

Assays for nuclease activity. The nuclease activity of recombinant GST-MGA0676 was analyzed by both SDS-PAGE (13) and agarose gel electrophoresis (8) as described previously. Briefly, for SDS-PAGE analysis, 1 µg of recombinant protein was electrophoresed at a constant current of 12 mA at 4°C for 1.5 h through a 10% SDS-polyacrylamide gel containing 10 µg herring sperm DNA ml⁻¹ (Roche). The gel was washed 4 times and left overnight in incubation buffer (0.04 M Tris, pH 7.5, 1% skim milk, 0.04% beta-mercaptoethanol) before the protein was allowed to renature in renaturation buffer (0.04 M Tris, pH 7.5, 1% skim milk, 0.04% beta-mercaptoethanol, 2 mM CaCl₂, 2 mM MgCl₂) statically at 37°C for 8 h. DNA was visualized by staining with ethidium bromide.

For agarose gel electrophoresis analysis, approximately 4 µg of recombinant GST-MGA0676 was incubated at 37°C in 100 µl nuclease reaction buffer with magnesium chloride (25 mM Tris-HCl, pH 8.8, 10 mM CaCl₂, 10 mM MgCl₂) containing 1 µg of nucleic acid substrate. A 10-µl aliquot was removed at different time intervals, and EDTA was added to these aliquots to a final concentration of 20 mM to stop the reaction at these time points. Reaction products were analyzed by 1% agarose gel electrophoresis, and the DNA was visualized by staining with ethidium bromide. Exonuclease and endonuclease activities were analyzed with double-stranded lambda phage DNA (New England BioLabs) and closed circular plasmid DNA (pGEX-4T-1 containing the synthesized MslA gene purified as described above), respectively, as the substrates.

Nucleic acid binding studies. Nucleic acid binding studies were performed essentially as described previously (14, 15). Single-stranded oligonucleotides were synthesized with a 5' fluorescein label (Invitrogen). Purified recombinant protein or the GST control protein and the labeled oligonucleotides were incubated for 10 min at 25°C in binding buffer (20 mM Tris-HCl, pH 7.5, 40 mM NaCl, 1 mM EDTA), and the products were separated by electrophoresis at 60 V in 2% agarose gels for 1 h or in 15% polyacrylamide gels for 3 h at a constant temperature of 4°C. Both agarose and polyacrylamide gels were prerun before samples were loaded into the gel. The double-stranded DNA (dsDNA) substrates were generated by annealing 5'-end fluorescein-labeled (dT)₃₀ and unlabeled (dA)₃₀ or 5'-end fluorescein-labeled (dT)₁₅ and (dA)₁₅ in 10 mM Tris-HCl, pH 7.5, 50 mM NaCl, and 1 mM EDTA at 95°C for 2 min and then cooling the solution to 25°C over a period of 45 min. Oligonucleotide-protein complexes were visualized by UV transillumination.

Binding of exonuclease-digested oligonucleotides. Binding studies were performed as described above using 5'-end fluorescein-labeled oligonucleotides that had been digested with exonuclease I (New England BioLabs) according to the manufacturer's instructions.

Determination of optimal conditions for binding reactions. The optimal conditions for single-stranded (ssDNA)- and dsDNA-MslA protein binding reactions were determined using gel mobility shift assays. Approximately 40 pmol of recombinant GST-MslA was added to 20 µl of binding buffer containing 20 pmol of (dT)₃₀ or 20 pmol of annealed dsDNA composed of (dT)₃₀ and (dA)₃₀, and the mixture was incubated at 25°C for 10 min. The reaction products were then analyzed by agarose gel electrophoresis. This binding assay was also used to examine the effects of pH, incubation time, ionic strength (sodium and potassium), the concen-

FIG 1 SDS-PAGE of purified recombinant proteins. Lane 1, PageRuler prestained protein ladder (Pierce); lane 2, GST-MsIA fusion protein; lane 3, GST. Numbers to the left of the gel are molecular masses (in kilodaltons).

tration of divalent cations (magnesium and calcium), and temperature on binding.

Competition binding assays to determine substrate specificity and binding efficiency. To determine the ability of GST-MsIA to bind peptides and single nucleotides, binding studies were performed as described above with the addition of one to five molar equivalents of the unlabeled tripeptide glutathione (Boehringer Mannheim) or one of the four unlabeled deoxynucleoside triphosphates (dNTPs; Bioline). To determine the binding efficiency of each molecule of MsIA, binding studies were performed using multiple combinations of 5'-end fluorescein-labeled oligonucleotides of different lengths. Reaction products were analyzed by polyacrylamide gel electrophoresis, and the ratio of the fluorescence intensity of the unbound oligonucleotide for each of the reactions was compared to the ratio in a control reaction that did not include any competing substrate.

Exonuclease I protection assay. The exonuclease I (New England Biolabs) protection assays were performed in exonuclease reaction buffer (67 mM glycine-KOH, 6.7 mM MgCl₂, 10 mM beta-mercaptoethanol, pH 9.5). Approximately 0.04 U of exonuclease I was added to 20 μ l of buffer to which (dT)₃₀ and recombinant GST-MsIA had already been added. The reaction mixture was incubated at 37°C for 30 min and then for 20 min at 80°C to heat inactivate the exonuclease I. The reaction products were analyzed by 2% agarose gel electrophoresis (16–18).

Examination of cleavage products of MsIA. Cleavage products of GST-MsIA were identified by PAGE and immunoblotting. The products derived from the amino- and carboxyl-terminal ends of the protein were differentiated by probing Western blots with polyclonal antibodies against GST. The site of cleavage within MsIA was predicted by using mass spectrometry to determine the size of the carboxyl-terminal cleavage product of GST-MsIA.

RESULTS

Cloning and expression of the MsIA and MGA0676 genes. Six and four tryptophan codons were identified in the sequences of the MGA0674 and MGA0676 genes, respectively, of *M. gallisepticum* strain R_{low}. All tryptophan codons were altered to TGG to enable the full-length expression of the recombinant proteins. The genes encoding the mature forms of the proteins were cloned into the expression vector pGEX-4T-1, and these constructs were used to transform *E. coli* strain JM109. The recombinant GST-MsIA fusion protein, purified by glutathione affinity chromatography, had a molecular mass of approximately 100 kDa (Fig. 1, lane 2), as

FIG 2 Binding of GST-MsIA to RNA, ssDNA, and dsDNA. GST-MsIA-oligonucleotide complexes were separated from oligonucleotide in a 15% polyacrylamide gel in Tris-borate-EDTA buffer. Twenty picomoles of 5'-end fluorescein-labeled (U)₁₅, single-stranded (dT)₁₅, or annealed (dT)₁₅ and (dA)₁₅ was incubated with increasing concentrations of GST-MsIA in 20- μ l reaction mixtures for 10 min at 25°C. The amounts of MsIA in each reaction mixture were 0 pmol (lanes 1, 4, and 7), 20 pmol (lanes 2, 5, and 8), or 80 pmol (lanes 3, 6, and 9). Note that at similar concentrations of protein and oligonucleotide, relatively more dsDNA than ssDNA was incorporated within the less mobile product, while relatively more ssDNA was incorporated within the more mobile product.

determined by SDS-PAGE, while the molecular mass of GST-MGA0676 was approximately 55 kDa (data not shown).

Sequence analysis. The sequence of the cloned MGA0674 gene fragment from *M. gallisepticum* strain R_{low} was confirmed to be 2.169 kbp in length, as expected, which would encode a peptide of 723 amino acids with a predicted molecular mass of 79.308 kDa. This predicted molecular mass, plus that of GST (26 kDa), was consistent with that determined for GST-MsIA by SDS-PAGE (Fig. 1). The sequence of the cloned MGA0676 gene fragment from *M. gallisepticum* strain R_{low} was confirmed to be 0.768 kbp in length, as expected, which would encode a peptide of 253 amino acids with a predicted molecular mass of 28.441 kDa. This predicted molecular mass, plus that of GST (26 kDa), was consistent with that determined for GST-MGA0676 by SDS-PAGE.

A search of the MolliGen (version 3.0) database identified homologs of MGA0674 in *M. synoviae* MS53_0285/0329, *M. hyopneumoniae*, *M. bovis*, *M. genitalium*, *M. fermentans*, *M. conjunctivae*, and *M. pulmonis*. Multiple homologs of MGA0674 are found in *M. pneumoniae*, and these homologs have been classified as members of lipoprotein family 2. All these homologs possess two conserved domains, the *Mycoplasma* lipoprotein X central domain and the *Mycoplasma* lipoprotein 10 C-terminal domain. These domains have been detected only in the genomes of *Mycoplasma* spp. (10).

In *M. gallisepticum*, the gene encoding MGA0674 is located immediately 5' to MGA0676, which encodes a nuclease homolog, and immediately 3' to the nuclease homolog are three genes predicted to encode an ABC transport system. Some of the homologs of MGA0674 in other mycoplasmas are also located immediately 5' to homologs of the putative MGA0676 nuclease and the MGA0677, MGA0678, and MGA0679 ABC transport system genes.

Oligonucleotide binding properties of MsIA protein. To assess the ability of MsIA to bind to nucleic acid, we carried out gel mobility shift assays with 5'-end fluorescein-labeled oligonucleotides. These included ssRNA [(U)₁₅], ssDNA [(dT)₁₅], and dsDNA composed of (dT)₁₅ and (dA)₁₅. Two bands with reduced mobility that were not observed with the GST control protein were observed in binding reactions between GST-MsIA and either RNA or DNA oligonucleotides (Fig. 2). Most of the oligonucleo-

FIG 3 Binding of GST-MsIA to ssDNA and dsDNA. GST-MsIA-polynucleotide complexes were separated from oligonucleotides in a 2% agarose gel. Twenty picomoles of 5'-end-labeled single-stranded (dT)₃₀ or annealed (dT)₃₀ and unlabeled (dA)₃₀ was incubated under the same conditions described in the legend to Fig. 2. The amounts of GST-MsIA in each reaction mixture were 0 pmol (lanes 1 and 5), 20 pmol (lanes 2 and 6), 40 pmol (lanes 3 and 7), or 80 pmol (lanes 4 and 8). The intensity of fluorescence of dsDNA was expected to be half that of ssDNA because only one strand of the oligonucleotide (dT)₃₀ was labeled. Note that at similar concentrations of protein and oligonucleotide, relatively more dsDNA than ssDNA was incorporated within the less mobile product and more ssDNA than dsDNA was incorporated within the more mobile product.

tides had reduced mobility after the addition of one molar equivalent of MsIA, and the rate of migration of the bands with reduced mobility was not affected by higher protein concentrations. Although MsIA bound both RNA and DNA oligonucleotides, the highest binding affinity of MsIA was with ssDNA [(dT)₁₅], rather than with dsDNA or ssRNA oligonucleotides. There was a significant difference between ssDNA and dsDNA in the distribution of bound oligonucleotide between the two different complexes, with more ssDNA being bound within the more mobile complex and more dsDNA being bound within the less mobile complex at similar MsIA and oligonucleotide concentrations (Fig. 3). The minimal length of ssDNA bound by MsIA was examined by performing gel mobility shift assays with 5'-end fluorescein-labeled single-stranded dT oligonucleotides of 5, 10, 15, 20, 25, and 30 bases in length. Binding reactions with single-stranded 5- to 20-mer oligonucleotides generated two bands with reduced mobility with increasing concentrations of MsIA protein, while three or four bands with lower mobility were seen when 25- and 30-mer oligonucleotides were incubated with MsIA (the results of reactions with 5-, 10-, 20-, and 30-mers are shown in Fig. 4). Products with reduced mobility were not observed in binding reactions with exonuclease-digested 5'-end fluorescein-labeled 5-mer oligonucleotides. To examine whether binding to MsIA was influ-

FIG 4 Binding of GST-MsIA to ssDNA oligonucleotides of different lengths. Twenty picomoles of 5'-end fluorescein-labeled (dT)₅, (dT)₁₀, (dT)₂₀, or (dT)₃₀ oligonucleotides was incubated with 0 or 20 pmol of GST-MsIA. Binding products were analyzed by electrophoresis through a 15% polyacrylamide gel in Tris-borate-EDTA buffer.

FIG 5 Effect of ssDNA base composition on GST-MsIA binding. Twenty picomoles of 5'-end fluorescein-labeled (dT)₁₅, (dA)₁₅, (dC)₁₅, or (dG)₁₅ was incubated with 40 pmol of MsIA. Binding products were analyzed by electrophoresis through a 2% agarose gel. The brightness of the image was enhanced for the reaction with (dG)₁₅ to compensate for quenching of the end label on this oligonucleotide.

enced by the base composition of the oligonucleotide, 5'-end fluorescein-labeled single-stranded (dT)₁₅, single-stranded (dA)₁₅, single-stranded (dC)₁₅, or single-stranded (dG)₁₅ was incubated with MsIA, and the products were separated by 2% agarose gel electrophoresis. Although MsIA bound all four homopolymeric oligonucleotides, a greater proportion of (dT)₁₅ or (dC)₁₅ than (dA)₁₅ or (dG)₁₅ was bound at lower MsIA concentrations (Fig. 5). These results suggest that MsIA has a higher affinity for pyrimidines than for purines.

Competition binding assays with single nucleotides and peptides. The ability of MsIA to bind to peptides and single nucleotides was evaluated in competition binding assays between fluorescein-labeled ssDNA and unlabeled dNTPs or a tripeptide. No difference in the level of fluorescence of the DNA-protein complexes or the unbound oligonucleotide was observed when up to five molar equivalents of the competing substrate was included in the binding reaction between MsIA and 5'-end fluorescein-labeled (dT)₁₅ (Fig. 6). The means of the ratios of bound and unbound oligonucleotides (0.94 to 1.03) in competition reactions

FIG 6 Competition binding assays. The intensity of unbound (dT)₁₅ did not differ ($P < 0.001$) between the reaction mixture that contained 10 pmol GST-MsIA and 10 pmol (dT)₁₅ and the reaction mixtures that also contained 50 pmol of either glutathione, dATP, dCTP, dGTP, or dTTP as the competing molecule. The fluorescence of unbound (dT)₁₅ was more intense when GST rather than GST-MsIA was used in the reaction mixture. Oligonucleotide and oligonucleotide-protein complexes were separated in a 15% polyacrylamide gel in Tris-borate-EDTA buffer.

FIG 7 Effect of competing oligonucleotides of different lengths on GST-MslA binding. The intensities of unbound fluorescent oligonucleotides were compared to determine the binding efficiency of GST-MslA (14 pmol) for oligonucleotides of different lengths, with 10 pmol (dT)₅ allowed to compete with different combinations of 20 pmol of (dT)₁₀, (dT)₂₀, or (dT)₃₀. No differences in the intensity of different unbound oligonucleotides were observed, suggesting that there was a single binding site on MslA used for each oligonucleotide molecule. Oligonucleotide and oligonucleotide-protein complexes were separated in a 15% polyacrylamide gel in Tris-borate-EDTA buffer.

were not significantly different ($P < 0.001$, t test, performed in triplicate) from those in reactions performed in the absence of potential competitor substrates.

Competition binding assays to determine binding efficiency.

The ratios of the different unbound oligonucleotides in the competition binding reactions were similar to the ratios of the different oligonucleotides that were initially added to the binding reaction [1 (dT)₅:2 (dT)₁₀: (dT)₂₀, or (dT)₃₀] (Fig. 7).

Effect of incubation conditions on binding reaction. The effects of pH, ionic strength (sodium and potassium ions), the concentration of divalent cations (calcium and magnesium), incubation time, and incubation temperature on binding between recombinant GST-MslA and ssDNA and dsDNA were also assessed. Binding was examined at pH 6.3, 7.3, 8.3, and 9.3. The effect of varying the concentrations of sodium, potassium, calcium, or magnesium ions was examined at 0 mM, 20 mM, 40 mM, and 80 mM concentrations of each of the ions. Binding reaction mixtures were also incubated for 1, 5, 10, 15, or 20 min at 25°C and for 10 min at 25, 32, 37, or 42°C. No significant difference in binding was detected at the different pHs or temperatures assessed. The ssDNA-MslA complexes were produced within 1 min of incubation at 25°C. Variations in the concentrations of sodium, potassium, and magnesium ions did not affect binding, but improved binding of ssDNA by MslA was seen in 20 mM Ca²⁺ compared with that seen at lower concentrations. Higher concentrations of Ca²⁺ did not increase binding. Variations in pH, ionic strength, divalent cation concentration, incubation time, or incubation temperature had no discernible effect on binding between dsDNA and MslA.

Exonuclease I protection. To further characterize the properties of MslA, exonuclease I protection assays were carried out using single-stranded oligonucleotide (dT)₃₀. GST-MslA and (dT)₃₀ were incubated in exonuclease I reaction buffer at 25°C for 10 min, and then 0.04 U of exonuclease I was added. In the presence of MslA, (dT)₃₀ was partially protected from digestion by exonuclease I, and this protection was increased by increasing concentrations of MslA (Fig. 8).

FIG 8 Exonuclease I protection assay. Exonuclease I (0.04 U) was added 10 min after 20 pmol (dT)₃₀ and GST-MslA had been allowed to bind. Reaction products were analyzed by 2% agarose gel electrophoresis. The amount of GST-MslA in each reaction was 0 pmol (lane 1), 20 pmol (lane 2), 40 pmol (lane 3), or 80 pmol (lane 4).

Cleavage of GST-MslA *in vitro*. The fusion protein cleaved into two products over time (Fig. 9). The cumulative size of the observed products corresponded to the size of the whole fusion protein and, taking into account the GST fusion, also appeared to correspond to the sizes of the P22 and P57 products previously observed *in vivo*. The smaller fragment corresponding to P22 included the GST fusion partner, and the larger product (P57) appeared to preferentially bind to longer oligonucleotides (≥ 25 bp). Binding reactions run simultaneously on SDS-polyacrylamide and native polyacrylamide gels established that the multiple bands observed in the mobility shift assay on native polyacrylamide gels were not explained by cleavage of the MslA protein, as multiple bands of similar staining intensity were seen in the mobility shift assay even when the uncleaved fusion protein was the only band detectable by SDS-PAGE (Fig. 10).

The full-length protein could be stabilized by inclusion of phenylmethylsulfonyl fluoride during purification (data not shown), suggesting that the cleavage was likely to be catalyzed by a serine protease. The mass of the most abundant carboxyl-terminal fragment was determined to be 56.829 kDa by mass spectrometry, corresponding most closely to a cleavage site between threonine

FIG 9 SDS-PAGE and Western blot analysis of intact and cleaved GST-MslA. The gel on the left was stained with Coomassie blue, and the gel on the right was blotted and probed with anti-GST antibody to identify the amino-terminal cleavage fragment. A PageRuler prestained protein ladder (Pierce) was used as a molecular weight (MW) marker. Numbers to the right of the gel are molecular weights (in thousands).

FIG 10 Binding of full-length GST-MslA to ssDNA oligonucleotides of different lengths compared under different electrophoresis conditions. Reaction products were separated at 60 V for 3 h in corresponding lanes in a 15% polyacrylamide gel in Tris-borate-EDTA buffer (A and B) and a 15% SDS-polyacrylamide gel (C). The 15% polyacrylamide gel was visualized with UV transillumination (A) and then stained with Coomassie blue (B). The same reaction products were separated on a 15% SDS-polyacrylamide gel and stained with Coomassie blue (C), showing that only full-length GST-MslA was present, even though complexes with two different migration rates were detectable in native gels. A 5'-end fluorescein-labeled oligonucleotide ladder was run in lane 1. No oligonucleotide was included in the reactions in lanes 2 and 3. The reactions in lanes 4 to 8 each contained a single oligonucleotide of a different size in each lane.

231 and glutamate 232 in the prolipoprotein sequence (predicted to yield a carboxyl-terminal fragment of 56.823 kDa).

DISCUSSION

In a recent study of transposon-induced mutants of *M. gallisepticum*, MslA was shown to be a virulence factor, but its biochemical function has not previously been determined (10). We hypothesized that MslA is a nucleic acid binding protein because of the positional association between the gene that encodes it, a gene encoding a lipoprotein nuclease, and three genes encoding a putative ABC transporter (8). Using gel mobility shift assays, we showed that full-length mature MslA, expressed as a GST fusion protein, bound ssRNA, ssDNA, and dsDNA but bound ssDNA more efficiently than ssRNA or dsDNA. The efficiency of binding appeared to be equivalent for ssDNA oligonucleotides ranging in length from 5 to 30 bases, but MslA was unable to bind single nucleotides. The full-length protein produced two cleavage fragments, the larger of which preferentially bound ssDNA that was greater than or equal to 25 bp in length. Two binding complexes with distinct mobilities were detected in reactions between MslA and (dT)₁₅ or (dT)₂₀, while four distinct complexes were detected in reactions with (dT)₂₅ or (dT)₃₀ when products were analyzed by electrophoresis through 15% polyacrylamide gels. This suggested that the minimal binding site for MslA may be 5 to 10 nucleotides, but two distinct complexes were also seen in binding reactions with (dT)₅ and (dT)₁₀. These polyacrylamide gels were stained with Coomassie blue to compare the ratios of ssDNA and MslA in each of the complexes (data not shown), and each of the com-

plexes was found to contain similar relative amounts of MslA and ssDNA, suggesting that a single molecule of MslA bound to a single molecule of ssDNA. Thus, we conclude that different conformations of MslA-ssDNA can be formed, even with 5-mers, resulting in different rates of migration in polyacrylamide gels. The different conformations were observed even in the absence of the two cleavage products, suggesting that they are not the result of protein breakdown. A greater number of conformations appeared to be generated when reactions included oligonucleotides longer than 20 bases, possibly suggesting that MslA may bind these longer oligonucleotides at different points. The presence of the P57 cleavage product accounted for the most mobile conformation that was observed in binding reactions with longer oligonucleotides.

Previous studies on a carboxyl-terminal cleavage product of MslA recovered in *M. gallisepticum* cultured *in vitro* identified a cleavage site between the two alanine residues at positions 224 and 225 (10). The cleavage site predicted in our studies by mass spectrometry on *E. coli*-expressed MslA differed from this by 6 residues, although a much less abundant product with a mass of 57.490 kDa was detected, as were several intermediate products, suggesting that cleavage *in vitro* may also occur between the two alanine residues (predicted to result in a carboxyl-terminal product of 57.483 kDa), with amino-terminal degradation resulting in the intermediate products.

While we found that MslA could bind ssDNA composed entirely of any of the four nucleotides, oligonucleotides composed of pyrimidines appeared to bind with a higher affinity than those composed of purines. Other ssDNA-binding proteins, including *E. coli* SSB (19) and hRP-A (14), show a similar preference for pyrimidines.

Binding to ss- and dsDNA was not significantly affected by variations in pH or ionic strength, and complexes formed in less than 1 min at 25°C. Similar broad tolerances of reaction conditions have been seen for the SSB of *Mycoplasma pneumoniae* MPN229 and for other recombinant SSB proteins (20). However, binding of ssDNA, but not dsDNA, by MslA was sensitive to low concentrations of Ca²⁺. A similar sensitivity to divalent cations has been seen for the SSB proteins of *E. coli* (21) and *Streptococcus pneumoniae* (22). Because MslA had been shown to be a DNA-binding protein, we examined its capacity to protect ssDNA from exonuclease I. Increasing the concentrations of MslA increased protection of ssDNA from digestion in a dose-dependent manner. This study has shown that the lipoprotein MslA, a virulence factor of *M. gallisepticum*, is a novel promiscuous polynucleotide binding protein. Homologs of MslA are found in the genomes of a number of different mycoplasma species (8). Notably, in the genomes of the human pathogen *M. pneumoniae* (3, 23, 24), the porcine pathogen *M. hyopneumoniae* (8, 9), and the avian pathogen *M. synoviae* (8, 25), there are multiple paralogs of MslA, suggesting selection for immune evasion, although studies of transcription of these genes in *M. pneumoniae* have failed to find any evidence of phase variation (24, 26). As MslA is cell surface exposed and appears to be located within an operon that we confirmed here contains an exonuclease gene as well as a putative ABC transport system, we predict that its role is likely to be as a binding protein that delivers oligonucleotides to the exonuclease, which processes these oligonucleotides to generate individual nucleotides for transport by the ABC transporter. As MslA is not required for growth *in vitro* and does not appear to be required for coloni-

zation by some attenuated strains of *M. gallisepticum*, it is presumably not essential for importation of nucleotides into the mycoplasma cell. However, its significance in virulence suggests that it is likely to be required for optimal importation of nucleotides and, thus, for optimal growth when concentrations of nucleotides become limiting, which may occur during inflammation in the respiratory tract.

ACKNOWLEDGMENTS

This study was partially supported by Australian Research Council Discovery Project grant DP1095772. K.M. was supported by an Australian Government Endeavor Research Fellowship.

We thank Paul O'Donnell of the Bio21 Molecular Science and Biotechnology Institute for performing the mass spectrometry analysis.

REFERENCES

- Razin S, Yagov D, Naot Y. 1998. Molecular biology and pathogenicity of mycoplasmas. *Microbiol. Mol. Biol. Rev.* 62:1094–1156.
- Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM, Fritchman RD, Weidman JF, Small KV, Sandusky M, Fuhrmann J, Nguyen D, Utterback TR, Saudek DM, Phillips CA, Merrick JM, Tomb JF, Dougherty BA, Bott KF, Hu PC, Lucier TS, Peterson SN, Smith HO, Hutchison CA, III, Venter JC. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* 270:397–403.
- Himmelreich R, Hilbert H, Plagens H, Pirk E, Li BC, Herrmann R. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.* 24:4420–4449.
- Himmelreich R, Plagens H, Hilbert H, Reiner B, Herrmann R. 1997. Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. *Nucleic Acids Res.* 25:701–712.
- McLeod MP, Qin X, Karpathy SE, Gioia J, Highlander SK, Fox GE, McNeill TZ, Jiang H, Muzny D, Jacob LS, Hawes AC, Sodergren E, Gill R, Hume J, Morgan M, Fan G, Amin AG, Gibbs RA, Hong C, Yu XJ, Walker DH, Weinstock GM. 2004. Complete genome sequence of *Rickettsia typhi* and comparison with sequences of other rickettsiae. *J. Bacteriol.* 186:5842–5855.
- Hopfe M, Henrich B. 2004. OppA, the substrate-binding subunit of the oligopeptide permease, is the major Ecto-ATPase of *Mycoplasma hominis*. *J. Bacteriol.* 186:1021–1028.
- Schmidt JA, Browning GF, Markham PF. 2004. *Mycoplasma hyopneumoniae* p65 surface lipoprotein is a lipolytic enzyme with a preference for shorter-chain fatty acids. *J. Bacteriol.* 186:5790–5798.
- Schmidt JA, Browning GF, Markham PF. 2007. *Mycoplasma hyopneumoniae* mhp379 is a Ca^{2+} -dependent, sugar-nonspecific exonuclease exposed on the cell surface. *J. Bacteriol.* 189:3414–3424.
- Minion FC, Lefkowitz EJ, Madsen ML, Cleary BJ, Swartzell SM, Mahairas GG. 2004. The genome sequence of *Mycoplasma hyopneumoniae* strain 232, the agent of swine mycoplasmosis. *J. Bacteriol.* 186:7123–7133.
- Szczepanek SM, Frasca S, Jr, Schumacher VL, Liao X, Padula M, Djordjevic SP, Geary SJ. 2010. Identification of lipoprotein MslA as a neoteric virulence factor of *Mycoplasma gallisepticum*. *Infect. Immun.* 78:3475–3483.
- Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison CA, III, Smith HO, Venter JC. 2006. Essential genes of a minimal bacterium. *Proc. Natl. Acad. Sci. U. S. A.* 103:425–430.
- Dybvig K, Lao P, Jordan DS, Simmons WL. 2010. Fewer essential genes in mycoplasmas than previous studies suggest. *FEMS Microbiol. Lett.* 311:51–55.
- Minion FC, Jarvill-Taylor KJ, Billings DE, Tigges E. 1993. Membrane-associated nuclease activities in mycoplasmas. *J. Bacteriol.* 175:7842–7847.
- Kim C, Snyder RO, Wold MS. 1992. Binding properties of replication protein A from human and yeast cells. *Mol. Cell. Biol.* 12:3050–3059.
- Dabrowski S, Olszewski M, Piatek R, Brillowska-Dabrowska A, Konopa G, Kur J. 2002. Identification and characterization of single-stranded-DNA-binding proteins from *Thermus thermophilus* and *Thermus aquaticus*—new arrangement of binding domains. *Microbiology* 148:3307–3315.
- Kushner SR, Nagaishi H, Clark AJ. 1972. Indirect suppression of *recB* and *recC* mutations by exonuclease I deficiency. *Proc. Natl. Acad. Sci. U. S. A.* 69:1366–1370.
- Kur J, Hasan N, Szybalski W. 1989. Physical and biological consequences of interactions between integration host factor (IHF) and coliphage lambda late p' promoter and its mutants. *Gene* 81:1–15.
- Goldmark PJ, Linn S. 1972. Purification and properties of the *recBC* DNase of *Escherichia coli* K-12. *J. Biol. Chem.* 247:1849–1860.
- Meyer RR, Laine PS. 1990. The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol. Rev.* 54:342–380.
- Sluijter M, Hoogenboezem T, Hartwig NG, Vink C. 2008. The *Mycoplasma pneumoniae* MPN229 gene encodes a protein that selectively binds single-stranded DNA and stimulates recombinase A-mediated DNA strand exchange. *BMC Microbiol.* 8:167. doi:10.1186/1471-2180-8-167.
- Lohman TM, Ferrari ME. 1994. *Escherichia coli* single-stranded DNA-binding protein: multiple DNA-binding modes and cooperativities. *Annu. Rev. Biochem.* 63:527–570.
- Hedayati MA, Grove DE, Steffen SE, Bryant FR. 2005. Expression and purification of the SsbB protein from *Streptococcus pneumoniae*. *Protein Expr. Purif.* 43:133–139.
- Dandekar T, Huynen M, Regula JT, Ueberle B, Zimmermann CU, Andrade MA, Doerks T, Sanchez-Pulido L, Snel B, Suyama M, Yuan YP, Herrmann R, Bork P. 2000. Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res.* 28:3278–3288.
- Hallamaa KM, Browning GF, Tang SL. 2006. Lipoprotein multigene families in *Mycoplasma pneumoniae*. *J. Bacteriol.* 188:5393–5399.
- Vasconcelos AT, Ferreira HB, Bizarro CV, Bonatto SL, Carvalho MO, Pinto PM, Almeida DF, Almeida LG, Almeida R, Alves-Filho L, Assuncao EN, Azevedo VA, Bogo MR, Brigido MM, Brocchi M, Burity HA, Camargo AA, Camargo SS, Carepo MS, Carraro DM, de Mattos Cascardo JC, Castro LA, Cavalcanti G, Chemale G, Collevatti RG, Cunha CW, Dallagiovanna B, Dambros BP, Dellagostin OA, Falcao C, Fantinatti-Garboggini F, Felipe MS, Fiorentin L, Franco GR, Freitas NS, Frias D, Grangeiro TB, Grisard EC, Guimaraes CT, Hungria M, Jardim SN, Krieger MA, Laurino JP, Lima LF, Lopes MI, Loreto EL, Madeira HM, Manfio GP, Maranhao AQ, Martinkovics CT, et al. 2005. Swine and poultry pathogens: the complete genome sequences of two strains of *Mycoplasma hyopneumoniae* and a strain of *Mycoplasma synoviae*. *J. Bacteriol.* 187:5568–5577.
- Hallamaa KM, Tang SL, Ficorilli N, Browning GF. 2008. Differential expression of lipoprotein genes in *Mycoplasma pneumoniae* after contact with human lung epithelial cells, and under oxidative and acidic stress. *BMC Microbiol.* 8:124. doi:10.1186/1471-2180-8-124.