Association of a Major Protein Antigen of *Mycoplasma arthritidis* with Virulence

A.-H. T. Tu,1 B. Clapper,2 T. R. Schoeb,1 A. Elgavish,1 J. Zhang,1 L. Liu,1† H. Yu,1 and K. Dybvig1,2*

Department of Genetics1 and Microbiology,2 University of Alabama at Birmingham, Birmingham, Alabama

Received 9 July 2004/Returned for modification 13 August 2004/Accepted 17 September 2004

*Mycoplasma arthritidis* causes acute polyarthritis in rats and chronic proliferative arthritis in mice. *M. arthritidis*-induced arthritis serves as a model for arthritis caused by infectious agents and as a model for examining the role of the superantigen MAM (*M. arthritidis* T-cell mitogen) in the development of autoimmunity. *M. arthritidis* strain 158-1 is a spontaneous mutant of strain 158 that has a drastic reduction in virulence. We show that the mutant is missing a major antigen of 47 kDa (P47) and has acquired a protein of 67 kDa (P67), P47 and P67 partitioned into the detergent phase by extraction with Triton X-114. Coomassie blue staining of sodium dodecyl sulfate-polyacrylamide gels show that P67 is produced in abundance. Analysis of gel-purified P67 by mass spectrometry led to its identification as a lipoprotein (the open reading frame [ORF] 619 gene product) predicted from the genome sequence of *M. arthritidis*. PCR analysis of genomic DNA from 158 and 158-1 indicates that P47 and P67 are encoded by the same ORF 619 gene and differ only in the number of repeats in a tandem repeat region. By two-dimensional polyacrylamide gel analysis, no protein differences were detectable between 158 and 158-1 other than P47 and P67. Collectively, the data suggest that the tandem repeat region of P47 and P67 influences disease outcome.

Mycoplasmas are small, self-replicating organisms that lack a cell wall and cause disease in a wide variety of animal species. Mycoplasmas rely on the animal host to provide nutrients because they are deficient in many biochemical pathways such as the tricarboxylic acid cycle, an electron transport system, and de novo biosynthesis of amino acids, purines, and pyrimidines. *Mycoplasma arthritidis* is a natural pathogen of rats, and experimental diseases can be induced by intravenous or intra-peritoneal injection of a large number of organisms into rats and mice (6). Because a large number of mycoplasmas are required for experimental infection, it is unclear how disease occurs in the natural environment. Previous studies show that different strains of *M. arthritidis* vary in their ability to cause disease (5, 7, 9, 12, 13, 23). Several factors such as the MAM (*M. arthritidis* T-cell mitogen) superantigen (5), bacteriophage MAV1 (22), and the MAA cytadhesins (23) have been associated with virulence.

In this study, we describe a spontaneous mutant, 158-1, of *M. arthritidis* that is avirulent compared to the wild-type parental strain, 158. No protein or antigenic differences between the two strains were identifiable other than a 67-kDa (P67) protein in 158-1 that is missing in 158 and the P47 protein of 158 that is missing in 158-1. We show that P47 and P67 are major surface antigens encoded by the same gene and differ only with respect to the number of repeats in a tandem repeat region.

**MATERIALS AND METHODS**

**Bacterial strains.** Three strains of *M. arthritidis* were used in this study. Strain 158-1 (referred to previously as 1581) was selected as a random filter clone of a stock of strain 158 that had been passed 100 times (20). Strain 158L3-1 is a MAV1 lysogen of 158 (22). Mycoplasmas were grown overnight at 37°C in EB medium, as previously described (21). In preparation for infection of animals, cells were harvested by centrifugation, washed once in fresh EB medium, suspended in EB medium supplemented with 15% glycerol, and stored at −80°C. Before the animals were inoculated, an aliquot was thawed, serially diluted, and assayed for CFU to determine viability.

**Filter cloning of strains 158 and 158-1.** A mid-logarithmic growth-phase culture of 158 or 158-1 was filtered through a 0.45-μm-pore-diameter Acrodisc syringe filter (Gelman Sciences, Ann Arbor, Mich.) and assayed for CFU. Agar plugs containing individual colonies were picked, inoculated into fresh EB broth, and grown at 37°C to late-logarithmic growth phase. Glycerol was added to a final concentration of 15%, and the cultures were stored at −80°C for immunoblot analysis.

**Animals.** Female DBA/2NCr mice (4 to 8 weeks old, 10 to 20 g) were obtained from the National Cancer Institute (Frederick, Md.). Male Lewis rats (35 to 40 days old, 120 to 145 g) were ordered from Charles River Laboratories (Wilmington, Mass.). Animals were purchased from colonies free of rodent pathogens and maintained in a pathogen-free environment in Microisolator cages within the animal facilities at the University of Alabama at Birmingham.

**Induction and assessment of disease.** Mice or rats were injected intravenously with 200 μl of mycoplasma stock containing 10⁷ or 10⁸ CFU, respectively, or control broth. Each group included five to six animals. Each experiment was repeated at least once.

To evaluate arthritis in mice, animals were sacrificed 3 weeks after inoculation. Blood from infected and uninfected animals was collected and allowed to clot to obtain serum for immunoblot analysis. The wrists and ankle joints were excised, fixed in 70% alcohol formalin, decalcified, sectioned at a 5-μm thickness, and stained with hematoxylin-eosin for histopathologic evaluation. Histopathologic examination was done by the Comparative Pathology Laboratory at the University of Alabama at Birmingham. Severity of arthritis was quantified by a pathologist (T.R.S.) who was not provided with the identity of the experimental groups. As described elsewhere (19), joints were subjectively assigned a score from 0 (normal) to 5 (severe) based on five characteristic histologic lesions of arthritis: inflammation, joint-capsule thickening, tendonitis, bone destruction/remodeling, and ankylosis. For each animal, an overall pathology score was obtained by averaging the scores from the wrist and ankle joints. Thus, the theoretical maximum arthritis score for each animal was 5. For most experiments, one wrist and one ankle joint from each animal were evaluated by histopathology. The remaining wrist and ankle joint were cultured for mycoplasmas as described below for culture analysis of the rat joints.

To evaluate disease in rats, animals were weighed every 2 days and sacrificed at the end of 2 weeks. All four wrist and ankle joints were collected, skinned, and...
pooled from each animal. The joints were minced in 4 ml of EB medium, sonicated as described previously to liberate mycoplasmas from the tissue (19), and assayed for CFU.

**PCR and sequence analysis of ORF 619 of M. arthritidis genome.** PCR analysis was performed according to standard procedures. Genome sequence data from *M. arthritidis* strain 158L3-1 were obtained from The Institute for Genomic Research through the web site at http://www.tigr.org. Open reading frame (ORF) 619 is located from nucleotide position 525815 to position 526579. Primers were synthesized with the following 5'-to-3' sequences: P47F, GCT AAACAGGATTATCAAACGG; P47R, CGCTTCAAATAGGTCAATGTGCC; P47-3 (forward primer), CTCACGGTTGACAGCGCCGTG; and P47-4 (reverse primer), CTCACGATGAAAGATCGTTG. P47F and P47R anneal at positions 525376 and 526913, respectively. P47-3 and P47-4 anneal at positions 525882 and 526497, respectively. Different combinations of forward and reverse primers were used to determine the size of the repeat region of ORF 619 in strains 158, 158L3-1, and 158L3-1. PCR products were purified from an agarose gel and sequenced by automated dye terminator methods at the Iowa State University DNA Synthesis and Sequencing Facility, Ames. The complete sequence of the PCR product was determined for ORF 619 from strains 158 and 158L3-1. The complete sequence of the PCR product of ORF 619 from strain 158L3-1 could not be determined because of the large number of tandem repeats.

**Western analysis and TX-114 phase extraction.** Protein samples were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (10% polyacrylamide) and electrotransferred onto nitrocellulose membranes for immunoblot analysis (18). Protein sizes were estimated by comparison to Bio-Rad Kaleidoscope prestained molecular weight standards. To determine if P47 and P67 are associated with the membrane, mycoplasmal proteins from 25 ml of cell culture were extracted with the nonionic detergent Triton X-114 (TX-114) (24). Aliquots of the whole-cell lysate, the aqueous phase, and the detergent phase were resolved by SDS-PAGE and either stained with Coomassie blue to detect the relative abundance of the proteins or transferred to nitrocellulose for immunoblot analysis. Sera for Western analysis consisted of pools of serum collected from mice or rats at 3 or 2 weeks postinoculation (p.i.), respectively. Pooled sera at a dilution of 1:100 were used for Western analysis. The secondary antibody was goat anti-mouse or goat anti-rat, conjugated with alkaline phosphatase.

For two-dimensional (2D) PAGE, cells from a 1-ml culture were harvested by low-speed centrifugation and analyzed by the Mass Spectrometry and Proteomics Shared Facility at the Comprehensive Cancer Center of the University of Alabama at Birmingham. Proteins were solubilized in 7 M urea, 2 M thiourea, 4% ((3-cholamidopropyl)dimethylamino)-1-propane-sulfonate (CHAPS), Complete Mini protease inhibitor cocktail (Roche), 2% amphotelys (pH 3 to 10), and 5 mM tributyl phosphine (TBP). Samples were sonicated on ice by using quick bursts to minimize heating, set at room temperature for 20 min, and centrifuged at 100,000 × g for 30 min at 22°C to remove any nucleic acids or insoluble material. For isoelectric focusing, samples were separated with Bio-Rad 11-cm immobilized pH gradient strips (pH 3 to 10) with the Multiphor II apparatus (Amersham). The focusing conditions were 0 to 800 V with a 5-min ramp constant 300 V for 5 min, 300 to 3,500 V with a 1.5-h ramp, and 3,500 V held constant for 4.5 h. Precast Criterion gradient gels (8 to 16%) were used to resolve proteins for the second dimension. Gels were stained with Sypro Ruby (Molecular Probes, Inc., Eugene, Oreg.) to visualize proteins.

**Mass spectrometry.** P67 was excised from SDS-PAGE gels, trypsinized, and analyzed by the Comprehensive Cancer Center Mass Spectrometry Shared Facility, using matrix-assisted laser desorption ionization–time-of-flight mass spectrometry (MALDI-TOF MS with a Voyager DE-Pro instrument; Applied Biosystems) and liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI MS/MS; Micromass Q-TOF2 instrument; Micromass, Manchester, United Kingdom). The results obtained from MALDI-TOF and LC-ESI MS/MS were compared, and the computer-generated trypsin fragments of each of the putative lipoproteins were predicted from the *M. arthritidis* genome sequence (http://www.tigr.org) by using MacVector (Accelrys, Inc., San Diego, Calif.) and ProteinProspector (http://prospector.ucsf.edu/).

**Statistical analysis.** Parametric data were analyzed by analysis of variance with supplemental mean comparisons by Bonferroni’s method (*P < 0.05*). Nonparametric data were analyzed by the Mann-Whitney rank-sum test.

## Results

**Comparison of virulence of strains 158 and 158-1.** Mice were inoculated with either *M. arthritidis* strain 158 or 158-1. At the end of 3 weeks, animals were sacrificed and joints were collected for histopathological studies. Strain 158 caused greater disease than did 158-1. Histopathology results indicated an increased level of exudates in the joint cavity, synovial and capsular inflammation, subchondral bone involvement, and tendonitis for mice inoculated with 158. These mice had a mean arthritis score of 1.7 ± 0.14, whereas the joints of mice infected with 158-1 were normal, with statistically lower arthritis scores of 0.34 ± 0.05 (*P < 0.05*). Similar experiments were performed with Lewis rats. Weight loss was used as an indicator of the level of disease since previous studies have shown that weight loss corresponds with severity of arthritis (22). Rats infected with strain 158 lost weight at 3 days p.i., whereas rats infected with strain 158-1 continued to gain weight (Fig. 1). Weight loss in 158-infected animals occurred during the first 8 days p.i. and was followed by a recovery period. Mycoplasmas were recovered from the joints of each of the infected rats at 2 weeks p.i., but 1,000-fold more CFU were obtained from rats infected with 158 than from those infected with 158-1 (*P = 0.004*). Collectively, these results show that strain 158 is highly virulent compared to 158-1.

**Antigenic differences among various isolates of *M. arthritidis*.** Sera from mice inoculated with either strain 158 or 158-1 were used to examine immunoblots of *M. arthritidis* proteins, the results being identical regardless of the source of sera. An immunodominant protein of 47 kDa was detected in 158 but not in 158-1. A larger protein of 67 kDa was present in 158-1. Eleven isolates of *M. arthritidis* from the joints were selected for further study (4 isolates from mice inoculated with 158-1 and 7 isolates from mice inoculated with 158) by immunoblot analysis. Except for minor variation in the size of the P47 and P67 proteins as noted below, the protein profile of the 11 isolates was essentially identical to that of the parent inoculum. A protein smaller than P67 (about 64 kDa) was detected for two of the four isolates obtained from mice inoculated with 158-1 (Fig. 2, lanes 1 and 2). Filter cloning of the isolate analyzed in Fig. 2, lane 1, revealed that some subclones produced P67 and others produced P64 (data not shown). One of the seven isolates from mice inoculated with 158 had a 49-kDa protein, slightly larger than P47 (Fig. 2, lane 10). These results suggest that P47 and P67 undergo size variation, a common occurrence among mycoplasmal surface proteins (1).
proline linkages are generally refractory to digestion with trypsin because of the internal lysine residues in this fragment because lysine-O
ORF 619 gene product. Trypsin should not cleave at some of the amino-terminal fragment of the ORF 619 gene product, which is predicted to have covalently attached fatty acid (see below). The LC-ESI MS/MS sequences confirmed the identity of the 8-amino-acid KPEGETPK repeat of ORF 619 sometimes has a single-amino-acid substitution to KPESETPK, and the mass spectrum also identified this peptide (data not shown). The predicted ORF 619 gene product has a consensus sequence corresponding to a prokaryotic lipoprotein signal peptide (identified from PROSITE at http://au.expasy.org/prosite/), consistent with the TX-114 extraction data shown in Fig. 4.

P47 and P67 are encoded by the same gene but differ in the number of tandem repeats. To determine whether P47 is encoded by the same gene as P67, PCR analyses were performed using primers that span the repeat region of ORF 619 (Fig. 6). The PCR products of 1,410 and 1,215 bp were obtained from 158L3-1 and 158 template DNA, respectively, using primer pair P47F and P47-4 (Fig. 7). The nucleotide sequences of the PCR products confirmed that both products corresponded to ORF 619. By calculation, the 195-bp difference in the size of the PCR products from 158L3-1 and 158-1 templates corresponds to eight copies of the 24-bp tandem repeat of ORF 619. From the genome sequence of strain 158L3-1 (a MAV1 lysogen of 158 that produces P47), P47 is predicted to have 23 tandem repeats, 19 of which are PKKPEGET and 4 of which are PKKPESET. From the PCR data shown in Fig. 7, we predict that P67 has a total of 31 tandem repeats.

PCR amplification of ORF 619 from filter clones of 158L3-1 and from mycoplasmas isolated from animals that had been infected with 158-1 revealed further variation in the number of tandem repeats in the gene. The size of the PCR product from ORF 619 indicates that the P47 protein in Fig. 2, lanes 1 and 2, has about 29 tandem repeats. The protein produced by the filter clones analyzed in Fig. 3, lanes 6 and 8, had about 44 and 27 tandem repeats, respectively.

The gene product, minus the signal peptide sequence, predicted for P47 from the nucleotide sequence of ORF 619 of strain 158L3-1 is a 24.8-kDa lipoprotein with a pI of 4.8. P67, with 31 tandem repeats, is predicted to be 31.8 kDa with a pI of 5.0. There are numerous examples of mycoplasmal lipopo-
teins that have extensive tandem repeat domains. Like P47, these proteins generally migrate on SDS-PAGE gels with an apparent molecular mass considerably higher than the mass predicted from sequence data (2). Database comparisons revealed no homologues of the ORF 619 gene product, and no orthologue in *M. arthritidis* was identified from the complete genome sequence of strain 158L3-1.

**DISCUSSION**

Our studies of the *M. arthritidis* phage MAV1 (4) led to the realization that 158-1 was a relatively avirulent, spontaneous mutant of strain 158 (see Fig. 1). The ability of 158 to cause disease may be associated with the P47 protein that is replaced with the elongated P67 protein in the mutant. The P47 and P67 proteins are highly immunogenic, being the predominant antigens recognized by sera from mice infected with either of these two strains. We propose referring to P47 and P67 as the MIA (mycoplasma immunodominant antigen) proteins. The MIA proteins are abundant. The predominant band on Coomassie-stained gels of 158-1 proteins extracted with TX-114 corresponded to P67 (data not shown). P47 and P67 are size-variable proteins that differ from one another by eight tandem repeat units, with each repeat unit having the sequence PKKPEGET or PKKPESET. Variation in the number of tandem repeats presumably occurs as a result of slipped-strand mispairing during DNA replication of the tandem repeat region (8). Size variation through slipped-strand mispairing is a common occurrence with mycoplasmal lipoproteins such as

![Image](http://iai.asm.org/)

**FIG. 5.** Tandem mass spectrum for the doubly charged ion (876.4) of the tryptic peptide KPEGETPKKPEGETPK derived from P67.

![Image](http://iai.asm.org/)

**FIG. 6.** Schematic diagram of the ORF 619 gene and predicted amino acid sequence of its product. The white region of the gene encodes the amino-terminal signal peptide indicative of lipoproteins (bold amino acids in the sequence). The gene’s striped region encodes the tandem repeats (thin underline PKKPEGET repeats or thick underline PKKPESET repeats in the amino acid sequence). The black region encodes nonrepetitive portions of the protein. The positions of the primers used for PCR analysis of the gene are indicated by arrows above the diagram. The schematic of the gene and flanking primer-binding sites is to scale, with the coding region consisting of 765 bp.

Previous work indicates that pathogenic mycoplasmas use size variation to alter the surface-exposed proteins that are major targets of the host immune response. In *M. hyorhinis*, the elongated Vlp molecule protects the organism from host growth-inhibiting antibodies (3). In *M. pulmonis*, the elongated VsaA protein protects the mycoplasma from complement-mediated killing (16). The difference in virulence between *M. arthritidis* strains 158 and 158-1 is not due to differences in complement susceptibility. Preliminary data in our laboratory indicate that 158 and 158-1 resist complement killing under conditions in which *M. pulmonis* cells producing a short Vsa protein are killed (B. T. Matthews and K. Dybvig, unpublished data). Unlike Vlp and Vsa, in which the elongated forms of the protein are associated with resistance to host defenses, *M. arthritidis* strain 158 with the shorter P47 protein was more virulent than 158-1 producing P67.

The relationship of the repeats of P47 and P67 with virulence is reminiscent of the alpha C proteins of group B streptococci (10, 11). The alpha C protein is a protective surface-associated antigen found in many clinical isolates. By reducing the number of repeats in the alpha C protein, the pathogenicity of the strain was enhanced due to a reduction of the host immune response to the protein (11). The association of the shorter alpha C protein with virulence appears to be dependent on the production of specific host antibody. In contrast, the difference in virulence between strains 158 and 158-1 is apparent as early as 3 days p.i. and should thereby be independent of specific antibody to P47 and P67.

Whether strain 158-1 is relatively avirulent because it has an elongated ORF 619 or because it has some other unidentified mutation in its genome is unknown. However, by 2D-PAGE, no protein differences were noted between 158 and 158-1. If the loss of virulence is a result of the increased number of tandem repeats in the mia gene (ORF 619), it is possible that P67 masks other mycoplasmal surface components, similar to the examples of epitope masking described in other species of mycoplasma (17, 25). If masked by P67, surface components important for antiphagocytosis, colonization, or another aspect of virulence may not properly interact with host factors. Another possibility is that cells producing P67 could be at a growth disadvantage in vivo compared to cells producing P47, but no obvious differences were noted during growth in vitro. To investigate whether P47 and P67 affect phagocytosis, future experiments could involve testing the virulence of 158 and 158-1 in mice deficient in inducible nitric oxide synthase. A thorough study of size-variable surface proteins of *M. arthritidis* will provide an understanding of the mechanism(s) in which the mycoplasm establishes a chronic, persistent infection.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant AR44252 from the National Institutes of Health. Preliminary sequence data were obtained from The Institute for Genomic Research through the web site at http://www.tigr.org.

We thank Portia Caldwell for technical help and Marion Kirk for assistance with electrospray analysis.

REFERENCES