INFECTION AND IMMUNITY, Mar. 1995, p. 891–898 Vol. 63, No. 3
0019-9567/95/$04.00+0
Copyright © 1995, American Society for Microbiology

Small Repeating Units within the *Ureaplasma urealyticum* MB Antigen Gene Encode Serovar Specificity and Are Associated with Antigen Size Variation

XIAOTIAN ZHENG,[1] LEE-JENE TENG,[1]† HAROLD L. WATSON,[1] JOHN I. GLASS,[1] ALAIN BLANCHARD,[2] AND GAIL H. CASSELL[1]*

Department of Microbiology, University of Alabama at Birmingham Schools of Medicine and Dentistry, Birmingham, Alabama 35294,[1] and Unité d'Oncologie Viral, Institut Pasteur, Paris, France[2]

Received 17 October 1994/Returned for modification 12 December 1994/Accepted 28 December 1994

***Ureaplasma urealyticum* is a common commensal of the female lower urogenital tract, yet it has been shown to be an important cause of chorioamnion infection, respiratory and central nervous system disease, and death in premature infants. It has been suggested that only certain serovars are capable of producing invasive disease. However, we previously showed that many serotypes are invasive and that perhaps antigen variability and host factors are more important determinants of ureaplasma infections than are different serotypes per se. The molecular characterization in this report describes a mechanism available to ureaplasmas for producing antigen variation. That antigen, designated MB and previously identified on *U. urealyticum*, contains serovar-specific and cross-reactive epitopes, is produced both in vitro and in vivo, is a predominant antigen recognized during ureaplasma infections of humans, undergoes a high rate of size variation in vitro, and is size variable on invasive ureaplasma isolates. In the present study, we cloned and sequenced the gene of the MB antigen from serovar 3, the serovar most commonly isolated from humans. The 3′ two-thirds of the gene was shown to contain identical 18-nucleotide tandem repeats. PCR analysis and direct sequencing of two variants indicated that alterations within this repeat region are responsible for the size variation of the MB antigen. Intact recombinant serovar 3 MB antigen and truncated products, expressed by coupled in vitro transcription and translation of the cloned gene, were immunoprecipitated by both a serovar-specific monoclonal antibody and the serum of a *U. urealyticum*-infected patient, and these results identified the repeat region of the MB antigen as serovar defining. Resolution of the precise amino acids responsible for specific epitopes and characterization of similar genes in the other serovars should yield reagents useful in elucidating the role of antigen size variants in disease production and the role of specific antibody in protection from ureaplasma disease.**

The exact significance of *Ureaplasma urealyticum* in human disease is controversial, particularly as it relates to pregnancy outcome (7). Recent studies indicate that *U. urealyticum* is a definite cause of chorioamnionitis (6) and is a significant cause of respiratory disease (3), meningitis (24), and death in infants (5). A principle repeatedly illustrated in these studies is that only in a subpopulation of individuals infected in the lower genitourinary tract do the organisms reach the upper tract and only in some of these individuals does disease ensue (4). Furthermore, it is apparent that only a subpopulation of infected infants, primarily those with <30 weeks of gestation, are at risk for development of disease (5, 24). Identification of either patient or *U. urealyticum* markers that correlate with disease or increased susceptibility to *U. urealyticum* disease would eliminate much of the controversy about the pathogenic nature of *U. urealyticum.*

The initial difficulty in proving an etiologic role for ureaplasmas in adverse pregnancy outcome supplied a major impetus for the hypothesis that only certain subgroups of the species are truly disease associated. *U. urealyticum* has been divided by a number of serological methods into 14 serovars (16), but more recently we have found that within individual serovars a major antigen, designated MB (multiple banded), is present in different sizes in different clinical isolates (25, 27, 32). Thus, there is an additional level of genetic variability beyond the serovar category which may contribute to ureaplasmal pathogenesis. Significantly, size variation of an analogous antigen in *Mycoplasma pulmonis*, a pulmonary and genitourinary pathogen of rodents, is the only observed difference between virulent and avirulent strains related to differences in disease severity and incidence of death (9, 19, 26).

Our previous phenotypic characterization of MB antigen shows that it (i) is species specific; (ii) contains both serovar-specific and cross-reactive epitopes, i.e., it can be recognized by both serovar-specific and cross-reactive monoclonal antibodies (MAbs); (iii) is produced not only in vitro but also in vivo; (iv) undergoes a high rate of size variation in vitro; (v) is present and size variable on invasive ureaplasma isolates (i.e., those from placenta, lung tissue, and cerebrospinal fluid); and (vi) is among the predominant antigens recognized during infections in humans (25). Furthermore, we have shown that MAbs to these antigens can inhibit the growth of the organisms in vitro, indicating the potential importance of this antigen for host defense (25). By using MAbs to the MB antigen, we have recently shown that invasive infection is not limited to one serovar or a few particular serovars and that the serovars that are isolated from patients with invasive *U. urealyticum* express size-variant MB antigens (7, 25, 32). The purpose of the present investigation was to clone and sequence the MB antigen gene to gain further understanding of its variation within *U. urealyticum* and to provide a rational basis for development

* Corresponding author. Mailing address: BBRB, Department of Microbiology, University of Alabama at Birmingham, Birmingham, AL 35294. Phone: (205) 934-9339. Fax: (205) 934-9256.

† Permanent address: School of Medical Technology, College of Medicine, National Taiwan University, Taipei, Taiwan, Republic of China.

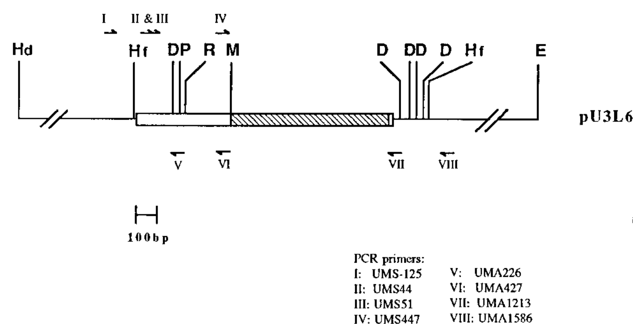

FIG. 1. Physical map of the recombinant plasmid. The position of the serotype 3 *U. urealyticum* MB antigen gene is indicated by a bar in which the shaded area represents the repeat region. The arrows indicate the positions and orientations of PCR primers used to study size variation. Restriction sites: Hd, *Hin*dIII; Hf, *Hin*fI; D, *Dra*I; P, *Pvu*II; R, *Rsa*I; M, *Mva*I; E, *Eco*RI.

of reagents to address the roles of different serovars, different antigen size variants, and specific antibody (20, 21, 23) in invasive disease.

MATERIALS AND METHODS

Ureaplasmas. The *U. urealyticum* serovar 3 reference strain was obtained from E. A. Freundt (Institute of Medical Microbiology, University of Aarhus, Aarhus, Denmark). Two serovar 3 clinical isolates were isolated at the University of Alabama at Birmingham (UAB). Isolate 1546 was from lung tissue of a newborn infant, and isolate 8315 (M472 in reference 27) was from the placenta of another patient.

Antibodies. The four MAbs used were described previously (25). MAb 3B1.5 reacted strongly with serovar 3 with faint cross-reactions with other serovars. MAb 5B1.1 cross-reacted equally with serovars 3 and 14. MAb 10C6.6 was serovar 3 specific. MAb 8B5.2 showed distinct cross-reactions with 12 of the 14 serovars (only serovars 2 and 5 did not react). Sera from two patient were used in this study. The positive patient serum was reactive to the *U. urealyticum* serovar 3 MB antigen on an immunoblot described in a previous publication (25), in which it was designated serum B2. The negative patient serum had no immunoblot reactivity to the *U. urealyticum* serovar 3 antigen (data not shown).

Antigen purification, protease digestion, and peptide sequencing. MAb 3B1.5, directed to the serovar 3 MB antigen, was used for affinity purification of the MB antigen as described previously (25). The purified antigen was lyophilized and trypsin digested, and fragments were separated by reverse-phase high-pressure liquid chromatography (C_{18}). Fractions were collected individually, and the N-terminal sequence was determined by the UAB Protein Analysis and Peptide Synthesis Core Facility.

Oligonucleotide primers. Oligonucleotides for sequencing, Southern blots, and PCR were synthesized by the Oligonucleotide Synthesis Core Facility in the UAB Cancer Center. The oligonucleotide nomenclature is as follows: molecules are named UMS or UMA for ureaplasma MB antigen coding sense primers or antisense primers, respectively. The number following UMS or UMA refers to the location on the MB antigen gene map corresponding to the most 3′ base of the oligonucleotide (the positions of PCR primers are shown in Fig. 1).

DNA hybridization. Southern blot analysis was used to identify DNA fragments of interest from genomic DNA of the *U. urealyticum* serovar 3 reference strain and for screening of positive transformants. After restriction endonuclease digestion, *U. urealyticum* genomic DNA or plasmid DNA with inserts was electrophoresed in agarose gel and transferred onto nylon membranes (Hoefer Scientific Instruments). The prehybridization-hybridization solution contained 1 M NaCl, 50 mM Tris-HCl (pH 7.4), 0.5% sodium dodecyl sulfate (SDS), 5× Denhardt's solution, and 0.02% salmon sperm DNA. The hybridization temperature for the 18-nucleotide probe was 42°C.

Molecular cloning and sequencing. The MB antigen genes of the *U. urealyticum* serovar 3 reference strain and isolate 8315 were cloned by different methods. Genomic DNA of the reference strain was double digested with *Hin*dIII and *Eco*RI and analyzed by Southern blot hybridization with a ^{32}P-labeled oligonucleotide probe whose design was based on amino acid sequencing information. DNA fragments of interest were purified from agarose gels with Gene Clean (Bio 101, Inc., La Jolla, Calif.) and ligated to plasmid vector pUC19, which had been digested with the same two enzymes. The ligation product was transformed into *Escherichia coli* DH5α. Positive transformants were identified by Southern hybridization with the probe described above.

Cloning of the MB antigen gene from *U. urealyticum* isolate 8315 was achieved by ligation of PCR products (see Results) into the pCRII plasmid with a TA Cloning Kit (Invitrogen, San Diego, Calif.).

We sequenced the MB antigen genes by using plasmid templates. Both DNA strands were sequenced with Sequenase version 2.0 (United States Biochemicals). To obtain the sequence of a region of the MB antigen gene containing many copies of the same 18-bp repeat, we made a nested set of exonuclease III deletions beginning at a *Dra*I site upstream of that repeat region (18). Sequence data were analyzed with McVector (IBI, New Haven, Conn.) and the Genetics Computer Group Sequence Analysis Software Package, version 7.2 (University of Wisconsin Genetics Computer Group).

PCR. Template DNA was obtained as previously described (2). The denaturation, annealing, and elongation temperatures and times used were 94°C for 20 s, 62°C for 1 min, and 72°C for 1 min, respectively, for 35 cycles. The conditions for amplification of regions containing repeats were 94°C for 50 s, 57°C for 2 min, and 72°C for 2 min, respectively, for 30 cycles.

Synthetic peptide analysis. The peptide of interest was chemically synthesized (solid phase) by the UAB Protein Analysis and Peptide Synthesis Core Facility (430A Peptide Synthesizer; Applied Biosystems), and antibody reactivity was examined by enzyme-linked immunosorbent assay (ELISA) (25). The wells were coated with 80 μl of a solution containing 50 μg of protein per ml.

Eukaryotic in vitro translation of the MB antigen. We subcloned and mutagenized the MB antigen gene by PCR so that it could be expressed in a eukaryotic coupled in vitro transcription-translation system. As the template for the PCR we used the plasmid that contained the 6-kbp *U. urealyticum* serovar 3 chromosomal fragment, and as primers we used a coding sense 59-base oligonucleotide with the sequence 5′-ATGTCGACGCCAGCCATGAAATTATTA AAAAATAAAAAATTCTGGGCTATGACATTAGG-3′ (UMS44) and an antisense 57-base oligonucleotide with the sequence 5′-GTCGACTAGTCACT TAGAGCTCCCGAGGTAACTAAAGTATTATTTTCCAGTAGTTTC-3′ (UMA1213). DNA fragments produced in the PCR with these primers contained a *Sal*I restriction enzyme site and a consensus eukaryotic translation initiation signal (13) upstream of the MB antigen gene start codon. The antisense primer was used to add *Xma*I, *Sac*II, and *Sal*I restriction sites, as well as stop codons in all three reading frames, to the 3′ end of the gene. This PCR product was ligated into plasmid pCRII (TA Cloning Kit; Invitrogen) behind a T7 RNA polymerase promoter. That plasmid, designated MB, was used in a coupled in vitro transcription-translation system (TNT; Promega Biotech) that uses T7 RNA polymerase and mRNA-dependent rabbit reticulocyte lysates. The reactions were performed in accordance with the manufacturer's instructions at 30°C for 90 min. The in vitro translation products were labeled with either [^{3}H]leucine or [^{3}H]proline. Aliquots of the reactions were either immunoprecipitated and then analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) or analyzed directly by SDS–15% PAGE. The protein gels were fluorographed (8) and used to expose X-ray film.

Two MB antigen genes with deletions were also created for analysis in the eukaryotic coupled in vitro transcription-translation system. The plasmid containing deletion mutant R, which had only the 3′ two-thirds of the gene, was constructed by the method described in the previous paragraph. A coding sense primer, 5′-GCGAATTCGCCGCCAGCCATGCATCTAAGAGAAAAAGTG AACTTTGAAACTACACAA-3′ (UMS447), was used along with antisense primer UMA1213 to generate a PCR product with a start codon in a consensus eukaryotic translation initiation context that was 9 codons upstream of the MB antigen gene repeat region. A second deletion mutant, plasmid ΔMB, was derived by partially digesting plasmid MB with restriction enzyme *Bsp*MI. Selected partial digestion products were treated with T4 DNA ligase and transformed into competent *E. coli* cells. On the basis of nucleotide sequence analysis, a transformant was selected whose plasmid, ΔMB, contained only three 18-nucleotide repeats. This corresponded to a deletion of 711 ± 18 nucleotides from the MB antigen gene.

Immunoprecipitation. Twenty-microliter aliquots of in vitro transcription-translation reactions were mixed with 600 μl of MAbs or patient sera diluted 1:200 in phosphate-buffered saline (PBS). These antibody-antigen mixtures were incubated at 4°C. After overnight incubation, 25 μg of protein G-Sepharose 4 Fast Flow (Pharmacia) in 100 μl of PBS was added to each sample. The protein G-Sepharose beads had been preabsorbed overnight in PBS containing 2% rabbit reticulocyte lysate. The samples were mixed gently at room temperature for 2 h. The Sepharose beads were pelleted from suspension and were washed three times with PBS. The beads were resuspended in 25 μl of SDS-PAGE sample buffer and boiled for 3 min prior to being loaded onto a protein gel.

Nucleotide sequence accession number. The GenBank accession number for the MB antigen gene of the serovar 3 reference strain is L20329.

RESULTS

Cloning of the MB antigen gene of the *U. urealyticum* serovar 3 reference strain. Affinity-purified MB antigen was digested with trypsin, the fragments were separated by high-pressure liquid chromatography, and a single oligopeptide was analyzed by amino acid sequencing. The peptide had the sequence, listed from the N terminus, Glu-Gln-Pro-Ala-Gly. Because trypsin cleaves peptide bonds at the carboxy terminus of either lysine or arginine, it followed that either lysine or arginine

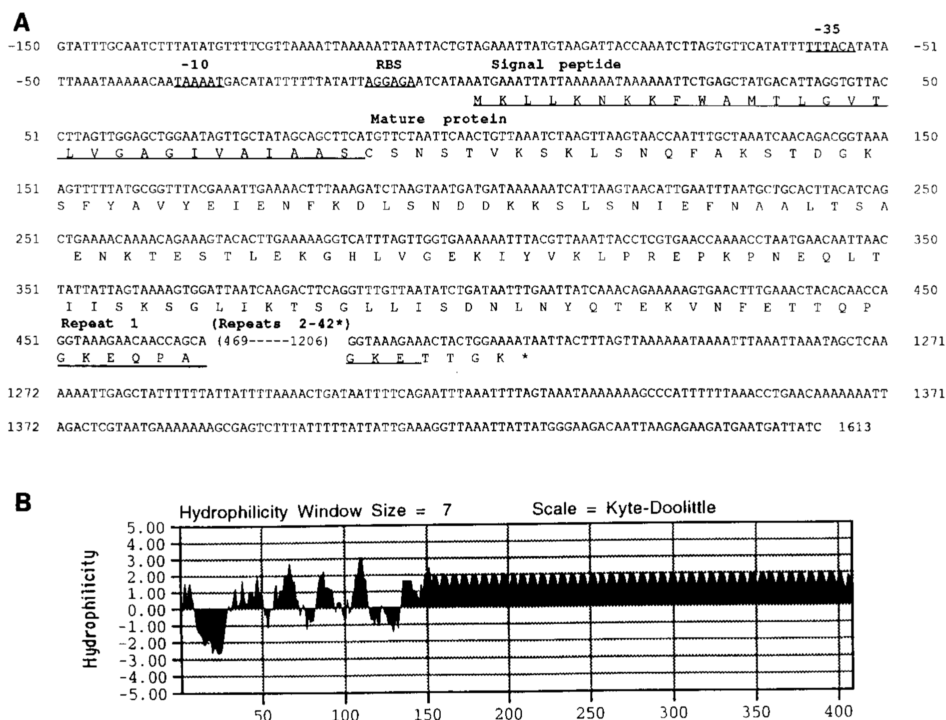

FIG. 2. Nucleotide and amino acid sequence analysis. (A) Nucleotide and deduced amino acid sequence of serovar 3 *U. urealyticum* MB antigen gene and flanking regions. The areas underlined are the predicted promoter consensus site, the ribosomal binding site (RBS), the signal sequence, and the repeat region. (B) Kyte-Doolittle hydrophilicity plot generated with the McVector software. Positive values on the *y* axis indicate predicted hydrophilic regions with a greater than 50% chance of surface probability. Values on the *x* axis indicate amino acid numbers in the sequence. The total copy number of repeats was obtained by estimating on an agarose gel the size of the DNA fragment containing the repeat region; this value has an uncertainty of ±1 repeat unit.

should be located upstream of the glutamic acid. An 18-nucleotide degenerate DNA probe (with lysine or arginine upstream of glutamic acid) was designed relying on the codon usage previously described for the cloned urease gene of *U. urealyticum* (1). The nucleotide sequence is 5′-AA(or G)A(or T)-GAA-CAA-CCA-GCA(or T)-GGT-3′. This probe was used to locate the gene of interest in *U. urealyticum* chromosomal DNA. Southern blot hybridization showed that the probe hybridized to a single *Eco*RI-*Hin*dIII DNA fragment. This 6-kbp fragment was cloned into plasmid pUC19 (Fig. 1).

Analysis of the MB antigen gene sequence. The gene encoding the MB antigen consists of one large open reading frame of 1,230 bp (Fig. 2A) which encodes 409 amino acid residues, giving a predicted molecular mass of 42,863 Da. The N terminus of the predicted peptide consists of a typical prokaryotic precursor signal peptide followed by a consensus membrane lipoprotein lipid attachment site (cysteine). This putative signal peptide would be 29 amino acids long and would contain a hydrophobic core region (Fig. 2B).

Beginning at bp 451 of the DNA sequence (Fig. 2), there is a region of identical 18-nucleotide repeat units with a deduced amino acid sequence of GKEQPA. This repeated sequence is identical to that of the purified peptide (EQPAG), with a lysine residue upstream, that was used to design the aforementioned 18-nucleotide degenerate probe. To obtain the complete DNA sequence of this region, we made nested DNA fragments that contained various lengths of the repeat region by digesting the linearized plasmid with exonuclease III (18). Our analysis showed the reiterated 18-bp units to be completely uniform. The copy number of the repeats was calculated on the basis of agarose gel electrophoretic analysis of restriction endonuclease fragments containing the repeat region. This region contained an estimated 42.5 repeats encompassing 765 bp of the gene. Genetic data base searches for both nucleic acid and derived amino acid homologies showed no significant matches for these repeats. Analysis of the deduced amino acid sequence of the MB antigen predicted the carboxy repeat region to be hydrophilic, surface localized, and antigenic.

***U. urealyticum* serovar 3 contains a single copy of the MB antigen gene.** We searched the *U. urealyticum* serovar 3 genome for other copies of the MB antigen gene. Southern blot analysis of *U. urealyticum* serovar 3 genomic DNA digested with different restriction enzymes is shown in Fig. 3. That blot was probed with a *U. urealyticum* serovar 3 MB antigen gene DNA fragment that was generated by PCR (with primers UMS44 and UMA427, i.e., primers II and VI, respectively, in Fig. 1) and nick translation. This probe contained 427 bp and hybridized to the 5′ end of the MB antigen gene. The 427-bp MB antigen gene probe hybridized to a single band in the lanes digested with enzymes *Mva*I, *Hin*fI, and *Ssp*I. Lanes containing *U. urealyticum* serovar 3 genomic DNA digested with *Pvu*II and *Rsa*I showed hybridization to two bands; however, because both of those enzymes cut within the MB antigen gene, the presence of two bands was consistent with the presence of a single copy of the MB antigen gene. In addition, with an 18-nucleotide probe (one repeat unit) and by single and double digestion with different combinations of restriction enzymes *Eco*RI, *Hin*dIII, *Bam*HI, and *Xba*I, Southern blot analysis showed a single hybridization band (data not shown), also indicating that there is only one copy of the repeat region.

Immunological detection of in vitro translation products from the MB antigen gene and synthetic peptides. Confirmation that our cloned gene was indeed the MB antigen gene

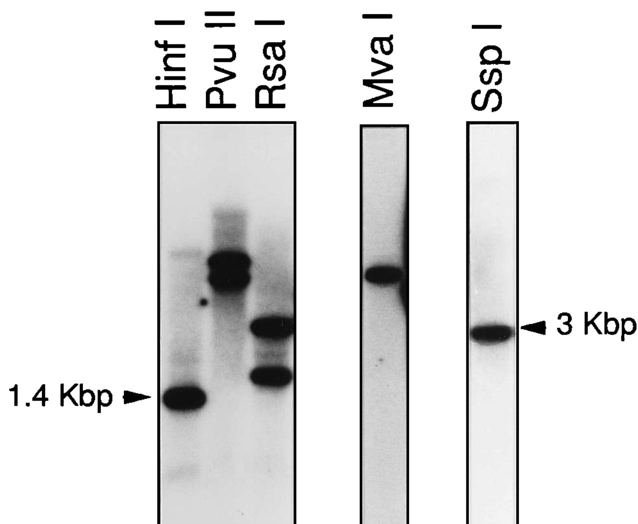

FIG. 3. Southern blot hybridization indicating that the MB antigen gene exists as a single copy in the chromosome. Genomic DNA of the serovar 3 reference strain was digested with restriction enzymes as indicated. The 383-bp DNA probe used for hybridization was the N-terminal nonrepeat region of the gene generated by PCR with primers UMS44 and UMA427. *Pvu*II and *Rsa*I cut within the gene.

required immunological recognition of the expressed recombinant protein with reagents specific for the MB antigen. There was also a conflict between the predicted molecular mass of the MB antigen (39.8 kDa) and its apparent molecular mass as determined by SDS-PAGE (67.8 kDa). To address these questions, the MB antigen gene and MB antigen gene deletions were expressed by using a coupled in vitro transcription-in vitro translation system which used T7 RNA polymerase and mRNA-dependent rabbit reticulocyte lysate. To do this, it was necessary to insert the gene in an expression vector plasmid so that it was behind a T7 RNA polymerase promoter and a consensus eukaryotic translation initiation sequence (13); additionally, we changed the single TGA tryptophan codon to a TGG tryptophan codon. As diagrammed in Fig. 4A, three expression plasmids were constructed: MB, which codes for a 409-amino-acid peptide with a molecular mass of 42,863 Da (including the putative signal peptide); R, the 3′-end 63% of the gene, which contains 42.5 ± 1 repeat units; and ΔMB, which has 39.5 ± 1 repeats deleted, leaving just 3 repeats between the intact amino and carboxy termini.

Aliquots of the in vitro expression reactions, which were metabolically labeled with ^{3}H-labeled amino acids, were analyzed by SDS-PAGE (Fig. 4B). We found that our recombinant MB antigen, with a predicted molecular mass of 42,863 Da, migrated at the same rate as that described for authentic MB (~68 kDa).

On the basis of the amino acid sequence, the most plausible explanation for the MB antigen's aberrant electrophoretic migration rate was the potential for an unusual or denaturation-resistant protein conformation imparted by the high density of prolines and charged residues in the repeat region of the MB antigen. By using a deletion mutant, R, which lacks expression of the nonrepetitive amino region of the molecule and by varying the number of 6-amino-acid repeats in the MB antigen as done in a second deletion mutant, ΔMB (three repeats), we hoped to investigate the contribution of the repeat region to the MB antigen SDS-PAGE migration rate. As shown in Fig. 4B, the R in vitro translation product migrated at an aberrantly slow rate; however, the ΔMB product migrated at its predicted molecular weight. On the basis of these results, the aberrant migration of MB and related proteins in SDS-PAGE can be modified by the presence or absence of the repeat units.

Previous analysis of *U. urealyticum* MB genes from the different serovars (22) indicated that the nonrepetitive N-terminal region of the MB gene product is likely to contain serovar-conserved determinants. Therefore, to test the implied hypothesis that serovar specificity is encoded by the repeat region, we used deletion mutant R, which expresses only the repeat region of the MB antigen gene, for a crude form of epitope mapping. As shown in Fig. 5, lanes 3 and 6, serovar 3-specific MAb 10C6.6 immunoprecipitated both the complete MB antigen and the repeat region fragment of the MB antigen. To determine if epitopes in this repeat region are recognized in vivo, we also tested sera from two patients. On the basis of their *U. urealyticum* serovar 3 reactivity as determined by immunoblotting, the two sera used were designated positive and negative patient sera. As shown in Fig. 5, lane 7, the positive patient serum immunoprecipitated the in vitro-translated repeat region whereas the negative patient serum did not (Fig. 5,

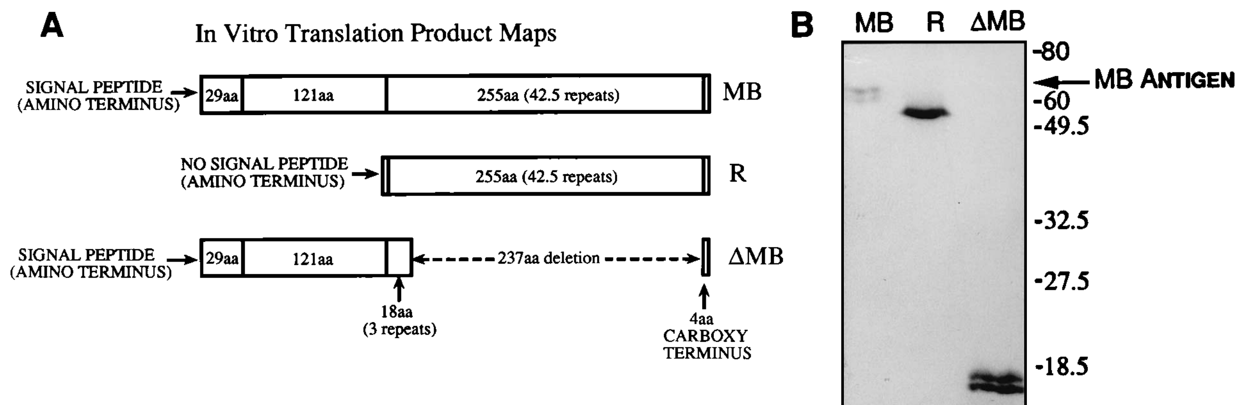

FIG. 4. In vitro translation products from the MB antigen gene and truncated MB antigen genes. In panel A, three plasmids were assayed: MB, which codes for a 409-amino-acid (aa) peptide with a molecular weight of 42,863 (including the signal peptide); R, the carboxy 63% of the gene, which contains 42.5 ± 1 repeat units; and ΔMB, which has 39.5 ± 1 repeats deleted, leaving 3 repeats between the intact amino and carboxy termini. The numbers to the right of the fluorograph shown in panel B mark the locations and sizes, in kilodaltons, of protein molecular size markers. The arrow indicates the size of the authentic *U. urealyticum* serovar 3 MB antigen. The doublets in lanes MB and ΔMB were likely due to inefficient cleavage of signal peptides.

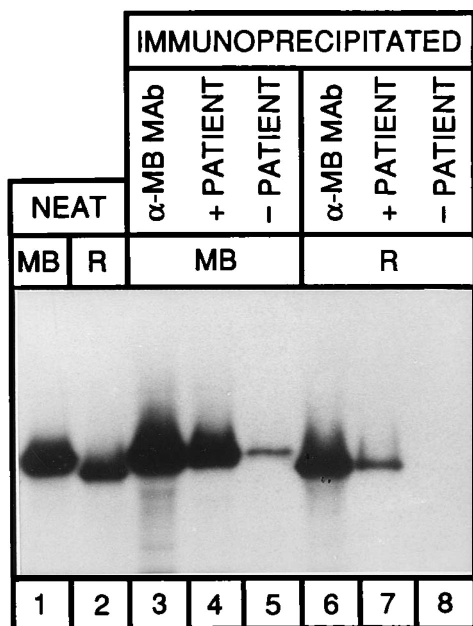

FIG. 5. Immunoprecipitation of in vitro-translated recombinant MB antigen. In vitro translational products were labeled with [^{3}H]leucine (for the whole MB molecule) or [^{3}H]proline (for the repeat region fragment). Samples were resolved by SDS–10% PAGE. The gel was fluorographed and exposed to X-ray film. Two in vitro translation products were analyzed: MB was the in vitro-translated product of the complete MB antigen gene, and R was a deletion mutant that had only the 3′ two-thirds of the gene and therefore contained the repeating sequence only. Lanes 1 and 2 were loaded with in vitro translation products without being immunoprecipitated. The anti-MB MAb used was serovar 3 specific (MAb 10C6.6). The positive and negative patient sera refer to patient serum which was either reactive or not reactive to serovar 3 antigen in our previous immunoblotting tests.

lane 8). These data indicated that patient serum antibodies recognized epitopes contained within the serovar-specific repeat sequence.

Although not directly relevant to the task of localizing serotype-specific epitopes within the MB antigen, it was interesting that the negative patient serum which did not react with the serovar 3 antigen in an immunoblot did appear to precipitate the complete in vitro-translated MB antigen. This discrepancy between immunoblot and immunoprecipitation is likely due to the higher sensitivity of the latter approach resulting from antigen concentration and lack of denaturation prior to analysis.

Added support for the immunoreactivity of the repeat region was provided by testing these two patient sera and four previously described MAbs (25) for reactivity to both affinity-purified serovar 3 MB antigen and a synthetic peptide containing three complete repeats (18 amino acids, starting with the glycine residue). As shown in Fig. 6, three MAbs and the positive patient serum reacted to both the purified MB antigen and the synthetic peptide, while one MAb (8B5.2) and the negative patient serum reacted only with the whole MB antigen. This weak reaction of the negative serum with the whole molecule agreed with the in vitro translation experiments described above. Importantly, since this negative serum did not precipitate the in vitro-expressed repeat fragment or react with the synthetic repeats in an ELISA, it is likely that the epitope recognized by this serum is located in the N-terminal nonrepeat region; the region that we predict contains a high density of common sequences among different serovars of the organisms (22). Theoretically, it is possible that the patient had been

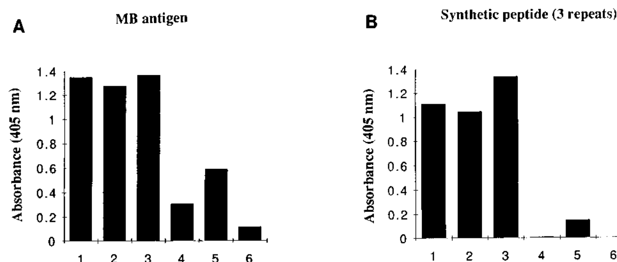

FIG. 6. ELISA results of antibody reactivities to affinity-purified serovar 3 MB antigen (A) and a synthetic peptide containing three repeats (B). Columns 1 to 4 indicate optical density readings for MAbs 10C6.6, 3B1.5, 5B1.1, and 8B5.2, respectively. Columns 5 and 6 are for the positive and negative patient sera, respectively.

infected by other serovars of the organism and had antibodies directed to the common region of the MB antigen. Additional sequencing and epitope mapping of the MB genes from other serovars eventually should allow resolution of this question.

The repeat region is associated with antigen size variation. Four *U. urealyticum* serovar 3 variants were chosen for PCR analysis of MB antigen size variation. These variants were subclones from parent strain 1546, which was isolated from lung tissue of a newborn infant. This parent strain is a mixture of size variants (53 and 68 kDa are the major species) which can be easily separated by subcloning, as previously demonstrated (25). Among the variants isolated, 21 and 1 had the same-size MB antigen, whose molecular weight was lower than that of 16 and 19 (Fig. 7). When PCR primers UMS51 (5′-CTGAGCTATGACATTAGGTGTTACC-3′) and UMA427 (primer VI in Fig. 1; 5′-ACCTGGTTGTGTAGTTTCAAAG TTCAC-3′) were used to amplify most of the 5′ nonrepeat region, the lengths of the PCR products were the same among the four variants. In an analysis of a larger region of these ureaplasma subclone MB antigen genes, we used PCR primer UMS-125 (primer I in Fig. 1; 5′-GTATTTGCAATCTTT ATATGTTTTCG-3′) and aforementioned primer UMA427. This pair defined a DNA fragment of 602 bp covering some of the 5′ upstream region and the entire N-terminal nonrepeat region from positions −150 to 450 on the reference strain. When these primers were used, the lengths of the PCR products for the four variants were also the same. *Dra*I (position shown in Fig. 1) digestion of the PCR products also gave DNA fragments of the same length, indicating that this region is conserved among the variants. When primer UMS-125 (primer I in Fig. 1) was used with primer UMA1213 (primer VII in Fig. 1) to amplify the region from positions −150 to 1242 (including both repeat and nonrepeat regions), the lengths of the PCR products were different among the four variants. Subclones 21 and 1 gave shorter PCR products than did 16 and 19. These data indicated that the lengths of N-terminal nonrepeat regions were conserved among different variants and that those of the C-terminal repeat regions were variable and associated with size variation of this antigen.

Repeat copy number was the only difference between MB antigens from the reference strain and a clinical isolate. A PCR-based method was used to clone the MB antigen gene of clinical isolate 8315. The apparent molecular mass, as determined by SDS-PAGE, of the MB antigen of this isolate (39 kDa) (27) was markedly less than that of the reference strain (67.8 kDa). PCR primers were designed on the basis of the DNA sequence of the serovar 3 reference strain. The PCR product obtained from isolate 8315 with primers UMS-125 (primer I in Fig. 1) and UMA1586 (primer VIII in Fig. 1;

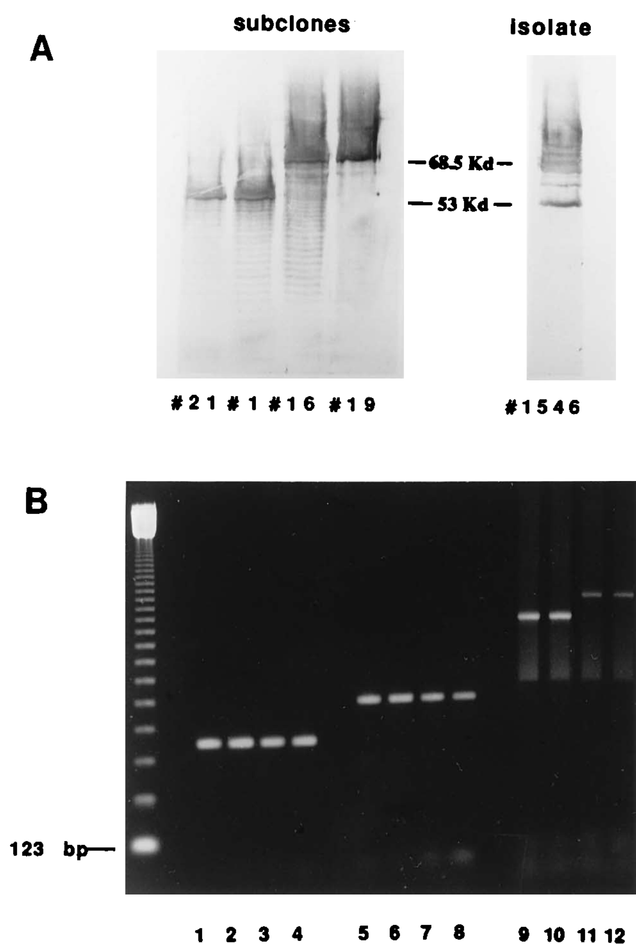

FIG. 7. PCR analysis of antigen size variation. (A) Immunoblot of four subclones and their parent strain, which is a serotype 3 *U. urealyticum* clinical isolate. The blot was developed with MAb 3B1.5. (B) Result of PCR analysis of MB antigen gene of the four variants described in panel A. PCR products were analyzed by agarose gel electrophoresis, and samples were loaded in the same order as the four subclone variants in panel A for lanes 1 to 4, 5 to 8, and 9 to 12. Locations of PCR primers are shown in Fig. 1. The combinations of primers for PCRs were as follows: primers III and VI, lanes 1 to 4; primers I and VI, lanes 5 to 8; primers I and VII, lanes 9 to 12. DNA size markers are on the left side. The space between bands equals 123 bp.

5′-GATAATCATTCATCTTCTCTTAATTGTC-3′) was cloned into plasmid pCRII. This cloned fragment started 150 bp upstream of the MB antigen gene's start codon and ended 383 bp downstream of the gene's stop codon. Two positive clones were chosen for DNA sequencing. The sequence of the first clone was identical to that of the reference strain, except that base 413 was changed from T to C (a silent mutation) and there were only 9.5 repeat units. The sequence of the second clone also was identical to that of the reference strain, except that base 1247 changed from A to G (17 bp downstream of the stop codon) and there were only 10.5 repeats. These transition mutations might exist in the original isolates or could have been introduced during PCR amplification. The finding of two different lengths for the repeat sequence was not surprising, since this isolate was most likely composed of multiple size variants.

DISCUSSION

Accumulating evidence suggests that the *U. urealyticum* MB antigens and the analogous antigens in other mycoplasmas (26–28, 30) are likely to be pivotal mediators in the interaction of mycoplasmas with their host cells. *U. urealyticum* MB surface antigens, which contain both serogroup- and serovar-defining regions, are among the major ureaplasmal antigens recognized by humans during infection (25). Additionally, the MB antigens undergo an extraordinarily high rate of size variation that is phenotypically similar to that documented for several other mycoplasmas (27, 28). These facts show the need for characterization of the epitopes eliciting the potentially protective antibody responses and determination of how these epitopes may be affected by these size changes. Such knowledge would be requisite to understanding the role that these antigens and their associated size variation may play in the success or failure of these organisms as pathogens. To this end, we have cloned and sequenced the gene encoding the *U. urealyticum* serovar 3 MB antigen, and although precise epitopes have not been defined, we have mapped the major immunoreactive and size-variable region to the carboxy two-thirds of the molecule.

There was a striking difference between the apparent molecular mass of the MB antigen determined by SDS-PAGE, 67.8 kDa, and the deduced molecular mass of 42,863 Da based on the predicted amino acid sequence of MB. Our data suggested that the structure of the six-amino-acid repeats could hinder MB denaturation in SDS-PAGE. Among the six amino acids were two charged residues, one positive (lysine) and one negative (glutamic acid), as well as a proline. In concert, the conformational effects of a multitude of beta-turns imparted by the proline and the potential for intra- and intermolecular interactions provided by the charged residues could render the MB protein resistant to denaturation with resultant aberrant electrophoretic migration. Future studies will elucidate the effect of the six-amino-acid repeats on the higher-order conformation of MB. In addition to answering the questions regarding the predicted size versus the observed size of MB, insight into MB conformation may show how *U. urealyticum* presents this molecule to a host.

DNA sequence analysis of the MB gene showed that it encodes a polypeptide with a putative signal peptide and acylation site and a C-terminal six-amino-acid direct repeat region. Analysis of the deduced amino acid sequence of the MB antigen's C-terminal repeat region indicated that the region is probably antigenic and surface exposed. Furthermore, the MAbs which were reactive in our synthetic peptide (repeat sequence) ELISA and which immunoprecipitated the in vitro-translated six-amino-acid repeat region were also previously shown to inhibit the in vitro growth of the organism (25). This inhibitory property of the antibody requires interaction with a surface-exposed epitope. In concert, these data suggest a probable orientation of the MB antigen at the surface of the ureaplasma membrane, i.e., an N-terminal membrane anchor resulting in exposure of the C-terminal repeat region to the microenvironment surrounding the organism.

Thus, we believe that the carboxy domain is most likely to encounter host defenses first and to elicit a dominant antibody response. In fact, our data demonstrated that this relatively simple structure, consisting of direct repeats of only six amino acids, provided a diversity of immunoreactive sequences. Three MAbs, including two serovar-cross-reactive MAbs, 3B1.5 and 5B1.1, and a serovar 3-specific MAb, 10C6.6, reacted with the synthetic repeats, whereas 8B5.2 did not. Previously we showed that the epitopes defined by these MAbs encompass a large part, and in some cases all, of the epitopes recognized by antibodies obtained from *U. urealyticum*-infected humans (25). Since our previous report reflected a high degree of conservation among serovars of the epitope defined

by MAb 8B5.2, this antibody's lack of recognition of the repeat region is not surprising. One might expect a higher degree of conservation of structure-related motifs in the membrane-proximal (anchoring) region of a surface-exposed antigen. Therefore, we expected to find the 8B5.2 epitope within the N-terminal nonrepeat region of our molecule. On the other hand, the less conserved epitopes, defined by MAbs 3B1.5, 5B1.1, and 10C6.6, were all located within the carboxy repeat region, which is likely to be surface exposed. The ability to diversify the regions of an antigen that first encounter a changing microenvironment would be of obvious advantage to a pathogen in establishing a successful infection. Further resolution of the precise amino acids responsible for each epitope within the *U. urealyticum* serovar 3 repeat region and characterization of similar genes in the other serovars will aid the development of specific antibody detection modalities. Additionally, because these are the predominant epitopes recognized by humans, we should be able to begin determining how or if a patient's antibody responses to this antigen affect the microorganism's pathogenic capacity.

PCR analysis of size variants obtained by subcloning indicated that alterations within this repeat region are responsible for the size variation of the antigen. Importantly, the identification of PCR primers and conditions capable of distinguishing size variants should provide the potential to directly analyze specimens from patients without prior growth of the organism in vitro. This is critical because the selective pressures on ureaplasmas cultured in vitro are almost certainly very different from the pressures encountered by organisms surviving in a mammalian host. In conjunction with patient history and disease outcome, the information obtained by PCR evaluation of infected patient tissue should allow eventual determination of whether certain size variants are more frequently associated with a specific pathological manifestation. Definition of the precise size variants present in vivo will also aid in the assessment of whether host defenses can generate pressure resulting in selection of size variants and whether a particular immune response can effectively prevent continued colonization and disease.

The phenomenon of protein size variation resulting from changes in repetitive structures has been seen in surface antigens on both eukaryotic and prokaryotic pathogens. Many surface antigens expressed in different stages of malaria have tandem repeats, and natural antibody responses are directed against epitopes encoded by these repeats (12). For bacterial pathogens, repetitive elements have been observed in the M protein of group A streptococci (11), the C protein alpha antigen of group B streptococci (14), outer membrane protein H.8 on pathogenic neisseriae (29), and a surface protein of *Streptococcus pneumoniae* (31). Most of these proteins are considered to be virulence associated; e.g., the M protein of group A streptococci has antiphagocytic activity (10) and the C protein alpha antigen of group B streptococci is associated with resistance to killing by polymorphonuclear leukocytes and affects the outcome of opsonophagocytic killing of the organism (14). Antigen size variation has also been found in mycoplasmas (28). Our previous studies showed that the difference in V-1 antigen in *Mycoplasma pulmonis*, a genital and respiratory tract pathogen of rodents, is the only difference detectable between virulent and avirulent strains and this difference is associated with differences in disease severity (26, 27). Antigen size variation also is documented for the surface lipoproteins encoded by a cluster of related but divergent genes of *M. hyorhinis*, a swine pathogen (17, 30). Changes within repetitive sequences encoding the surface lipoproteins are responsible for spontaneous size variation. Even though *U. urealyticum* and *M. hyorhinis* are not closely related phylogenetically (15), they may have similar systems leading to surface diversity and complexity. They do, however, differ in at least two features: (i) there is no diversity within the repetitive sequences of the ureaplasma MB antigen, and (ii) unlike that of *M. hyorhinis*, there appears to be only one copy of the MB antigen gene in the ureaplasma chromosome, as well as no significantly similar genes. Perhaps most importantly, the present study localized a major immunoreactive domain of this antigen to a region of epitopes that are directly affected by the observed size variation. The sequence integrity defining these epitopes appears not to be affected; only the number of epitopes is altered. Although the function of this ureaplasma antigen has not been defined, previously described studies (10, 12, 14) with similar antigens of other organisms have at least implied the importance of varying the size of an antigen via repetitive units.

ACKNOWLEDGMENTS

We thank Kevin Dybvig for helpful discussions during the preparation of the manuscript and Elliot Lefkowitz of the UAB Biological Computer Research Facility for assistance with this project.

This work was supported by grant AI 28279 to G.H.C. from the National Institute of Allergy and Infectious Diseases. The UAB Protein Analysis and Peptide Synthesis Core Facility was supported by grant 5P30CA13148.22. Support for synthesis of oligonucleotides for DNA sequencing was obtained through NCI grant CA13148 to the UAB Comprehensive Cancer Center.

REFERENCES

1. **Blanchard, A.** 1990. *Ureaplasma urealyticum* urease genes; use of a UGA tryptophan codon. Mol. Microbiol. **4:**669–676.
2. **Blanchard, A., J. Hentschel, L. Duffy, and K. Baldus.** 1993. Detection of *Ureaplasma urealyticum* by polymerase chain reaction in the urogenital tract of adults, in amniotic fluid, and in the respiratory tract of newborns. Clin. Infect. Dis. **17**(Suppl. 1)**:**S148–S153.
3. **Cassell, G. H., D. T. Crouse, K. B. Waites, P. T. Rudd, and J. K. Davis.** 1988. Does *Ureaplasma urealyticum* cause respiratory disease in the newborn? Pediatr. Infect. Dis. **7:**535–541.
4. **Cassell, G. H., R. O. Davis, K. B. Waites, M. B. Brown, P. A. Marriott, S. Stagno, and J. K. Davis.** 1983. Isolation of *Mycoplasma hominis* and *Ureaplasma urealyticum* from amniotic fluid at 16–20 weeks of gestation: potential effect on outcome of pregnancy. Sex. Transm. Dis. **10:**294–302.
5. **Cassell, G. H., K. B. Waites, D. T. Crouse, P. T. Rudd, K. C. Canupp, S. Stagno, and G. R. Cutter.** 1988. Association of *Ureaplasma urealyticum* infection of the lower respiratory tract with chronic lung disease and death in very low birth weight infants. Lancet **ii:**240–244.
6. **Cassell, G. H., K. B. Waites, R. S. Gibbs, and J. K. Davis.** 1986. Role of *Ureaplasma urealyticum* in amnionitis. Pediatr. Infect. Dis. J. **5:**S247–S252.
7. **Cassell, G. H., K. B. Waites, H. L. Watson, D. T. Crouse, and R. Harasawa.** 1993. *Ureaplasma urealyticum* intrauterine infection: role in prematurity and disease in newborns. Clin. Microbiol. Rev. **6:**69–87.
8. **Chamberlain, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. Anal. Biochem. **98:**132–135.
9. **Davidson, M. K., S. E. Ross, J. R. Lindsey, and G. H. Cassell.** 1988. Colony opacity, hemadsorption, hemolysis, and mitogenicity are not associated with virulence of *Mycoplasma pulmonis*. Infect. Immun. **56:**2169–2173.
10. **Fischetti, V. A.** 1991. Streptococcal M protein. Sci. Am. **264:**32–39.
11. **Hollingshead, S. K., V. A. Fischetti, and J. R. Scott.** 1987. Size variation in group A streptococcal M protein is generated by homologous recombination between intragenic repeats. Mol. Gen. Genet. **207:**196–203.
12. **Kemp, D. J., R. L. Coppel, and R. F. Anders.** 1987. Repetitive proteins and genes of malaria. Annu. Rev. Microbiol. **41:**181–208.
13. **Kozak, M.** 1989. The scanning model for translation: an update. J. Cell Biol. **108:**229–241.
14. **Michel, J. L., L. C. Madoff, K. Olson, D. E. Kling, D. L. Kasper, and F. M. Ausubel.** 1992. Large, identical tandem repeating units in the C protein alpha antigen gene, *bca*, of group B streptococci. Proc. Natl. Acad. Sci. USA **89:** 10060–10064.
15. **Razin, S.** 1992. Mycoplasma taxonomy and ecology, p. 3–22. *In* J. Maniloff, R. N. McElhaney, L. R. Finch, and J. B. Baseman (ed.), Mycoplasmas: molecular biology and pathogenesis. American Society for Microbiology, Washington, D.C.
16. **Robertson, J. A., and G. W. Stemke.** 1982. Expanded serotyping scheme for

Ureaplasma urealyticum strains isolated from humans. J. Clin. Microbiol. **15:**873–878.
17. **Rosengarten, R., and K. S. Wise.** 1990. Phenotypic switching in mycoplasmas: phase variation of diverse surface lipoproteins. Science **247:**315–318.
18. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
19. **Talkington, D. F., M. T. Fallon, H. L. Watson, R. K. Thorp, and G. H. Cassell.** 1989. *Mycoplasma pulmonis* V-1 surface protein variation: occurrence *in vivo* and association with lung lesions. Microb. Pathog. **7:**429–436.
20. **Taylor-Robinson, D., P. M. Furr, and A. D. B. Webster.** 1985. *Ureaplasma urealyticum* causing persistent urethritis in a patient with hypogammaglobulinaemia. Genitourin. Med. **61:**404–408.
21. **Taylor-Robinson, D., P. M. Furr, and A. D. B. Webster.** 1986. *Ureaplasma urealyticum* in the immunocompromised host. Pediatr. Infect. Dis. J. **5:**S236–S238.
22. **Teng, L.-J., X. Zheng, J. I. Glass, H. L. Watson, J. Tsai, and G. H. Cassell.** 1994. *Ureaplasma urealyticum* biovar specificity and diversity are encoded in multiple-banded antigen gene. J. Clin. Microbiol. **32:**1464–1469.
23. **Volgler, L. B., K. B. Waites, P. F. Wright, and G. H. Cassell.** 1985. *Ureaplasma urealyticum* polyarthritis in agammaglobulinemia. Pediatr. Infect. Dis. J. **4:**687–691.
24. **Waites, K. B., P. T. Rudd, D. T. Crouse, K. C. Canupp, K. G. Nelson, C. Ramsey, and G. H. Cassell.** 1988. Chronic *Ureaplasma urealyticum* and *Mycoplasma hominis* infections of the central nervous system in preterm infants. Lancet **ii:**17–22.
25. **Watson, H. L., D. K. Blalock, and G. H. Cassell.** 1990. Variable antigens of *Ureaplasma urealyticum* containing both serovar-specific and serovar-cross-reactive epitopes. Infect. Immun. **58:**3679–3688.
26. **Watson, H. L., D. K. Blalock, and G. H. Cassell.** 1990. Characterization of the variable V-1 antigen of *Mycoplasma pulmonis*, p. 529. *In* G. Stanek, G. H. Cassell, J. G. Tully, and R. F. Whitcomb (ed.), Recent advances in mycoplasmology. Gustav Fischer Verlag, New York.
27. **Watson, H. L., X. Zheng, and G. H. Cassell.** 1993. Structural variations and phenotypic switching of mycoplasmal antigens. Clin. Infect. Dis. **17**(Suppl. 1)**:**S183–S186.
28. **Wise, K. S., D. Yogev, and R. Rosengarten.** 1992. Antigenic variation, p. 473–489. *In* J. Maniloff, R. N. McElhaney, L. R. Finch, and J. B. Baseman (ed.), Mycoplasmas: molecular biology and pathogenesis. American Society for Microbiology, Washington, D.C.
29. **Woods, J. P., S. M. Spinola, S. M. Strobeland, and J. G. Cannon.** 1989. Conserved lipoprotein H.8 of pathogenic *Neisseria* consists entirely of pentapeptide repeats. Mol. Microbiol. **31:**43–48.
30. **Yogev, D., R. Rosengarten, R. Watson-McKown, and K. S. Wise.** 1991. Molecular basis of *Mycoplasma* surface antigenic variation: a novel set of divergent genes undergo spontaneous mutation of periodic coding regions and 5′ regulatory sequences. EMBO J. **10:**4069–4079.
31. **Yother, J., and D. E. Briles.** 1992. Structural properties and evolutionary relationships of PspA, a surface protein of *Streptococcus pneumoniae*, as revealed by sequence analysis. J. Bacteriol. **174:**601–609.
32. **Zheng, X., H. L. Watson, K. B. Waites, and G. H. Cassell.** 1992. Serotype diversity and antigen size variation among invasive isolates of *Ureaplasma urealyticum* from neonates. Infect. Immun. **60:**3472–3474.