Epitope Mapping of the Variable Repetitive Region within the MB Antigen of *Ureaplasma urealyticum*

XIAOTIAN ZHENG,† KEITH LAU, MONICA FRAZIER, GAIL H. CASSELL, AND HAROLD L. WATSON*

> *Department of Microbiology, University of Alabama at Birmingham Schools of Medicine and Dentistry, Birmingham, Alabama 35294*

Received 3 May 1996/Returned for modification 13 June 1996/Accepted 9 August 1996

One of the major surface structures of *Ureaplasma urealyticum* **recognized by antibodies of patients during infection is the MB antigen. Previously, we showed by Western blot (immunoblot) analysis that any one of the anti-MB monoclonal antibodies (MAbs) 3B1.5, 5B1.1, and 10C6.6 could block the binding of patient antibodies to MB. Subsequent DNA sequencing revealed that a unique six-amino-acid direct tandem repeat region composed the carboxy two-thirds of this antigen. In the present study, using antibody-reactive peptide scanning of this repeat region, we demonstrated that the amino acids defining the epitopes for MAbs 3B1.5, 5B1.1, and 10C6.6 are EQP, GK, and KEQPA, respectively. Peptide scanning analysis of an infected patient's serum antibody response showed that the dominant epitope was defined by the sequence PAGK. Mapping of these continuous epitopes revealed overlap between all MAb and patient polyclonal antibody binding sites, thus explaining the ability of a single MAb to apparently block all polyclonal antibody binding sites. We also show that a single amino acid difference in the sequence of the repeats of serovars 3 and 14 accounts for the lack of reactivity with serovar 14 of two of the serovar 3-specific MAbs. Finally, the data demonstrate the need to obtain the sequences of the** *mba* **genes of all serovars before an effective serovar-specific antibody detection method can be developed.**

Ureaplasma urealyticum has been implicated in perinatal morbidity and mortality, including chorioamnionitis, premature birth, and respiratory disease in newborns (1, 2). Additional studies have also shown that *U. urealyticum* is the single most common microorganism isolated from the central nervous system of newborn infants (13, 14). Since disease occurs in only a subpopulation of infected individuals, it has been postulated that only some of the 14 established serovars may be responsible for causing diseases. Earlier studies suggested that a serovar-specific antibody may be required for protection against invasive diseases caused by this organism. Our preliminary studies (18) suggested that invasiveness was not likely to be limited to one or a few particular serovars and may be mediated by an adaptive property inherent in all serovars, i.e., the ability to accommodate a new environment via easily alterable surface antigens. This ability to provide a highly plastic interface between organism and host has been found in many mycoplasmas (16). The change of topography of *U. urealyticum* consequent to changes of its MB antigen may reflect such a property of this organism, thus providing any serovar with a possible immune system avoidance mechanism and the potential ability to alternate between a commensal and a pathogenic state. We have shown that this size-variable antigen, which presents a multiple-banding pattern on an immunoblot, is one of the major antigens recognized during infection by antibodies of patients and that antibodies against this antigen have metabolism inhibition activity in vitro (11, 15). Recently, we cloned and sequenced the *mba* gene of the serovar 3 reference strain and found that the associated antigen contains a region

at the carboxy end of the molecule which consists exclusively of direct tandem repeat units (17). Each repeat unit consists of six amino acids with the sequence of GKEQPA. The corresponding genetic region encompassed 62% of the gene and contained 42 1/2 repeating units. Our results showed that alteration of the repeat copy number accounted for the observed size variation. Using synthetic peptides and coupled in-vitro transcription and translation of the cloned gene, we showed that this repetitive region contained both serovar-specific and cross-reactive epitopes and was recognized by ureaplasma-infected patient antibodies.

In the present study we used antibody-reactive peptide scanning to further elucidate the immunoreactive profile of this carboxy domain (12). The commonly used approach is to produce a series of overlapping peptides covering the entire sequence of interest and determine which regions are reactive with an antibody preparation. However, with a simple sixamino-acid repeat this approach was unlikely to yield the maximum amount of information. Therefore, we chose to synthesize peptides consisting of all unique 4-mers through 18-mers (a total of three direct repeats) to more finely map the epitopes within this region. The complete list of peptides used is shown in Table 1. In all, 75 unique 4- to 18-mer pin-coupled synthetic peptides were prepared by and purchased from Chiron Mimotopes, Clayton, Australia. All peptides received N-terminal acetylation and were covalently bound to pins via their C termini. Reactivities of the peptides with three monoclonal antibodies (MAbs), with serum from one *U. urealyticum*-infected patient, with pooled antisera from three rabbits hyperimmunized with serovar 3, and with the same rabbit anti-serovar 3 sera absorbed with whole serovar 3 organisms (to remove antiserovar 3 antibodies) were determined by enzyme-linked immunosorbent assay with peroxidase conjugates as described in the instructions from the manufacturer. Rabbit antisera were prepared as previously described (15). For preparation of absorbed rabbit antisera, whole serovar 3 organisms were sus-

^{*} Corresponding author. Mailing address: Department of Microbiology, BBRB 618, University of Alabama at Birmingham, Birmingham, AL 35294. Phone: (205) 934-9188. Fax: (205) 975-5965. Electronic mail address: harold_watson@micro.microbio.uab.edu.

[†] Present address: Division of Clinical Microbiology, Hilton 470, Mayo Clinic, Rochester, MN 55905.

^a Peptide numbers are as in Fig. 1.

^b Major epitope reactive with patient serum.

^c GKEQPA is one complete repeat unit.

pended in a 1:10 dilution of rabbit anti-serovar 3 serum at a concentration of 1 mg of organism protein per ml of diluted serum. The suspension was kept at 4° C overnight with rocking. The absorbed samples were centrifuged at $10,000 \times g$ for 25 min at 4° C. The supernatant (absorbed serum) was removed and stored at -20° C until use.

The three MAbs and patient serum tested had been characterized previously by immunoblotting (15). Among the MAbs, 3B1.5 reacts strongly with serovar 3 and reacts weakly with other serovars, 5B1.1 recognizes serovars 3 and 14, and 10C6.6 is serovar 3 specific. All of these MAbs were found to react with a synthetic peptide composed of three repeats in our previous study (17). As determined by immunoblotting, the patient serum (patient B2) showed a strong cross-reaction between serovars 3 and 14 (also a less intense reaction with serovar 1), similar to that seen for MAb 5B1.1. Like the patient serum, the rabbit anti-serovar 3 sera also showed strong crossreactivity between serovars 3 and 14 (data not shown).

The results of the antibody reactivity testing of the pincoupled peptides are shown in Fig. 1. Separate experiments were performed for each antibody tested. A comparison of Fig. 1a and Table 1 (i.e., identification of the longest common sequence in all peptides reactive with a single MAb) indicated that specific amino acid sequences were required for the epitopes for each MAb, as follows: MAb 5B1.1 required GK, MAb 3B1.5 required EQP, and MAb 10C6.6 required KEQPA. It appears unusual for an epitope defined solely by GK, such as that required by MAb 5B1.1, to display the degree of specificity seen for this MAb (15). The MAb had been previously shown to react only with serovars 3 and 14 and not to react with other mycoplasmas, which implies that GK is a rare amino acid combination, although this is unlikely. Therefore, verification of the epitopes for 5B1.1 and 3B1.5 will necessitate the evaluation of unique di- and tripeptides. We anticipate that these twoand three-residue sequences will be found to be critical but not sufficient for the formation of the complete epitopes; conformational or charge properties of neighboring residues will contribute to the total definition of these epitopes. The schematic diagram of the serovar 3 sequence (Fig. 2) indicates that all epitopes detected are present in only two direct repeats and that, as previously demonstrated, binding of any one MAb should result in blocking of the other MAbs.

As seen in Table 1 and Fig. 1a, the major patient antibodyreactive epitope, PAGK, does not correspond to that of any of the MAbs. However, PAGK does overlap epitopes defined by

FIG. 1. Peptide scan of anti-*U. urealyticum* serovar 3 antibodies. Pin-bound peptides listed in Table 1 were reacted with antibodies according to the manufacturer's instructions. (a) Reactions with MAbs and patient serum. MAb ascites fluid was diluted 1:2,000 and patient serum was diluted 1:800 before reaction with the pin-bound peptides. The patient serum responses (A_{405} , measured as optical density [O.D.]) to each individual peptide (1 to 75) are shown. The peptides reactive with each MAb are indicated by the bars below the patient scan. For this preliminary scan, only the easily detected peak responses (>0.25 nm) are defined for the patient serum. These patient responses are outlined in Table 1 and include responses to peptides 5, 10, 20, 25, 36, 41, 46, 51, 56, and 61. The peptides reactive with MAb 5B1.1 were peptides 1, 5, 6, 7, and 10 to 75 (corresponding to GK); those reactive with MAb 3B1.5 were peptides 2, 3, 7, 8, 9, 13, 14, 15, and 18 to 75 (corresponding to EQP); and those reactive with MAb 10C6.6 were peptides 8, 13, 14, 19, 20, 24 to 26, 29 to 32, and 34 to 75 (corresponding to KEQPA) (see Table 1). Peptide 14, which contains no GK, was positive with MAb 5B1.1, and peptide 22, which contains no EQP, was positive with MAb 3B1.5. These positive results were most likely due to nonspecific antibody interaction with the pins, and with a total of 75 reactions per MAb, these discrepancies do not alter the data interpretation. (b) Individual pin reactions of MAb 10C6.6 (black bars), rabbit anti-serovar 3 antisera (shaded bars), and rabbit anti-serovar 3 antisera absorbed with serovar 3 whole cells (open diamonds).

the three MAbs; therefore, any one of the MAbs could block this patient response as previously demonstrated (15).

Figure 1b shows the individual pin reactions of the rabbit anti-serovar 3 sera, absorbed rabbit anti-serovar 3 sera, and MAb 10C6.6. It is noteworthy that in the peptide scans, the intensity of the signal showed a trend to increase for all antibodies analyzed (Fig. 1) as the number of repeats (number of epitopes) increased; the exception, as expected, was the flat signal for the absorbed rabbit sera, which were depleted of serovar 3 antibodies. The rabbit antiserum profiles did not show any well-resolved individual peaks as did the patient serum profile. However, the rabbit anti-serovar 3 peaks for the shorter peptides (1 to 19) correlate well with the MAb 10C6.6 peaks. As the peptides increase in length, new epitopes are formed and the signal becomes more complex. For example, peptide 21 provides a new dominant epitope (EQPAGKE), which reappears in peptides 26 and 27 and then again in peptides 31 through 33. After this point, the specificity of the signals is less clear, most likely because additional overlapping epitopes are recognized by the polyclonal sera.

If the peak patient response had been directed to the serovar-specific epitope KEQPA, it would have indicated that the

Serotype 14

FIG. 2. Immunoreactive sites of the repeat region in the MB antigen. One repeat unit is GKEQPA for serovar 3 and GKEQQPA for serovar 14. MAb 10C6.6 is serovar 3 specific, MAb 5B1.1 cross-reacts between serovars 3 and 14, and MAb 3B1.5 reacts strongly with serovar 3 and shows weak nonspecific cross-reactions with other serovars. The patient serum response is to an epitope common to serovars 3 and 14. All identified epitopes are either overlapping or juxtaposed. The dashed arrows indicate the disruption of two epitopes in the serovar 14 sequence compared with the serovar 3 sequence.

patient had been exposed to *U. urealyticum* serovar 3. In fact, this epitope is represented by peptide 8 and shows marginal reactivity with the patient antibody, whereas the major patient response is to the epitope PAGK, which may be either specific or cross-reactive. Immunoblot analysis of this patient's serum had previously shown strong reactivity not only with *U. urealyticum* serovar 3 but also with serovar 14 and to a lesser extent serovar 1. Whether this particular patient's peak response was serovar specific cannot be determined until corresponding regions of the other serovars have been sequenced. We have now sequenced the *mba* gene of serovar 14. Previously described PCR primers UMS-125 and UMA1213 were used to amplify the serovar 14 *mba* gene, and then the product was cloned and sequenced by our standard procedures (17). The serovar 14 *mba* gene sequence is identical to that for serovar 3 with the exception of an additional glutamine residue (Q) in each repeat unit (Fig. 2). It can now be seen (Fig. 2) that the epitope recognized by patient serum, PAGK, is common to serovars 3 and 14. It can also be seen that MAbs 10C6.6 and 3B1.5 should not recognize serovar 14 and that one of the major rabbit epitopes discussed above, EQPAGKE, appears not to be found in serovar 14.

Sequence analysis was performed with the Genetics Computer Group program (University of Wisconsin—Madison) on a VAX-VMS computer at the University of Alabama at Birmingham. Interestingly, the antigenic index (AI) of the MB antigen amino acid sequence (Fig. 3) calculated as described by Jameson and Wolf (6) (Table 2) predicts that sequences containing the residues PAG should be the most antigenic. Both the patient (PAGK) and rabbit (EQPAGKE) major epi-

1	MKLLKNKKFW	AMTLGVTLVG	AGIVAIAASC	SNSTVKSKLS	NOFAKSTDGK
51.	SFYAVYEIEN	FKDLSNDDKK	SLSNIEFNAA	LTSAENKTES	TLEKGHLVGE
101	KIYVKLPREP		KPNEOLTIIS KSGLIKTSGL	LISDNLNYOT	BKVNFETTOP
151			GKEQPAGKEQ PAGKEQPAGK EQPAGKEQPA GKEQPAGKEQ		PAGKEQPAGK
	201 EQPAGKEQPA GKEQPAGKEQ PAGKEQPAGK EQPAGKEQPA GKEQPAGKEQ				
	251 PAGKEQPAGK EQPAGKEQPA GKEQPAGKEQ PAGKEQPAGK EQPAGKEQPA				
	301 GKEQPAGKEQ PAGKEQPAGK EQPAGKEQPA GKEQPAGKEQ PAGKEQPAGK				
351			EQPAGKEQPA GKEQPAGKEQ PAGKEQPAGK EQPAGKEQPA GKEQPAGKEQ		
401	PAGKETTGK*				

FIG. 3. Complete amino acid sequence of the *U. urealyticum* serovar 3 *mba* gene product. Calculation of the AI for this sequence was based on the method of Jameson and Wolf (6). The repetitive region begins at residue 151 and extends to residue 405. A symmetric distribution of antigenic sites is predicted for the repeat region. The highest AI is indicated by the underline at alternating PAG residues, beginning at residue 161.

topes contain these residues. It is also predicted (Fig. 3) that only every other PAG will have this property because alternating PAG residues are involved in loops (turns likely to be more antigenic, residues 161 to 163) or the less antigenic helices (residues 153 to 160). Residues 152 and 164 represent random coil. If the specific epitopes with the sequence EQPAGKE (loop flanked by alpha helices), defined by the rabbit antisera, are separately expressed along the repeat region of this molecule, then each individual copy of this epitope would be separated from the next by the less antigenic sequence QPAGK (only alpha helices).

The sequence of the native carboxy terminus deduced from the *mba* gene sequence of serovar 3 (17), including the terminal 2 1/2 repeat units, is GKEQPAGKEQPAGKETTGK. Careful examination of Fig. 1 and Table 1 reveals that the peak patient response was to the epitope defined by PAGK only if it occurred at the C terminus of the peptide. If PAGK was followed by E, as always occurs in the native protein, it was not recognized, implying that this particular subset of the patient's polyclonal antibodies would never recognize the native *U. urealyticum* serovar 3 molecule. The only exception to this was the

TABLE 2. Parameters used to calculate the AI

Residue position ^a	aa^b	Hydro- philicity ^{c}	Surface probability ^d	Chain flexibility ^e	Secondary structure f	AI ^g
151	G	2.443	3.644	1.111	т	1.300
152^h	K	2.571	3.253	1.111		0.900
153	Е	1.814	2.126	1.106	h	0.900
154	О	1.643	2.126	1.103	h	0.900
155	P	2.143	2.126	1.089	h	0.900
156	A	2.086	2.126	1.078	h	0.900
157	G	2.086	2.126	1.079	h	0.900
158	K	1.814	2.126	1.085	h	0.900
159	E	1.329	2.126	1.093	h	0.900
160	О	1.643	2.126	1.097	h	0.900
161	P	2.143	2.126	1.089	t	1.100
162	A	2.086	2.126	1.078	t	1.100
163	G	2.086	2.126	1.079	t	1.100
164	K	1.814	2.126	1.085		0.900
165	Е	1.329	2.126	1.093	h	0.900
166	О	1.643	2.126	1.097	h	0.900

^a Residue position in primary sequence.

^b aa, amino acid.

^c Determined by the method of Kyte and Doolittle (9).

^d Determined by the method of Emini et al. (5).

^e Determined by the method of Karplus and Schulz (8).

 f Predicted by the method of Chou and Fasman (3). T, beta turn; h or t, weak

secondary structure.
^{*g*} Determined by the method of Jameson and Wolf (6).

h The pattern between this residue and residue 163 (inclusive) is repeated to residue 403.

lack of reactivity with the C-terminal PAGK of peptide 15. This observation may be explained by a previous report (4, 10) demonstrating that the presence of a glutamic or aspartic acid residue at the amino terminus of a short peptide (two to seven residues) decreases antibody binding because of the ionic nature of these amino acids. A potential scenario for this patient is that the epitope PAGK may have resulted from partial processing of the *U. urealyticum* serovar 3 MB antigen and that this altered antigen actually elicited the antibody response. This processing could be a result of the normal defenses of the infected host or a consequence of organism-directed self processing of a premature protein, as, for example, in certain viruses (7). Alternatively, this population of antibodies may have been elicited by infection with another serovar in which the sequence PAGK is expressed at the carboxy terminus of the native protein.

The above discussions reveal the difficulty one encounters when trying to detect a serovar-specific polyclonal antibody response even if purified antigen containing serovar-specific epitopes and its amino acid sequence are available. These data also demonstrate the usefulness of one of our goals: once sequence data have been obtained for the variable region of all serovars, we should be able to effectively characterize a patient antibody response to the *U. urealyticum* MB antigens. Sequence comparisons of the different serovars will allow the selection of nonredundant peptides among the serovars to be used for serovar-specific antibody detection. This approach will allow us to discern differences in populations of patients with respect to responses to specific serovars as well as to specific MB epitopes within a serovar and to determine if there is any correlation with invasive versus noninvasive infections.

The serovar 14 sequence obtained in this study was assigned GenBank accession number U50462.

This work was supported by PHS grant AI 28279 from the National Institutes of Allergy and Infectious Diseases. MAbs were produced by the Hybridoma Core Facility of the Multipurpose Arthritis Center at the University of Alabama at Birmingham, which is supported by PHS grant 5P 60AR20614.

REFERENCES

1. **Cassell, G. H., K. B. Waites, D. T. Crouse, P. T. Rudd, K. C. Kanupp, 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:**240–245.

- 2. **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.
- 3. **Chou, P. Y., and G. D. Fasman.** 1978. Prediction of the secondary structure of proteins from their amino acid sequence. Adv. Enzymol. Relat. Areas Mol. Biol. **47:**45–148.
- 4. **Day, E. D.** 1990. The immunochemistry of some non-viral multiepitopic antigens, p. 521–602. *In* E. D. Day (ed.), Advanced immunochemistry, 2nd ed. Wiley-Liss, Inc., New York.
- 5. **Emini, E. A., J. V. Hughes, D. S. Perlow, and J. Boger.** 1985. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J. Virol. **55:**836–839.
- 6. **Jameson, B. A., and H. Wolf.** 1988. Predicting antigenicity from protein primary structure: a new algorithm for the prediction of antigenic sites. Comput. Appl. Biosci. **4:**181–186.
- 7. **Kaplan, A. H., M. Manchester, and R. Swanstrom.** 1994. The activity of the protease of human immunodeficiency virus type 1 is initiated at the membrane of infected cells before the release of viral proteins and is required for release to occur with maximum efficiency. J. Virol. **68:**6782–6786.
- 8. **Karplus, P. A., and G. E. Schulz.** 1985. Prediction of chain flexibility in proteins. Naturwissenschaften **72:**212–213.
- 9. **Kyte, J., and R. F. Doolittle.** 1983. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157:**105–132.
- 10. **Shi, P. T., J. P. Riehm, P. E. E. Todd, and S. J. Leach.** 1984. The antigenicity of myoglobin-related peptides synthesised on polyacrylamide and polystyrene resin supports. Mol. Immunol. **21:**489–496.
- 11. **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.
- 12. **Tribbick, G., B. Triantafyllou, R. Lauricella, S. J. Rodda, T. J. Mason, and H. M. Geysen.** 1991. Systematic fractionation of serum antibodies using multiple antigen homologous peptides as affinity ligands. J. Immunol. Methods **139:**155–166.
- 13. **Waites, K. B., D. T. Crouse, and G. H. Cassell.** 1990. Ureaplasma and mycoplasma CNS infections in newborn babies. Lancet **335:**658–659.
- 14. **Waites, K. B., L. B. Duffy, D. T. Crouse, M. E. Dworsky, M. J. Strange, K. G. Nelson, and G. H. Cassell.** 1990. Mycoplasmal infections of cerebrospinal fluid in newborn infants from a community hospital population. Pediatr. Infect. Dis. J. **9:**241–245.
- 15. **Watson, H. L., D. K. Blalock, and G. H. Cassell.** 1990. Variable antigens of *Ureaplasma urealyticum* containing both serovar-specific and serovar-crossreactive epitopes. Infect. Immun. **58:**3679–3688.
- 16. **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.
- 17. **Zheng, X., L.-J. Teng, H. L. Watson, J. I. Glass, A. Blanchard, and G. H. Cassell.** 1995. Small repeating units within the *Ureaplasma urealyticum* MB antigen gene encode serovar specificity and are associated with antigen size variation. Infect. Immun. **63:**891–898.
- 18. **Zheng, X., H. L. Watson, K. B. Waites, and G. H. Cassell.** 1992. Serotype diversity and antigen variation among invasive isolates of *Ureaplasma urealyticum* from neonates. Infect. Immun. **60:**3472–3474.