Expression, Structure, and Location of Epitopes of the Major Surface Glycoprotein of Pneumocystis carinii f. sp. carinii

MICHAEL J. LINKE,1,2* SUSAN M. SUNKIN,3 RYAN P. ANDREWS,1 JAMES R. STRINGER,3 AND PETER D. WALZER1,2

Cincinnati Veterans Affairs Medical Center,1 Department of Internal Medicine, University of Cincinnati College of Medicine,2 and Department of Molecular Genetics, Biochemistry, and Microbiology, University of Cincinnati College of Medicine,3 Cincinnati, Ohio

Received 29 May 1997/Returned for modification 31 July 1997/Accepted 2 October 1997

The major surface glycoprotein (MSG) of Pneumocystis carinii f. sp. carinii consists of a heterogeneous family of proteins that are encoded by approximately 100 unique genes. A genomic expression library was screened with a panel of MSG-specific monoclonal antibodies (MAbs) to identify conserved and rare epitopes. All of the antibodies reacted with epitopes that are encoded within the 5′ end of MSG. The results from the expression screening identified antibodies that recognize highly conserved, moderately conserved, and rare epitopes. Four MAbs (MAbs RA-F1, RA-E7, RA-G10, and RB-E3) reacted with a maltose binding protein–MSG-B fusion protein (MBP-MSG-B41-106E) by immunoblotting and enzyme-linked immunosorbent assay. Three of the MAbs (MAbs RA-F1, RA-G10, and RA-E7) reacted with the same continuous epitope that was localized to amino acids 278 to 290 of MSG-B. Comparison of the sequence of the RA-F1-, RA-G10-, and RA-E7-reactive epitope to the deduced amino acid sequences of multiple MSGs demonstrated that it is highly conserved. The reactivity of RB-E3 with MSG-B was shown to be dependent on amino acids 184 to 192, which may comprise a portion of a discontinuous epitope.

Pneumocystis carinii f. sp. hominis is an important cause of pneumonia in patients with human immunodeficiency virus infection, cancer, and organ transplantation and in other immunocompromised hosts. The host factors that predispose individuals to the development of P. carinii f. sp. hominis pneumonia involve impaired cellular and humoral immunity; however, the specific immune defects are poorly understood. All forms of P. carinii that have been examined contain a 95,000- to 140,000-kDa glycoprotein, termed either the major surface glycoprotein (MSG) or glycoprotein A (gpA), which plays an important role in the immunobiology of the organism (12, 17, 20, 34, 42, 50). MSG is highly immunogenic and stimulates the production of various cytokines, contains protective B- and T-cell epitopes, and facilitates the interaction of P. carinii with host cells (9, 30, 33, 43, 54).

MSG actually consists of a heterogeneous family of proteins encoded by about 100 genes. DNA sequence analysis of MSG cDNAs and/or genes has demonstrated that the gene family encodes many distinct MSG isoforms, which are very similar in size but whose sequences vary (10, 18, 21, 26, 48). Transcription of MSG genes appears to occur only within a single telomeric expression site (7, 39, 49). The isolation of multiple unique MSG cDNAs from a single P. carinii f. sp. carinii-infected rat lung suggests that multiple MSG mRNAs and proteins may be present at the same time within a population of P. carinii f. sp. carinii (26). The ability to alter the expression of different MSG molecules raises the possibility of antigenic variation occurring in P. carinii f. sp. carinii.

MSG appears to be involved in both the cellular and the humoral immune responses of the host to P. carinii f. sp. carinii. Spleen cells isolated from rats environmentally exposed to P. carinii f. sp. carinii proliferate in response to MSG, and sera from these animals contain MSG-specific antibodies (44). MSG is also recognized by humans exposed to P. carinii f. sp. hominis. MSG-specific antibodies have been detected in approximately 30 to 70% of human serum specimens that contain P. carinii f. sp. hominis antibodies (28, 31). Animals generate a vigorous immune response against MSG upon immunization with purified antigens or P. carinii f. sp. carinii preparations (7). Passive immunotherapy with MSG-specific monoclonal antibodies (MAbs) and/or hyperimmune polyclonal antiserum has also been shown to modulate P. carinii infection in ferrets, rats, and mice with severe combined immunodeficiency, supporting the significance of MSG in the host response to P. carinii (13, 14, 35). The involvement of MSG in the host immune response suggests that immunization with it may provide protection against P. carinii infection. Because multiple isoforms of MSG can be expressed within a population of P. carinii, it will be important to identify immunoreactive regions common to all MSGs.

Several groups have described the production and characterization of P. carinii-specific MAbs (11, 16, 19, 23, 24, 27, 29). MSG-specific MAbs have also been used to distinguish rat P. carinii f. sp. carinii and P. carinii f. sp. ratii strains at a phenotypic level (47). A MAb specific to MSG purified from P. carinii f. sp. hominis has been used in the development of a new method of diagnosis of P. carinii f. sp. hominis pneumonia by radioimmunodetection (15).
neally 3 days prior to fusion. Spleen cells from the mice were fused to a myeloma hominis.

The pellet was washed twice, and the organisms were recovered from infected P. carinii. 0.85% ammonium chloride to lyse the erythrocytes. The pellet was washed twice, P. carinii used to produce the MAbs used in this study: (i) P. carinii variant; the organism from human beings will be referred to as P. carinii hominis. f. sp.

Pneumocystis f. sp. will be used in this report (32). Provisional tripartite names denoting the mam-

The MAbs will greatly improve their usefulness in the analysis of different MSG isoforms.

### MATERIALS AND METHODS

**Pneumocystis nomenclature.** The Pneumocystis nomenclature proposed at the 3rd International Workshop on Pneumocystis in Cleveland, Ohio, in June 1994 will be used in this report (32). Provisional tripartite names denoting the mammal of origin were given to the organism populations isolated from various mammalian hosts. Pneumocystis isolated from rats will be referred to as either P. carinii f. sp. carinii, designated the prototype, or P. carinii f. sp. rats, designated a variant; the organism from human beings will be referred to as P. carinii f. sp. hominis.

**Source of organisms.** P. carinii f. sp. carinii pneumonia was induced by corticosteroid treatment of rats, and the organisms were recovered from infected lungs and were quantitated as described previously (4). Briefly, infected lungs were removed en bloc, minced, and homogenized. The homogenate was centrifuged at 1,000 × g for 10 min at 4°C, and the resulting pellet was treated with 0.85% ammonium chloride to lyse the erythrocytes. The pellet was washed twice, and the organisms were resuspended in phosphate-buffered saline.

**Source of antibodies.** MSG-specific MAbs were produced and characterized as described previously (27). Three different P. carinii antigen preparations were used to produce the MAbs used in this study: (i) P. carinii f. sp. carinii, (ii) MSG purified from P. carinii f. sp. carinii, and (iii) detergent-solubilized P. carinii f. sp. hominis (Table 1). Cells were immobilized three times subcutaneously with the antigen preparations, and antigen preparations were then injected intraperito-

### TABLE 1. Characterization of MSG-specific MAbs and frequency of epitope expression

<table>
<thead>
<tr>
<th>MAb</th>
<th>Immunogen*</th>
<th>P. carinii f. sp. carinii</th>
<th>P. carinii f. sp. hominis</th>
<th>No. of positive plaques</th>
<th>MAb Immunogen</th>
<th>P. carinii f. sp. carinii</th>
<th>P. carinii f. sp. hominis</th>
<th>No. of positive plaques</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA-F1</td>
<td>Rat MSG</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>738</td>
<td>RA-F1</td>
<td>Rat MSG</td>
<td>+</td>
</tr>
<tr>
<td>RA-E3</td>
<td>Rat PC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>286</td>
<td>RA-E3</td>
<td>Rat PC</td>
<td>+</td>
</tr>
<tr>
<td>RA-C6</td>
<td>Rat MSG</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>46</td>
<td>RA-C7</td>
<td>Rat MSG</td>
<td>+</td>
</tr>
<tr>
<td>RB-C8</td>
<td>Rat PC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>112</td>
<td>RA-C1</td>
<td>Rat MSG</td>
<td>+</td>
</tr>
<tr>
<td>RB-2F</td>
<td>Rat PC</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>307</td>
<td>RA-C11</td>
<td>Rat MSG</td>
<td>+</td>
</tr>
<tr>
<td>HB-G6</td>
<td>Human PC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>3</td>
<td>RA-F8</td>
<td>Rat PC</td>
<td>+</td>
</tr>
<tr>
<td>RA-C7</td>
<td>Rat MSG</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>23</td>
<td>RA-G10</td>
<td>Rat MSG</td>
<td>+</td>
</tr>
<tr>
<td>RA-C1</td>
<td>Rat MSG</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>55</td>
<td>RA-E7</td>
<td>Rat MSG</td>
<td>+</td>
</tr>
<tr>
<td>RA-C11</td>
<td>Rat MSG</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>6</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA-F8</td>
<td>Rat PC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>7</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA-G10</td>
<td>Rat MSG</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>ND</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA-E7</td>
<td>Rat MSG</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>ND</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Rat MSG, P. carinii f. sp. carinii MSG purified by electrophoresis from an SDS-polyacrylamide gel; Rat PC, SDS-solubilized P. carinii f. sp. carinii; Human PC, SDS-solubilized P. carinii f. sp. hominis.

**ND,** determined by immunoblotting reactivity.

**ND,** not done.

### TABLE 2. Primers used to amplify target regions of MSG-B for cloning and expression in pMAL-c2

<table>
<thead>
<tr>
<th>Primer</th>
<th>Location (nucleotides)</th>
<th>Sequence</th>
<th>Restriction site</th>
</tr>
</thead>
<tbody>
<tr>
<td>msg1a</td>
<td>1–20</td>
<td>AATTGATGGCCAGCGCCCGTTAAGAGG</td>
<td>MunI</td>
</tr>
<tr>
<td>msg375a</td>
<td>370–391</td>
<td>GAGTCCTGCTGCAAGTTGAGGG</td>
<td>BamHI</td>
</tr>
<tr>
<td>msg515a</td>
<td>497–520</td>
<td>GCCATGTGAAGCCTGAGAAGGGG</td>
<td>HindIII</td>
</tr>
<tr>
<td>msg552a</td>
<td>538–555</td>
<td>TTCAGACTCAGTGATAGGGCAAGAGG</td>
<td>XbaI</td>
</tr>
<tr>
<td>msg579a</td>
<td>579–594</td>
<td>GAGCTCGGTGGGAGATTTGCCG</td>
<td>HindIII</td>
</tr>
<tr>
<td>msg586a</td>
<td>575–597</td>
<td>GGCGCGCAAGCTTCAACCAATT</td>
<td>HindIII</td>
</tr>
<tr>
<td>msg834a</td>
<td>819–835</td>
<td>TTCAGACTCCTCTTCCTGGCCT</td>
<td>HindIII</td>
</tr>
<tr>
<td>msg1049a</td>
<td>1034–1052</td>
<td>TTTCAGACTCTCCGATGTTGTTT</td>
<td>HindIII</td>
</tr>
<tr>
<td>msg3801a</td>
<td>3762–3801</td>
<td>ACGTCAACTATGTCATCATGATTCAATAACACGTCCTGATCAATG</td>
<td>MunI</td>
</tr>
</tbody>
</table>

* Relative to MSB-B sequence.

**Nucleotides in boldface denote restriction enzyme site.**
were blocked with ELISAmate bovine serum albumin blocking reagent for 2 h.

Temperature. The wells were washed twice with ELISAmate wash buffer and

performed as described previously (25, 45).

SDS-polyacrylamide gel to a nitrocellulose membrane, and immunoblotting was

10% (125 mM Tris-HCl [pH 6.8], 4.0% sodium dodecyl sulfate (SDS), 20% glycerol,

supernatants by using an amylose resin column (New England Biolabs) according

restriction enzymes.

and ligated into the polylinker of pMAL-c2 digested with the corresponding

PCR cloning vector, pGEM-T (Promega). Ligation and transformation of

pMAL-c2 was unsuccessful; therefore, these products were initially cloned into a

PBS-gene fragment. The reactivity ( + ) or nonreactivity ( - ) of the plaques produced by each clone is indicated for each of eight MAbs.  

RESULTS

Assessment of MAb epitope prevalence in the MSG family. A panel of MAbs known to react with MSG had been generated previously (27). To estimate the number of MSG isoforms recognized by each of these MAbs, we used a genomic expression library that had been made by insertion of random fragments from the *P. carinii* f. sp. *carinii* genome into *gt11*. A genomic expression library could be used because previous studies had shown that MSG genes lack introns (41). It was determined that 4,000 to 5,000 plaques contain one genome equivalent in this library. However, because the inserts in the library are in random orientations and reading frames with respect to the vector DNA sequence, only one of six MSG gene fragments would be expected to be capable of expressing an MSG epitope. Therefore, 24,000 plaques would be expected to express every unique epitope at least once.

Ten previously described MAbs (MAbs RA-F1, RA-E7, RA-C6, RB-C8, RB-F9, HB-G6, RA-C7, RA-C1, RA-C11, and RB-F8) were each used to screen 300,000 plaques in the expression library (27). Table 1 shows that the number of reactive plaques ranged from three positive plaques with MAb HB-G6 to 738 positive plaques with MAb RA-F1.

The plaque data suggested that MAb RA-F1 recognized an epitope found on many MSGs. To examine this possibility, four RA-F1-reactive clones were chosen for further analysis. After plaque purification, the insert size was determined by EcoRI digestion of the phage DNA. The four phages contained different inserts, which ranged in size from 1.3 kb (clone 2) to 4 kb (clone 1) (Fig. 1). Analysis of the DNA sequence showed that each clone contained a different MSG gene (data not shown).

To further characterize these RA-F1-positive phages, their reactivities were assessed with the 11 other MAbs. Four of these 11 MAbs (MAbs RA-C7, HB-G6, RA-C11, and RB-F8) did not react with any of the clones, which is consistent with the fact that each of these four MAbs recognized few plaques in the *gt11* library screen (fewer than 23 plaques) (Table 1). The remaining seven MAbs each reacted with at least one RA-F1-positive phage (Fig. 1). Three MAbs (MAbs RA-F1, RA-E7, and RA-G10) reacted with all four clones, suggesting that they recognize the same epitope, and subsequent studies (described below) showed that this was the case. The reactivities of MAbs RA-F1, RA-E7, RA-G10, RA-C1, RA-C6, and RB-C8 with clone 2 indicates that the epitopes for these MAbs are con-
tained within the region from amino acids 227 to 660. The location of the epitopes for MAbs RB-E3 and RB-F9 could be localized to the first 600 amino acids of MSG on the basis of their reactivities with clone 3.

Localization of two MSG-specific MAb epitopes. To facilitate identification of the MAb epitopes, the msg-b isoform was produced in a bacterial expression system as described in the Materials and Methods. The MBPMSG-B fusion protein (MBPMSG-B41–1065) was assayed by immunoblot analysis for its reactivity to the 12 MAbs (data not shown), and it was found to be reactive with MAbs RA-F1, RA-E7, RA-G10, and RB-E3 (Fig. 1).

An epitope mapping strategy was designed to monitor the loss of reactivity of the MAbs with truncated forms of the MBPMSG-B fusion protein. MAb RA-E7 was chosen to represent the RA-F1, RA-E7, and RA-G10 group because of the availability of sufficient quantities of this MAb for epitope mapping studies. RB-E3 was selected because it appeared to recognize a unique epitope. The regions of MSG-B recognized by MAbs RA-E7 and RB-E3 were initially localized by expression of truncated MBPMSG-B fusion proteins produced by digestion of pMAL/msg-b124–3366 with ExoIII. MAbs RA-E7 and RB-E3 remained reactive with the two shortest ExoIII-generated truncations that stopped at MSG-B amino acids 563 (MBPMSG-B41–563) and 373 (MBPMSG-B41–373) (Fig. 2 and 3, lanes 3 and 4). These results indicated that the epitopes were...
contained in the region from amino acids 41 to 373. Further attempts to produce shorter proteins by ExoIII digestion were unsuccessful.

To identify the epitopes recognized by MAbs RB-E3 and RA-E7, MBPMSG-B fusion proteins that covered amino acids 41 to 373 were prepared by amplifying the target regions of msg-b by PCR and cloning the products into pMAL-c2. Evaluation of the fusion proteins expressed from the PCR-generated pMAL/msg-b constructs by SDS-PAGE (Fig. 2A and 3A) and immunoblotting with polyclonal antisera against MBP (Fig. 2B and 3B) demonstrated that they were of the predicted size.

The epitope for RA-E7 was mapped to MSG-B amino acids 279 to 290 by monitoring the loss of MAb reactivity of the fusion proteins by immunoblotting (Fig. 2C) and ELISA (Fig. 2D). The epitope was deduced from variations in reactivity between MBPMSG_{41-278} and MBPMSG_{41-290}. The MBPMSG_{41-278} construct contains MSG-B amino acids 41 to 278, and MBPMSG_{41-290} covers amino acids 41 to 290. The loss of reactivity of MAb RA-E7 with MBPMSG_{41-278} identifies the reactive amino acids as amino acids 279 to 290 (Fig. 2E). A construct with a large 5' deletion (MBPMSG_{193-587}) also reacted with MAb RA-E7, indicating that the upstream portion of MSG-B from amino acids 41 to 193 is not required for recognition by MAb RA-E7. This finding was also supported by the reactivity of phage clone 2 with MAb RA-E7; clone 2 contained MSG amino acids 227 to 660 (Fig. 1). MAbs RA-G10 and RA-F1 demonstrated reactivity patterns with the truncated fusion pro-
The identification of conserved and variable epitopes on MSG molecules provides a method of studying the expression of different isoforms through the use of either polyclonal antisera produced against peptides or MSG-specific MAbs. Previously, three different MSG variants were identified in a single lobe of an infected lung with epitope-specific polyclonal antisera by immunohistochemistry (1). In another study two MSG-specific MAbs (MAbs RA-C6 and RA-C11) were able to identify antigenic differences between genetically distinct P. carinii f. sp. carinii and P. carinii f. sp. ratti populations and within a genetically defined population of P. carinii f. sp. carinii.

Two additional MAbs (MAbs RA-F1 and RA-C7) reacted with all P. carinii f. sp. carinii and P. carinii f. sp. ratti populations examined (47).

In this study MSG-specific MAbs were initially characterized on the basis of the frequency with which their epitopes are encoded within the P. carinii f. sp. carinii genome by screening an expression library made by insertion of randomly sheared genomic P. carinii f. sp. carinii DNA. The MAbs could be separated into four groups on the basis of the number of plaques that each one recognized. MAb RA-F1 appeared to react with a conserved epitope on the basis of its reactivity with 738 plaques. MAbs RB-E3, RB-2F9, and RB-C8 each recognized between 100 and 300 plaques, suggesting that these epitopes are encoded in multiple MSG genes. The three MAbs in the third group (MAbs RA-C6, RA-C7, and RA-C1) reacted with 46, 23, and 55 plaques, respectively. This group of MAbs appeared to recognize a less conserved epitope. The final group of MAbs (MAbs HB-G6, RA-C11, and RB-F8) recognized a rare epitope on the basis of the low number (three to seven) of plaques that they detected.

In addition to characterizing the frequency with which MAbs are expressed in the Pneumocystis genome, analysis of reactive λgt11 phage established that the MAbs react with MSG. Characterization of the MSG-specific MAbs was previously based on their reactivity with a 116,000-molecular-weight glycoprotein that was specific to Pneumocystis (27). The reactivities of the MAbs with MSG were confirmed through DNA sequence analysis of the λgt11 clones. The four reactive clones that were analyzed by DNA sequencing all contained pieces of DNA that were homologous with previously identified msg genes.

The reactivities of four MAbs with recombinant MSG provided a method for mapping their epitopes. The epitope reactive with MAbs RA-F1, RA-G10, and RA-E7 was localized to a highly conserved region in the amino-terminal end of MSG-B and was identified as amino acids 278 to 290. Comparison of the identified epitope with deduced amino acids from previously cloned msg genes, cDNAs, or cloned regions of MSG from reactive λgt11 phage demonstrated that this sequence is highly conserved. The reactivity of MAb RB-E3 was also localized in the amino-terminal portion of MSG-B approximately 100 amino acids upstream from the epitope reactive with MAb RB-E3. Comparison of the amino acid sequence of the epitope reactive with MAb RB-E3 with that of previously cloned msg genes demonstrated that this region is not as well conserved as the epitope reactive with MAbs RA-F1, RA-G10, and RA-E7. Alignment of the amino acids in the RB-E3-reactive epitope regions from a reactive λgt11 phage and a nonreactive λgt11 phage did not reveal the presence of critical residues that would be essential for reactivity.

There are two basic types of epitopes, continuous and dis-
Continuous (3.46). Continuous epitopes consist of short linear sequences of amino acids. Discontinuous epitopes involve distant residues brought together by protein folding. The immunoblottig, ELISA, and sequence comparison data indicate that MSG-B amino acids 279 to 290 represent the epitope reactive with MABs RA-F1, RA-G10, and RA-E7, and deletion of these amino acids clearly results in the loss of MAB reactivity. In addition, the high degree of conservation of the epitope among other MSGs indicates that it is a continuous epitope. The functional identification of the RB-E3-reactive epitope by immunoblottig and ELISA suggests that it is continuous; however, the lack of a correlation of MAB reactivity with the conservation of specific amino acids indicates that additional residues may be required for high-affinity binding of MAB RB-E3. The decreased reactivity of the fusion protein also indicates that additional amino acids removed from the epitope region are required for antibody recognition.

The presence of a cysteine at residue 187 that could be involved in the folding of MSG through the formation of disulfhydryl bonds is noteworthy. The location of cysteine residues is maintained among all MSGs analyzed to date (26, 52). The conservation of the cysteine positions suggests the importance of disulfhydryl bonds in maintaining a conserved higher-order structure that could be critical to the function of MSG. MSG has been implicated in the binding of P. carinii to host cells and molecules. The extracellular MSG domains involved in these interactions have not been identified, and little is known about the position or orientation of MSG within the cell membrane or cell wall of P. carinii. As described previously, MABs RA-E7, RA-G10, RA-F1, RB-E3, RA-C1, RA-C6, RB-C8, and RB-F9 react with the surface of P. carinii i, sp. carinii by immunofluorescence (27). Localization of the MAB reactivity within the first 600 amino acids demonstrates that at least portions of the amino terminus of MSG are exposed on the surface of P. carinii i, sp. carinii. These results suggest that the surface-exposed amino terminus of MSG may also be involved in binding to host proteins.

The presence of hundreds of MSG genes in the genome suggests that P. carinii is capable of undergoing a form of antigenic variation and that the ability to alter this abundant surface protein is critical to its survival. The MABs characterized in these studies will provide useful tools for analyzing the expression of particular MSG isoforms on P. carinii.

ACKNOWLEDGMENTS

This work was supported by the Medical Research Service of the U.S. Department of Veterans Affairs and Public Health Service contract AI 25139 and grant AI 36701 from the National Institutes of Health.

REFERENCES