Multi-Gene Family of Major Surface Glycoproteins of *Pneumocystis carinii*: Full-Size cDNA Cloning and Expression

Kazuhiro KITADA, Miki WADA, and Yoshikazu NAKAMURA*
Department of Tumor Biology, The Institute of Medical Science, The University of Tokyo, P.O. Takanawa, Tokyo 108, Japan
(Received 25 March 1994)

Abstract

The major surface glycoprotein (MSG) of *Pneumocystis carinii* plays a crucial role in the fatal pneumonia caused by this organism in AIDS patients. A cDNA encoding a full-length MSG polypeptide was isolated from a λ phage library of rat-derived *P. carinii* cDNAs. The deduced MSG, referred to as the MSG5 subtype, is a 120,765-Da protein composed of 1,076 amino acids and contains an anchoring hydrophobic sequence at the C-terminus of the protein. Sequence analyses of cloned MSG-cDNAs revealed an MSG-gene family with ~70% protein sequence identity between subtypes. *P. carinii* karyotype hybridization analyses indicated that the MSG gene family members are scattered throughout most of the *P. carinii* chromosomes. These recombinant MSG proteins reacted with the antiserum from *P. carinii*-infected rats, as expected, and antiserum generated against *P. carinii*-infected mice, indicating the existence of common determinants in MSG polypeptides. The family of MSG proteins is rich in cysteine residues and these cysteines are highly conserved in all MSG subtypes regardless of species specificity, suggesting the structural and/or functional importance of these cysteines. The pathobiological significance of the MSG gene family and its sequence diversity in *P. carinii* is discussed.

Key words: *Pneumocystis carinii*; major surface glycoprotein; MSG family; cDNA

1. Introduction

*Pneumocystis carinii* is an opportunistic pathogen that often causes fatal pneumonia in immunosuppressed or immunodeficient patients with conditions or treatments such as AIDS, cancer chemotherapy or organ transplantation. More than 60% of AIDS patients suffer from *P. carinii* pneumonia at some time in the course of the disease.1

*P. carinii* is a eukaryotic microbe that is classified with the ‘protista fungi’ or ‘fungi’ group2-3 and infects many mammalian hosts. *P. carinii* derived from rats has an unusually abundant and highly immunogenic surface glycoprotein(s) named MSG (formerly called gp120 or P115) ranging from 100 to 125 kDa.4-5 The MSG molecules are highly glycosylated with mannose and are composed of several isoelectric variants.5 The ease of generating anti-MSG monoclonal antibodies indicates the strong antigenicity of MSG.5-8 Gigliotti and Hughes9 found that passive immunization with an anti-MSG monoclonal antibody partly protects against the progression of *P. carinii* pneumonia in animal models. Fisher et al.10 demonstrated a specific T-cell response to MSG molecules after immunization and natural infection. Ezekowitz et al.11 showed that uptake of *P. carinii* by alveolar macrophage is mediated by the mannose receptor which interacts with the mannose moiety of MSG. Pottratz et al.12,13 provided evidence that MSG participates in the attachment of the organism via interaction with fibronectin on alveolar epithelia and macrophages. These observations indicate that the MSG antigen plays a crucial role in the pathobiology of *P. carinii*.

MSG-cDNA cloning has been recently performed by three groups. Haidaris et al.14 have reported the partial cloning and sequence of a cDNA encoding a gpA antigen, analogous to MSG, of *P. carinii* from ferrets. Kovacs et al.15 and Wada et al.16 have reported the cloning and sequence of cDNAs encoding the MSG polypeptides of rat *P. carinii*. The latter two groups have found multiple genes encoding MSG proteins of *P. carinii* from rats, while multiple genes have not been observed in *P. carinii* from ferrets. To investigate the structure and pathobiological function of the MSG gene family, we have cloned and analyzed a cDNA encoding a full-length MSG polypeptide of rat-derived *P. carinii*.
2. Materials and Methods

2.1. P. carinii organisms

P. carinii was cultivated in nude rats and the trophozoites were separated from bronchoalveolar lavage essentially as described previously. Briefly, nude rats (F344/NKyo-rnu/rnu, from the Institute of Laboratory Animals, Kyoto University; or Rowett hooded, Central Institute of Experimental Animals, Kawasaki, Japan) were inoculated intranasally with a suspension of passed lung homogenate containing ~10^6 rat P. carinii organisms and were treated with cortisone acetate until the onset of pneumonia. These animals were maintained free from other bacterial and viral infections by routine microbiological monitoring on the sentinel animals. Bronchoalveolar lavage was obtained from heavily infected nude rats which showed typical manifestation of P. carinii pneumonia. The lavage fluid was centrifuged at 100 x g for 5 min, the resulting supernatant was layered on lymphocyte separation medium, LymphoSepal (Immu-no-biological Laboratory, Shizuoka, Japan) and centrifuged at 1000 x g for 30 min. Cells at the interface were collected and diluted with Dulbecco’s MEM (Nissui Seiyaku, Tokyo, Japan). The cells were incubated in cell culture dishes at 37°C for 1 h in a humidified atmosphere with 5% CO2. Non-adherent cells were then collected and washed twice in Dulbecco’s phosphate saline (PBS at pH 7.4). These P. carinii samples were 99.6% pure as judged by microexamination of cytochemically stained cells and were used in all experiments but the genomic DNA analysis, which employed further purified samples (99.96% purity in microexamination) by filtration through Nucleopore membranes at 5 mm porosity (Nucleopore, Pleasanton, CA).

2.2. Antibodies

Polyclonal and monoclonal (MA-2) antibodies against deglycosylated MSG molecules as well as rabbit serum against the recombinant LacZ-MSG1 fusion protein were described previously.

2.3. Western blot analysis

For SDS-PAGE, proteins were solubilized directly in 0.1 M Tris (pH 6.8), 1% SDS and 1% β-mercaptoethanol for 3 min at 95°C, and analyzed on SDS-polyacrylamide gels followed by immunoblotting with antibodies. Immunostaining was performed using peroxidase-conjugated goat, anti-rabbit or anti-mouse IgG (heavy and light chain specific, Amersham). After washing, the membranes were developed in diaminobenzidine (DAB) and H2O2 solution.

2.4. P. carinii libraries and screening

A P. carinii cDNA library made from P. carinii poly(A)+ RNA cloned into the Agt11 expression vector was described previously. The cDNA library was induced with isopropyl-β-d-thiogalactopyranoside (IPTG) and directly screened with polyclonal rabbit antisera against the deglycosylated MSG antigen according to standard methods. Immune-reactive clones were further characterized for reactivity to the monoclonal antibody MA-2 and subjected to DNA sequence analysis. The plaque hybridization of the cDNA library was carried out according to standard methods using a probe proximal to the 5′ terminus of the MSG gene. The probe was prepared by polymerase chain reaction (PCR) of genomic P. carinii DNA using N-terminal (antisense; 5′-GCCAAAGGTGGTTCTC-3′) and C-terminal (sense; 5′-GAGGTAGTTGTTAGGGG-3′) primers toward the flanking regions between MSG genes in tandem clusters. One of the amplified DNAs, a 374-bp segment, was cloned in plasmid pUC118, digested with HindIII and the resulting 205-bp 5′-terminal fragment was used as the hybridization probe. PCR amplification was performed essentially according to standard methods. Briefly, PCR reaction (100 μl) contained 10 mM Tris (pH 9.0), 50 mM KCl, 1.5 mM MgCl2, 0.1% Triton X-100, 200 μM of each deoxyribonucleoside triphosphate, 2.5 U Taq DNA polymerase (Promega, Madison, WI), 100 ng template DNA (bulk P. carinii DNA), and 50 pmol each of sense and antisense primers. Amplification was performed in a ZYMOREACTOR thermal cycler (ATTO, Tokyo, Japan) with a three-step cycling program: 1 min of denaturation at 94°C, followed by 2 min of annealing at 55°C and 2 min of elongation at 72°C, for a total of 30 cycles.

2.5. RNA blot hybridization

RNA was isolated from P. carinii organisms using the RNA extraction kit (Amersham Laboratories, Buckinghamshire, UK) according to the manufacturer’s instructions. Briefly, the organisms were homogenized in guanidinium thiocyanate solution and nucleic acids were precipitated with ethanol from the bulk of ribonucleases. Then, RNA was selectively precipitated using lithium chloride solutions, extracted with phenol/chloroform and precipitated with ethanol. The resulting RNA was subjected to agarose gel electrophoresis in the presence of formaldehyde, and blotted onto HybondTM.N+ filters (Amersham Laboratories) as described by Sambrook et al. The filter was hybridized to the 2.4-kb [32P] MSG1-cDNA probe, washed according to standard procedures, and exposed to X-ray film.
2.6. Other DNA procedures

Double-stranded or single-stranded DNAs were sequenced by the dideoxynucleotide chain-termination method with appropriate synthetic primers and \([\alpha-^{32}P]\) dCTP by the manual method, or with fluorescent universal primers by the automatic DNA sequencer (A.L.F.™ DNA Sequencer; Pharmacia, Uppsala, Sweden). DNA blot hybridization and plaque hybridization were conducted according to standard methods except that hybridization signals were detected by the nonradioisotope method using ECL Direct Nucleic Acid Labelling and Detection Systems (Amersham Laboratories) according to the manufacturer’s instructions.

3. Results

3.1. Full length cDNA structure of MSG

The \(\alpha\)gt11-expression library of \(P.\) carinii cDNA made by the random priming method was immunoscreened using anti-MSG-peptide polyclonal and monoclonal antibodies. These antibodies were raised against deglycosylated MSG molecules to allow recognition of the MSG peptide moiety, and to react with both the intact and deglycosylated proteins (Fig. 1A). In the initial screening of ~10,000 recombinant phages, 47 were positive. Figure 1B shows the immunoblot of MSG-LacZ fusion products synthesized in \(E.\) coli lysogens for two positive phages, AMW3 and AMW4. Although these 47 cDNAs were classified as members of the MSG gene family by DNA sequence analyses, none of them contained the entire sequence of any known MSG gene. We screened further the randomly-primed cDNA library by plaque hybridization with DNA probes proximal to the 5' or 3' terminus of the cloned MSG cDNAs. Thus, 20 5' and 35 3' cDNA clones were isolated and sequenced, but we failed to isolate the complete sequence of any MSG gene due to an unexpectedly high degree of sequence diversity.

Another trial was made to isolate a full-sequence cDNA clone by hybridization screening of the oligo(dT)-primed cDNA library with the 205-bp 5' terminus probe as described above. Three positive clones were selected and one of them carried a 3,578-bp cDNA fragment that encodes a complete reading frame (pHK168; Fig. 2A). The protein is composed of 1,076 amino acids and the deduced molecular weight was 120,765 Da, which is consistent with the protein size estimated by SDS-PAGE or cDNA sequences of partial fragments (see below) (DDBJ/EMBL/GenBank accession no. D21827). A fusion protein between a cloned MSG and glutathione S-transferase expressed in \(E.\) coli reacted with polyclonal antibody against deglycosylated MSG molecules, indicating that the cloned DNA encodes MSG (data not shown). The context around the predicted start codon is consistent with the eukaryotic translation initiator rule proposed by Kozak; \(^20\) i.e., a purine residue at position −4 and a guanine residue at position +4. However, the predicted stop codon, UAA, is followed directly by a stretch of adenine. This ambiguity suggests that this stop codon might have been artificially created by internal oligo(dT) priming due to the high AT content of \(P.\) carinii DNA, rather than by priming at the poly(A) tail. To test this possibility, we isolated and sequenced several cDNA clones for the C-termini of MSG polypeptides by plaque hybridization with the 3' MSG probe. As shown in Fig.
Figure 2. Sequence and hydropathy profile of the MSG5 protein. (A) DNA and deduced amino acid sequence the MSG5-cDNA cloned in pHK168 (single letter code). The nucleotide position is counted from the 5' end of the original clone and the amino acid position is counted from the initiation site of the MSG coding sequence. Potential N-linked glycosylation sites are underlined. (B) Hydrophobicity profile of the MSG5 sequence. The GCG program based on that of Kyte and Doolittle was used. The horizontal numbers refer to amino acid residues. Areas above the median line are hydrophobic, and areas below are hydrophilic.
3, all three clones examined were not identical to the cDNA in pHK168, but belonged to the same MSG gene family and retained a unique UAA stop signal at the position equivalent to that in pHK168, indicating that the UAA signal in pHK168 represents the translational stop site of the cloned MSG-cDNA. However, since a stretch of adenine directly follows UAA and there is no potential polyadenylation signal in the cloned DNA, it is unlikely that oligo(dT) priming was initiated at the poly(A) tail. The RNA detected by Northern blot hybridization was ~4 kb (Fig. 4). It is interesting that the transcripts appear to be homologous in size regardless of high sequence diversity.

The protein sequence is not identical to any of the MSG-polypeptide sequences so far reported but is highly homologous to all of them (see below). Analysis of hydropathicity and secondary structure revealed several areas that may be surface-exposed epitopes and the protein is hydrophilic except for the C-terminal region (Fig. 2B). The C-terminus is highly hydrophobic and likely to be an anchoring sequence in the cell membrane. Taking these results into consideration, we concluded that the cloned cDNA in pHK168 encodes a complete MSG polypeptide and therefore the cDNA was referred to as MSG5 following our previous designations of MSG-cDNA species, MSG1 through MSG4. It is noteworthy that MSG5 has a stretch of 25 threonine residues at the C-terminal region though their significance is not known (Fig. 2).

Figure 4. Northern blot analysis of MSG transcript. RNAs were isolated from rat P. carinii or uninfected rat lung (control), and subjected to RNA blot analysis using the MSG1 cDNA probe as described in Materials and Methods. Lanes 1 and 2, two RNA samples from rat P. carinii; lane 3, RNA sample from uninfected rat lung.

3.2. MSG gene family

Based on the results of Southern blot hybridization analysis using the MSG1 cDNA probe, we have previously suggested that P. carinii MSG forms a gene fam-
ili composed of multiple genes. We further explored the chromosomal localization(s) of the MSG gene family by karyotype hybridization analysis using pulse-field gel electrophoresis. It has been demonstrated that *P. carinii* has at least 13 chromosomes ranging in size from 0.3–0.7 megabase pairs. Under our experimental conditions, the MSG cDNA seemed to hybridize to most of the *P. carinii* chromosomes, in contrast to the dihydrofolate reductase probe that gave a single hybridization band on the same filter blot (data not shown). This is consistent with the recent observation by Kovacs et al. Therefore, it is clear that the MSG genes are distributed throughout most of the *P. carinii* chromosomes.

The amino acid sequence of MSG5 was compared to the complete or partial sequences of other MSG polypeptides reported previously. The protein alignment in Fig. 5 includes a full sequence of MSG from rat *P. carinii* (originally referred to as gp116) and partial sequences of the MSG counterpart, gpA, of ferret *P. carinii* and three MSG species, MSG1 through MSG3, of rat *P. carinii*. (We have shown only 600-bp sequences of MSG1 through MSG3 in the previous article and therefore include the entire sequences of the cloned MSG fragments: DDBJ/EMBL/GenBank accession no. D17438, D17439 and D17440, respectively.) The MSG subtypes of rat *P. carinii* are highly homologous over the entire length of the proteins and only a few small variable regions exist. One of them maps to the middle of the proteins between aligned positions 555–577, and another is the C-terminal stretch of threonine residues starting at position 1,087 in MSG5; this stretch does not exist in MSG1 and gp116, which instead carry TETK tetrapeptide repeats at the corresponding regions (Fig. 5). The percent identity between these polypeptides was calculated and is shown in Table 1. The MSG family of rat *P. carinii* conserves 65–80% identical amino acids among the subtypes and shares 32–36% identical amino acids with gpA of ferret-derived *P. carinii*. These results clearly demonstrate the conserved structural and functional organization of the MSG molecules.

The MSG proteins are rich in cysteine residues. The complete MSG5 and gp116 proteins each contain 59 cysteine residues, while the partial gpA polypeptide has 47. A most striking observation is that these cysteines are nearly perfectly conserved in all MSG subtypes regardless of species specificity. MSG5 from rat *P. carinii* shares only 32% identical amino acids with gpA from ferret *P. carinii*, however they share 94% (44/47 in gpA) and 86% (44/51 in MSG5) of cysteine residues within the overlapped protein sequences (Fig. 5). This highly conserved cysteine motif again emphasizes the conserved structural and functional organization of MSG.

### Table 1. Nucleotide and amino acid identities in the MSG genes*

<table>
<thead>
<tr>
<th></th>
<th>MSG1</th>
<th>MSG2</th>
<th>MSG3</th>
<th>MSG5</th>
<th>gp116</th>
<th>gpA</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSG1</td>
<td>100%</td>
<td>83%</td>
<td>85%</td>
<td>86%</td>
<td>89%</td>
<td>57%</td>
</tr>
<tr>
<td>MSG2</td>
<td>67%</td>
<td>72%</td>
<td>85%</td>
<td>87%</td>
<td>85%</td>
<td>59%</td>
</tr>
<tr>
<td>MSG3</td>
<td>68%</td>
<td>70%</td>
<td>22%</td>
<td>84%</td>
<td>85%</td>
<td>57%</td>
</tr>
<tr>
<td>MSG5</td>
<td>74%</td>
<td>73%</td>
<td>65%</td>
<td>1,076</td>
<td>88%</td>
<td>57%</td>
</tr>
<tr>
<td>gp116</td>
<td>80%</td>
<td>72%</td>
<td>69%</td>
<td>78%</td>
<td>1,089</td>
<td>57%</td>
</tr>
<tr>
<td>gpA</td>
<td>33%</td>
<td>36%</td>
<td>35%</td>
<td>32%</td>
<td>34%</td>
<td>790</td>
</tr>
</tbody>
</table>

a The entries (underlined) on the diagonal refer to the number of residues in each protein. Entries below the diagonal refer to the percentages of identical amino acid residues over the entire lengths of the proteins based on the alignments in Fig. 5. Entries above the diagonal refer to the percentages of identical nucleotide residues over the entire lengths of the genes. This work. Kovacs et al. 15. Haidaris et al. 14.

#### 3.3. Antigenic diversity

The presence of multiple MSG genes indicate possible antigenic diversity on the cell surface of *P. carinii*, which could play a role in immune evasion. While antigenic variability is not shown in the present study, recent immunological studies have demonstrated that *P. carinii* isolates from rats, mice and humans have species-specific, as well as shared, antigenic determinants. We asked the question if cloned MSG antigen is specific to *P. carinii* from rats or contains a common determinant recognized by other species. Thus, we examined immunoreactivities of the native and recombinant MSG molecules of rat-derived *P. carinii* by Western blotting with several antibodies against MSG of *P. carinii* from rats, mice and humans, including the polyclonal antibody against the recombinant MSG1 polypeptide. The representative immunoblots are shown in Fig. 6 and the immunoreactivities are summarized in Table 2. The rabbit antibody against the recombinant MSG polypeptide reacted with authentic MSG of rat *P. carinii* (data not shown; see ref. 16), and rabbit antibody or rat serum against authentic MSG of rat *P. carinii* also recognized the recombinant MSG polypeptide (data not shown). Similarly, mouse sera raised against mouse *P. carinii* recognized both authentic and recombinant MSG molecules from rat *P. carinii* (Fig. 6B). These results clearly demonstrated that the recombinant MSG polypeptide is not a rat *P. carinii*-specific immunodeterminant but also shares common epitopes with MSG from mouse *P. carinii*. On the other hand, antiserum derived from one AIDS patient recovering from *P. carinii* pneumonia reacted to the native MSG but not to the recombinant MSG (Fig. 6C), indicating that there exists an MSG subtype in rat *P. carinii* that differs antigenically from the cloned MSG but shares common epitopes with human *P. carinii*, or that the authentic glycosylation of the recombinant MSG provides the common immunodeterminant with human *P. carinii*
Alignment of the deduced amino acid sequences of six homologous clones encoding the major surface glycoproteins of P. gollai, as well as ferret-derived P. carinii, underlined * indicates that a residue is conserved with gpA. A # indicates that a cysteine residue is conserved in all MSG from rat-P. carinii. 

**Figure 5.** Alignment of the deduced amino acid sequences of six homologous clones encoding the major surface glycoproteins of P. carinii from rats and ferrets. gp116 denotes one MSG gene of rat-derived P. carinii and gpA denotes MSG from ferret-derived P. carinii. An * indicates that a residue is conserved among all MSG clones from rat that overlap in that region. An underlined * indicates that a residue is conserved with gpA. A # indicates that a cysteine residue is conserved in all MSG from rats as well as ferret-derived P. carinii.
Major Surface Glycoprotein Family of *P. carinii* [Vol. 1, 801-900]

**Figure 5.** Continued

**Figure 6.** Western blots of native and recombinant MSG polypeptides with several antibodies. Bulk proteins of rat *P. carinii* or induced *E. coli* lysogens were subjected to SDS-PAGE and analyzed by immunoblotting using several antibodies. Experimental procedures and conditions are essentially the same as described in Fig. 1 except that the immunoblots were stained by the nonradioisotope method in which the second antibodies were conjugated with the enzyme horseradish peroxidase and detected by enhanced chemiluminescence using the Hyperfilm™-ECL. Lane 1, rat *P. carinii*; lane 2, *E. coli* lysogen for AMW4 (MSG1); lane 3, *E. coli* lysogen for AgtII (control). The immunoreactive band above MSG in lane 1 represents MSG oligomers. (A) Monoclonal antibody MA-2 against deglycosylated MSG from rat *P. carinii* (control). (B) Mouse antiserum against mouse *P. carinii* MSG. (C) Human antiserum against human *P. carinii* MSG.

**4. Discussion**

While the incidence of *P. carinii* pneumonia is increasing due to AIDS, the difficulty in mass cultivation of *P. carinii* still impedes the progress of the fundamental studies of this organism. In the current study, we have successfully isolated the cDNA that encodes the complete genes for the major cell surface glycoproteins. Kovacs et al.15 have also reported the full sequence of an MSG-cDNA, but it is possible that their sequence does not actually exist in *P. carinii*, since the sequence was completed by PCR amplification of a missing part.
Table 2. Antigenicity of the recombinant MSG antigen to antisera from different hosts.

<table>
<thead>
<tr>
<th>Antisera</th>
<th>Immune reactivity of antigena</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Authentic MSG from</td>
</tr>
<tr>
<td></td>
<td>rat <em>P. carinii</em></td>
</tr>
<tr>
<td>Rabbit anti-recombinant-MSG, b</td>
<td>+</td>
</tr>
<tr>
<td>Rabbit anti-native-MSG, b</td>
<td>+</td>
</tr>
<tr>
<td>Rat anti-MSG, c</td>
<td>+</td>
</tr>
<tr>
<td>Mouse anti-MSG, d</td>
<td>+</td>
</tr>
<tr>
<td>Human anti-MSG, e</td>
<td>+/−</td>
</tr>
</tbody>
</table>

a Determined by Western blot analysis using the indicated antibodies as shown in Fig. 6.

b MSG represents MSG from rat *P. carinii*. Rabbit polyclonal antibodies were raised against the recombinant MSG1 or native MSG proteins. 16

c Antiserum was obtained from thymic sentinel animals maintained in the same room as *P. carinii*-infected nude rats. Latent infection in these animals was confirmed by PCR amplification of SS rDNA from *P. carinii* as described previously. 29
d MSG represents MSG from mouse *P. carinii*. Antiserum was obtained from mouse immunized with mouse *P. carinii*.

e MSG represents MSG from human *P. carinii*. Sera from five AIDS patients recovered from *P. carinii* pneumonia were tested for their reactivities to MSG. +/− denotes that only one antiserum reacted to the native MSG from rat *P. carinii*, while the others did not. This positive serum did not react with the recombinant MSG1.

of their clone which could generate artifacts due to the great variability among MSG or as a result of errors introduced during PCR. From this point of view, this report may contain the first complete structure of cDNA for an MSG polypeptide. These analyses clearly demonstrate the presence of multiple MSG genes that encode a family of related major cell surface glycoproteins and their distributions on multiple *P. carinii* chromosomes.

The MSG5 protein contains 6 potential N-linked glycosylation sites (Asn-X-Ser/Thr) as well as 59 cysteine residues (Fig. 2). These protein features enable complex structures of MSG molecules by glycosylation as well as by intra- or interpolyptide disulfide bond formation. The protein mass of complex and dense cell walls of *P. carinii* is composed of these highly abundant MSG molecules. There exists a hydrophobic domain at the C-terminus of the MSG protein that could be used for membrane anchorage. However, there is no hydrophilic region compatible with an intracytoplasmic domain, nor is there a hydrophobic leader sequence. The highly conserved cysteine residues in MSG suggest that these cysteines play a key role in building the functional structure of MSG on the cell surface. It seems that the cysteine residues do not occur at random but appear at some intervals in MSG. In most cases, adjacent cysteines are spaced 6 amino acids apart. Assuming that these regions are capable of forming α-helical structures, cysteines may be arranged in the same surface of the helix. Accordingly, one can speculate that the occurrence of intra- or interpolyptide disulfide bond formation may create or link the repetitive secondary structure on the *P. carinii* cell surface, which may be functionally important in host-organism interaction, immune response or other pathobiological activities of *P. carinii*.

Pottratz et al. 12,13 have demonstrated that MSG participates in the attachment of the organism to host lung cells via fibronectin, a host-derived adhesive glycoprotein. They reasoned that MSG has an integrin receptor-like domain susceptible to fibronectin since antiserum against a synthetic peptide for the cytoplasmic domain of the β1 subunit of integrin cross-reacted with rat *P. carinii* MSG. 13 However, MSG5 protein sequence (this work) do not carry any putative integrin receptor-like sequences. Therefore, these molecules may have more than one method for interacting with cell surface proteins. The MSG-fibronectin interaction remains to be clarified at the molecular level using the recombinant MSG molecules.

The presence of multiple MSG genes suggests the potential for antigenic variability used for immune evasion. 15,16 Antigenic variability of surface structures is observed in various pathogens from prokaryotes such as *Neisseria gonorrhoeae* 26 to eukaryotes such as *Trypanosomes* 27 and worm parasites such as *Schistosoma mansoni*. 28 This variability allows escape from cellular or humoral immunity of the host organism. Therefore, since *P. carinii* is a fatal opportunistic pathogen of AIDS patients, a particularly pressing question is do *P. carinii* MSG antigens vary in the cell cycle or during the proliferation of the organism? The deduced structure of the MSG reported here provides insight into the structure and function of these molecules, which is critical to the understanding of their role in host-organism interaction.
and immune response.

Acknowledgements: We thank K. Tanabe for antibodies, A. Yasuoka and S. Oka for patient sera, K. Kawakami and K. Ito for discussion, and Chris Olsson for critical reading of the manuscript. This work was supported by grants from The Ministry of Education, Science and Culture, Japan, Uehara Memorial Foundation, Tokyo Biochemical Research Foundation, and Japan Health Sciences Foundation.

References