ROSELY M. ZANCOPÉ-OLIVEIRA,^{1,2}* ERROL REISS,² TIMOTHY J. LOTT,² LEONARD W. MAYER,² and GEORGE S. DEEPE, JR.³

Laboratório de Micologia Médica, Hospital Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil,¹ Centers for Disease Control and Prevention, Atlanta, Georgia,² and University of Cincinnati College of Medicine, Cincinnati, Ohio³

Received 25 September 1998/Returned for modification 10 November 1998/Accepted 6 January 1999

The major diagnostic antigens of Histoplasma capsulatum are the H and M antigens, pluripotent glycoproteins that elicit both humoral and T-cell-mediated immune responses. These antigens may play a role in the pathogenesis of histoplasmosis. M antigen is considered immunodominant because antibodies against it are the first precipitins to arise in acute histoplasmosis and are commonly present during all phases of infection. The biological activity of monomolecular M antigen and its ability to elicit a protective immune response to H. capsulatum are largely unknown. A molecular approach was used to identify the biological nature of M antigen, including its purification from histoplasmin, partial digestion with proteinases, and reverse-phase highperformance liquid chromatography to separate the released peptides. The amino acid sequences of the purified peptides were obtained by Edman degradation, and using degenerate oligonucleotide primers for PCR, a 321-bp fragment of the gene encoding the M antigen was amplified from genomic H. capsulatum DNA. This fragment was used to screen an H. capsulatum genomic DNA library, leading to the isolation, cloning, and sequencing of the full-length gene. The M gene consists of 2,187-bp DNA encoding a protein of 80,719 Da, which has significant homology to catalases from Aspergillus fumigatus, Aspergillus niger, and Eimericella nidulans. A cDNA was generated by reverse transcription-PCR and cloned into the expression vector pQE40. The identity of the cloned, expressed protein was confirmed by Western blotting. The recombinant fusion protein was immunoreactive with monoclonal antibodies raised against M antigen, with polyclonal mouse anti-M antiserum, and with a serum sample from a patient with histoplasmosis. The gene encoding the major immunodominant M antigen of *H. capsulatum* is a presumptive catalase, and the recombinant protein retains serodiagnostic activity.

Histoplasmosis is a systemic fungal disease caused by *Histoplasma capsulatum* var. *capsulatum*, a dimorphic fungus which grows in the mycelial form at room temperature in the environment and in the yeast form at 37°C on enriched medium or in infected tissues. The significance of histoplasmosis results from its worldwide distribution associated with bird or bat excrement, its ability to mimic other serious diseases (23), and its recognition as a common and serious opportunistic infection in patients with AIDS (24).

The major diagnostic antigens of H. capsulatum are the H and M antigens, pluripotent glycoproteins that elicit both humoral and T-cell-mediated immune responses (9). These antigens may also play a role in the pathogenesis of histoplasmosis. M antigen is considered immunodominant because antibodies against it are the first precipitins to arise in acute histoplasmosis and are more commonly present during all phases of disease (11, 22). The biological activity of monomolecular M antigen and its ability to elicit a protective immune response to H. capsulatum are largely unknown because of difficulties in obtaining sufficient quantities of monomolecular M antigen by biochemical purification procedures. M glycoprotein has been proposed to be a catalase, with an activity that could protect yeast forms from oxidative fungicidal mechanisms, but definitive evidence to support this hypothesis awaits confirmation (8). Immunochemical analysis of M antigen indicated it is a glycoprotein with a molecular mass of \sim 70 to 94 kDa (8, 17, 25, 26) and an isoelectric point of 4.7 (7). M antigen contains species-specific protein epitopes and glycosidic epitopes; the latter are N linked to the peptide core and probably belong to the high-mannose-type glycan (26). The peptide epitopes react with human antibodies and are not affected by N deglycosylation (27). Biochemically purified M antigen has been shown to evoke proliferation of T cells (9).

The diagnosis of histoplasmosis is based on the results of careful clinical evaluation and associated laboratory tests. Isolation of H. capsulatum from biological specimens provides a definitive diagnosis, but this fungus frequently fails to grow under artificial culture conditions. In such cases, serologic tests, such as complement fixation and immunodiffusion, are indicated for detection of antibodies against M antigen. Serodiagnosis has some limitations, including misleading positive results for patients with other diseases caused by microbes that cross-react with H. capsulatum. Much of this cross-reactivity is due to carbohydrate determinants (1a, 7, 11, 12, 16, 27), and they can be reduced after chromatographic purification of the M antigen (25). Even purified M antigen evokes cross-reactivity in more-sensitive assays, such as enzyme immunoassay or Western blotting, owing to covalent carbohydrate determinants that are cross-reactive with similar determinants present among genera of dimorphic fungal pathogens and in certain bacteria (27). The detection of antibodies against peptide epitopes within native or recombinant M antigen could improve the specificity and sensitivity of the serologic tests and provide an earlier and more specific diagnosis of acute histoplasmosis. In this study, our objective was to identify the biological

^{*} Corresponding author. Mailing address: Laboratório de Micologia Médica do Hospital Evandro Chagas, Fundação Oswaldo Cruz, Av. Brasil 4365, Manguinhos, Rio de Janeiro 21045-900 Brazil. Phone: (55-21) 290-1943. Fax: (55-21) 590-9988. E-mail: zancope@fiocruz.br.

TABLE 1. Amino acid sequences of the NH₂ terminus and lysyl endopeptidase-digested fragments of M antigen

Origin	Amino acid sequence
NH ₂ terminus	S D P T D Q F L
2642-m1947/19	DFIFRQKIQHFDHER
5070-m1941/20	TLQGRAGLV
V22-m1947/20	AQALGGKNPDFHRQDL
V21-m1947/12	SGRYPE
V16-m1941/21	F D F D L L D P T K
V18-m1941/23	I I P E E L V P F T P I G K

nature of M antigen as a step in the eventual development of a recombinant antigen. With monomolecular M antigen, it would then be possible to conduct epitope mapping, improve the immunodiagnosis of histoplasmosis, and evaluate the role of M antigen in eliciting delayed-type hypersensitivity or its in vitro correlates and in the pathogenesis of histoplasmosis. Therefore, the gene encoding M glycoprotein was cloned, expressed, and sequenced.

MATERIALS AND METHODS

Strains, plasmids, and culture conditions. Yeast phase cells of *H. capsulatum* 6623 (ATCC 26320) were grown at 37° C in Pine's liquid medium (15, 16) for 48 h to late log phase. *Escherichia coli* q358 was used as the host for the bacteriophage λ Gem11, and *E. coli* INV F' (Invitrogen Co., Carlsbad, Calif.) was used as the recipient for the subcloning vector pBluescript SK(-) (Stratagene, La Jolla, Calif.).

Purification and amino acid sequence of M antigen. M antigen was purified from histoplasmin by cation-exchange chromatography in columns of carboxymethyl-Sepharose CL-6B, as described previously (25). Samples of M antigen were subjected to sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis (SDS-10% PAGE) and electrotransferred for 1 h at 400 mA to polyvinylidene difluoride membranes (Immobilon-P; Millipore Corp., Bedford, Mass.) in 25 mM Tris, 192 mM glycine, and methanol (20% [vol/vol]). The membrane was washed several times with 1 mM dithiothreitol (DTT), stained with Ponceau S, and destained with 10% aldehyde-free acetic acid-1 mM DTT. Several washings with 1 mM DTT were made to remove the acetic acid. The band was identified by its molecular mass, and its identity was confirmed by Western blotting. The band in the membrane corresponding to M antigen (200 pmol/protein band) was excised and submitted to Edman degradation without prior modification. To obtain the internal amino acid sequences, the band was digested in situ with lysyl endopeptidase (Boehringer Mannheim Corp., Indianapolis, Ind.), and peptides were purified by microbore reverse-phase high-performance liquid chromatography on reverse-phase C_{18} silica. All amino acid sequences were obtained with a model 477A protein sequencer or Procise (Perkin-Elmer/Applied Biosystems, Foster City, Calif.).

DNA isolation. Yeast form cells grown in 50 ml of Pine's broth were harvested by filtration on a 0.45-µm-pore-size membrane, washed three times with deion-ized H₂O, and blotted to remove excess moisture. Yeast form cells were placed in a sterile mortar with 1 g of glass beads (0.5-mm diameter); liquid N₂ was added, and the mixture was ground to a fine powder in a biological safety cabinet. Personal protective equipment was used, and other precautions were taken according to established biosafety guidelines (20). The triturate was resuspended in 20 ml of TE Buffer, pH 8.0 (10 mM Tris-1 mM EDTA), and DNA was extracted with phenol, ethanol (EtOH) precipitated, and then dried and redissolved in 0.05 M TE. The RNA was removed after digestion with RNase (final concentration, 10 µg/ml) (Boehringer Mannheim) at 37°C for 1 h, followed by proteinase K treatment (50 mg/ml) (Sigma Chemical Co., St. Louis, Mo.) for an additional 1 h at 37°C. The DNA was subjected again to phenol extraction, EtOH precipitated, and redissolved in TE (18).

Generation of M DNA probe by PCR. *H. capsulatum* genomic DNA was used as a template for PCR amplification of a DNA fragment encoding an internal portion of the M protein. Degenerate oligonucleotide primers were designed based on the amino acid sequences of the two internal peptides (V22 and V18) of M antigen because of initial ambiguity in the NH₂ terminal sequence (Table 1). The sense primer, M4F, 5'-AA(AG)AA(CT)CC(AGC)GA(CT)TT(CT)-3', was a 17-mer with 48-fold degeneracy, and the antisense primer, M8R, 5'-TT(A GCT)CC(AGT)AT(AGCT)GT(AG)AA-3', also was a 17-mer with 96-fold degeneracy. PCR was conducted in a total volume of 100 µl containing 100 ng of DNA as a template, 100 mM (each) deoxynucleoside triphosphate, 1 mM (each) oligonucleotide primer, and 10× PCR buffer containing 500 mM KCl, 100 mM Tris-HCl (pH 8.3), 25 mM MgCl₂, and 2.5 U of *Taq* polymerase (Boehringer Mannheim). The amplification conditions consisted of denaturation at 95°C for 5 min followed by 35 cycles of the succeeding steps: denaturation at 95°C for 5 min, annealing at 50°C for 1 min, and extension at 72°C for 1 min. A final elongation step was done at 72°C for 5 min. A 321-bp PCR product was subcloned into the pCRII vector with the TA cloning kit (Invitrogen, Inc., Carlsbad, Calif.) by procedures recommended by the vendor and sequenced with a dyelabeled terminator and automated sequencer (Perkin-Elmer/Applied Biosystems).

Screening of a *H. capsulatum* genomic library. The 321-bp amplicon was labeled with [α -³²P]dCTP by High Prime DNA labeling mix (Boehringer Mannheim), purified by passage through a DEAE column (NACS Prepac Convertible, Bio-Rad Laboratories Life Technologies, Inc., Gaithersburg, Md.), and used for screening the genomic library, derived from DNA partially digested with *Sau*3AI and cloned into λ Gem11 via the *XhoI* half-site, kindly provided by Glenmore Shearer. An *E. coli* q358 bacterium infected with the genomic library was replica plated onto nitrocellulose membranes. The plaques were lysed and then heat fixed. The filters were hybridized with a ³²P-labeled probe. Twelve positive colonies were purified and mapped by Southern analysis (18). These clones were digested with *Bam*HI, and one fragment of 4.0 kb was obtained.

Gene sequence analysis. The 4.0-kb fragment was subcloned into pBluescript II KS and sequenced by the strategy of "primer walking" by the dideoxy chain termination method. Oligonucleotides (22-mer) were synthesized on the basis of DNA sequence and applied to initiate the sequencing reaction. The clone was sequenced in both directions. To determine the sites of putative introns, 5 μ g of RNA was reverse transcribed, with oligo(dT) to initiate the cDNA reaction. The first strand of cDNA was amplified with a sense primer located at the start site of the mature protein: the sequence of this primer was 5'-CCGAACTCTCCCAA CCGCA-3'. The antisense primer was 5'-ACCAAGCTTCTATCCAA CGGGAACCGA-3'. A 5' *Eco*RI site (underlined) was added to the sense primer, and a *Hind*III site (underlined) was performed for 35 cycles of 94°C for 45 s, 50°C for 45 s, and 72°C for 2 min, with 5 U of *Vent* polymerase (New England Biolabs, Beverly, Mass.). The PCR product was digested with *Eco*RI and *Hind*III and cloned into pBluescript SK(–), restriction mapped, and sequenced in its entirety.

Southern hybridization. Standard conditions for electrophoresis and Southern blotting were used (18). Hybridization was performed with digoxigenin (DIG)-labeled 321-bp probe at 42°C overnight. All the prehybridization, hybridization, and membrane-washing steps were performed with DIG Easy Hyb kit and DIG Wash and Block buffer set (Boehringer Mannheim) according to the manufacturer's recommendations (1).

Expression of M gene. cDNA was synthesized and used as a template for PCR. The sense primer was 5'-CGGAATTCAGATCTGACCCTACGGACCAG-3', and the antisense primer was 5'-ACCAAGCTTCTAGCTTCTACCAACGGA AA-3'. The cDNA encoding M antigen was amplified by PCR with these primers in the presence of 2 U of *Vent* polymerase (New England Biolabs). The conditions were 94°C for 45 s, 55°C for 45 s, and 72°C for 1 min. Thirty-five cycles were performed. The amplified cDNA was digested with *Bg*/II and *Hind*III (Gibco BRL, Grand Island, N.Y.) and ligated into pQE40 (Qiagen, Valencia, Calif.) that had been digested with the same enzymes. The ligation reaction product was cloned into *E. coli* XL-1 Blue (Stratagene). Colonies of *E. coli* were examined for the presence of the proper plasmid by restriction mapping. One clone, pQE40M6, was randomly selected for expression. The recombinant protein was purified by affinity chromatography on a nickel-Sepharose column as described previously (4).

SDS-PAGE and Western blotting. Native and recombinant M antigens were first dissociated at 100°C for 5 min in 0.125 M Tris-HCl buffer (pH 6.8) containing 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, and 0.025% bromophenol blue. SDS-PAGE was then conducted with 7.5% separatory and 4% stacking gels in an electrophoresis cell (Mini-Protean II; Bio-Rad Laboratories, Richmond, Calif.) operated at 10 mA of constant current for stacking and 20 mA for protein separation. The gel contents were electrotransferred to nitrocellulose membranes in a Mini Trans-Blot cell (Bio-Rad Laboratories) containing the transfer buffer, 25 mM Tris-HCl, 192 mM glycine, and methanol (20% [vol/vol], pH 8.3) operated at 400 mA for 1 h. Free binding sites in the membranes were blocked by incubation for 30 min in 5% nonfat milk in phosphate-buffered saline (PBS)-0.3% Tween 20 (pH 7.5) (PBS-T). The membrane was sliced vertically, and the strips were incubated for 60 min at room temperature with (i) polyclonal mouse anti-M serum and polyclonal mouse anti-H serum, a gift from Sandra L. Bragg, Centers for Disease Control and Prevention; (ii) ascitic fluids containing murine monoclonal antibodies (MAbs) EC2-EC7 and EC2-AH12, a pool of MAbs specific for the M antigen, and MAb CA1-CB4, which binds C antigen (12, 17); and (iii) a serum specimen from a patient with culture-confirmed histoplasmosis. A preimmune mouse serum and a serum sample from a healthy individual were used as negative controls. Both polyclonal antisera and the MAbs were diluted 1:100 to 1:1,000 in PBS-T containing 5% nonfat milk. After each strip was washed in PBS-T four times for 20 min each time, horseradish peroxidaseconjugated goat anti-mouse immunoglobulin G and goat anti-human immunoglobulin G (Bio-Rad Laboratories), optimally diluted 1:1,000 in PBS-T, were added and incubated as described above. The blot strips were washed again and incubated with a substrate solution of 0.5 mg of 3,3'-diaminobenzidine tetrahydrochloride (Bio-Rad Laboratories) per ml plus 5 µl of H2O2 (30% [vol/vol]) per

FIG. 1. M gene cloning strategy. The M antigen peptide sequences are shown in their presumed positions, inferred from amino acid homology with fungal catalases. Primers M4F (sense primer) and M8R (antisense primer) were designed based on two internal peptides, V22-m1947/20 and V18-m1941/23, respectively (Table 1), and were used in PCR amplification of a 321-bp fragment of *H. capsulatum* genomic DNA.

 $50\ {\rm ml}$ of PBS. After color development, the strips were rinsed exhaustively in deionized water.

Nucleotide sequence accession number. The M gene sequence was deposited in GenBank, and its accession number is AF026268.

RESULTS

Amino acid sequencing of M antigen. The amino acid sequences of the NH_2 terminus and six internal peptides are shown in Table 1. The amino acid sequences of two internal peptides, V22 and V18, of the M protein (Table 1) showed 66 to 73% identity with sequences of catalases of *Schizosac*-charomyces pombe (gpD55675 YSPC-1) and Aspergillus niger (gpZ23138 ANCATRGNA-1).

Cloning and sequencing of the M gene. The significant degree of homology of these two internal peptides to the catalases suggested a certain arrangement in the protein (Fig. 1). Considering their positions, two degenerate oligonucleotides were designed to amplify a fragment from genomic *H. capsulatum* DNA. A 321-bp PCR product was achieved by using HistoM4F and HistoM8R as primers and was confirmed by Southern blotting to represent a unique gene of *H. capsulatum*. Sequence analysis of this amplicon obtained by the dideoxy chain terminator method enclosed the two native internal peptides, confirming that the PCR product encoded a region of the M gene.

The 321-bp PCR fragment was gel purified in 1% agarose and used to screen an *H. capsulatum* genomic DNA library to isolate the entire M gene. A *Bam*HI genomic fragment of 4.0 kb carrying the M gene was isolated and characterized. This fragment was subcloned into pBluescript II KS and was sequenced in its entirety in both directions. Figure 2 shows the complete nucleotide sequence of the M gene and the deduced amino acid sequence, consisting of 729 residues with an estimated molecular mass of 80,719 Da.

The coding region of the M gene is interrupted by five introns (Fig. 2), with the 5' and 3' extremities presenting the GT/AG consensus (14). The coding region of the M gene presented a 0.46 A/T ratio. The 5' 565-bp flanking sequence of this gene exhibited regions similar to the promoter regions of other eukaryotic genes. A putative TATA element is at position 318, and a T+C-rich pyrimidine block is found downstream at position 365. The CAA motif is found twice upstream of the T+C block at positions 34 and 341. The 3' region downstream from the M gene open reading frame contains a pentanucleotide (5'-AAATA-3') at position 3138, 19 nucleotides downstream from the termination codon, similar to the polyadenylation consensus sequence described in other eukaryotes (14). This sequence is known to be involved in termination codon.

nation of transcription, processing, and addition of poly(A) at the 3' terminus.

Protein structure. Sequencing of the N terminus of the native protein showed that the first residue of the mature protein is a serine residue at position 614, giving a mature protein of 689 amino acids with a predicted mass of 76,491 Da. Therefore, the expected M gene has a leader peptide composed of 40 amino acids. The initiator methionine is followed by an arginine (R), a positively charged residue, and a core of 10 or more hydrophobic residues following the positively charged amino acid, features described as characteristics of signal peptides (10). The sequence also has a purine (adenine) in the -3 position, a prerequisite for an initiation codon (13). Five potential N glycosylation sites (NXT or NXS) were predicted (Fig. 2).

Comparison of the deduced amino acid sequence of M gene with known sequences. The earlier database results showing that two peptide sequences of M protein had 66 to 73% identity with sequences of catalases of *S. pombe* (gpD55675 YSPC-1) and *A. niger* (gpZ23138 ANCATRGNA-1) suggested that M antigen could be a catalase. Comparison of the complete deduced amino acid sequence of M with known fungal catalases from *Aspergillus fumigatus*, *A. niger, Eimericella nidulans*, and *Saccharomyces cerevisiae* demonstrated 61.2, 53.2, 60.4, and 21.7% similarity, respectively, at the amino acid level. Figure 3 shows the multiple alignment of the amino acid sequences encoded by the M gene and those of fungal catalases. The M sequence can be divided into parts with high and low homology, suggesting functional domains.

Copy number of M gene. Southern blotting of *H. capsulatum* genomic DNA digested with various restriction enzymes was probed with the 320-bp PCR product in order to evaluate the genomic organization of the M gene. A single hybridized band of 4.0 kb was seen with the *Bam*HI-digested genomic DNA, which corresponded to the sizes of the λ Gem11-purified inserts. The hybridization profile obtained when genomic DNA was digested with *Bam*HI and with other restriction endonucleases (data not shown) suggested that a single copy or very few copies of the M gene occur(s) in the genome.

Immunologic reactivity of recombinant M protein. SDS-PAGE of recombinant M protein revealed a band with a molecular mass of \sim 73 kDa. The identity of the cloned, expressed protein was confirmed by Western blotting probed with MAbs specific for M antigen (EC2-EC7 and EC2-AH12), with polyclonal mouse anti-M antiserum, and with a serum sample from a patient with culture-confirmed histoplasmosis. The recombinant fusion protein was immunoreactive with all anti-M antigen MAbs and with polyclonal anti-M antiserum samples but not with the MAbs and polyclonal serum samples used as controls (Fig. 4).

DISCUSSION

Despite the accumulated evidence of its importance in humoral immune responses, measured in clinical serologic tests (1a, 16, 27), the biological function of M antigen remains obscure, and its interaction with cells of the immune system is largely uncharacterized. The lack of availability of monomolecular M antigen has delayed efforts to determine whether it can induce protective immunity. Recombinant DNA technology was used to begin to address these issues by cloning, expressing, and characterizing the gene encoding M protein. Amino acid sequences from the N-terminal and internal peptides were determined, and the significant sequence similarity of two internal peptides of M protein with sequences of catalases of other fungi was observed. This sequence homology facilitated the design of degenerate primers to isolate a gene

1		
101	$\begin{tabular}{lllllllllllllllllllllllllllllllllll$	
201	TGGTTTATTATAGCCGCCCGGAAGCCCTGGCCTGTTAAATTTTGGACCTCCTCCGCGCCATTCTTTCCAAACTTCGTGCGTCCGTTTCCCCCC	
401	TATATCGATATCTAACTGCCATTGATTATTTGCCTCCACTCAAGAGGGTCCCAGTCAATTTGGTCCCTCCAACACTATCTGCATATGTAATA	
	M R S	3
501	CGCTCAAGCTTATACTCGCCTCGGCGGGGTTTGTTCTGCAGCCTGTCCCTACATGTCAGGGGAGATGCCTAGCGGTCAGAAAGGCCCCCTCGATGGCG	26
601	L K L I L A S A G V V S A A C P Y M S G E M P S G Q K G P L D R R CCATGACACTCTCCCGACCACTACGACGATCTTACATTGACGATGAACAGTCGGTGGTCAACAACGGACGTGGGTGG	36
701	H D T L S D F I D Q F L S T I D D Q F L S T I D D C Q S V L I T D V G G F I E GACAACACACACACACACACACACACACACACACACACA	69
801	D Q H S L K A G N R G P T L L E D F I F R Q K I Q H F D H E R gatacAaaaatatgtgaccgtgttgcaaatccgctaattcaatttacgcagGTTCCTGAGCGCGCCGTCCATGCTCGAGGAGCTGGTGCCCATGGCGTA	100
001	V P E R A V H A R G A G A H G V	116
901	FT S Y N N W S N I T A A S F L N A A G K O T P V F V R F S T V A G	150
1001	${\tt GTAGCAGAGGCAGTGTTGACTCTGCTCGCGATATCCACGGATTTGCGACCCGTCTGTATACCGATGAAGGCAATTTTGgtaagcattatatcgtggtagt$	
1101	S R G S V D S A R D I H G F A T R L Y T D E G N F D	176
1101	CatactcataacagcacaacaatatgaatacgaatacgaccaggacctagcggctactcggcaatgtagATATCGTCGGAAACAACGTTCAGTCTTCTTC I V G N N V P V F F	186
1201	ATTCAGGACGCTATTCAATTCCCTGATTGATTCACGCTGTCAAGCCGGCAACCAGACAGTGAAATTCCCCCAGGCTGCAACTGCACATGATACGGCATGGG I O D A I O F P D L I H A V K P O P D S E I P O A A T A H D T A W D	220
1301	ATTTCCTCAGCCAGCCCAGCTCATTGCATGCCCTCTTCTGGGCAATGTCAGGACATGGAATCCCTCGCTCAATGCGTCATGTTGATGGGTGGG	
1 4 0 1	FLSQQPSSLHALFWAMSGHGIPRSMRHVDGWGV	253
1401	H T F R L V T D E G N S T L V K F R W K T L O G R A G L V W E E A	286
1501	${\tt CAGGCTCTTGGCGGAAAGAATCCCGACTTCCATCGACAAGACCTCTGGGATGCCATTGAATCTGGAAGGTACCCTGAGTGGGAggtaagatatgattccc$	
1601	Q A L G G K N P D F H R Q D L W D A I E S G R Y P E W E	314
1001		
1701	catataGCTGGGCTTTCAATTGGTGAATGAAGCAGATCAATCCAAGTTTGATTTCGATCTATTAGATCCCACCAAAATCATCCCAGAAGAACTTGTTCCT	0.45
1801	L G F Q L V N E A D Q S K F D F D L L D P T K I I P E E L V P	345
	FTPIGKMVLNRNPKSYFAETEQIM	369
1901	atgaatacatgtatagctagatgaagcgtatatctaaatattttccacagTTCCAACCAGGTCATGTAGTTCGCGGAATCGATTTCACGGATGACCCTT	396
2001	TGCTTCAGGGCCGCTTGTACTCCTACCTTGACACTCAATTGAATCGCCATGGAGGGTCCCAACTTCGAGCAACTGCCGATCAACAGACCCCCGCATCCCATT	500
2101	L Q G R L Y S Y L D T Q L N R H G G P N F E Q L P I N R P R I P F	419
2101	H N N N R D G A G	428
2201	tctgcacagGACAAATGTTCATCCCTCTAAACACGGCCGCATATACACCCAACTCAATGAGCAACGGATTCCCACAAAGACCAACCGGACCCCATAACAG	
		450
2301	Q M F I P L N T A A Y T P N S M S N G F P Q Q A N R T H N R AGGATTCTTCACCGCACCTGGGCGTATGGTAAATGGACCACTAGTGCGCGGAGCTCAGCCGAGCTTCAACGACGTCTGGTCCCAACCGCGTCTCTTCTAC	458
2301	Q M F I P L N T A A Y T P N S M S N G F P Q Q A N R T H N R AGGATTCTTCACCGCACCTGGGCGTATGGTAAATGGACCACTAGTGCGCGAGCTCAGCCGAGCTTCAACGACGTCTGGTCCCAACCGCGTCTCTTCTAC G F F T A P G R M V N G P L V R E L S P S F N D V W S Q P R L F Y	458 491
2301 2401	Q M F I P L N T A A Y T P N S M S N G F P Q Q A N R T H N R AGGATTCTTCACCGCACCTGGGCGTATGGTAAATGGACCACTAGTGCGCGAGCTCAGCCGAGCTTCAACGACGTCTGGTCCCAACCGCGTCTCTTCTAC G F F T A P G R M V N G P L V R E L S P S F N D V W S Q P R L F Y AACTCACTCACGGTCTTCGAGAAGCAATTCCTCGTCAACGCCATGCGCTTCGAAAACCTCCACGGGGGGGAAAACGTCCATCATCC N S L T V F F K O F L V N A M P F F N S H V P S F T V P K N V T L O	458 491
2301 2401 2501	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525
2301 2401 2501 2601	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558
2301 2401 2501 2601	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591
2301 2401 2501 2601 2701	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591
2301 2401 2501 2601 2701 2801	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591 625
2301 2401 2501 2601 2701 2801 2901	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591 625 658
2301 2401 2501 2601 2701 2801 2901	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591 625 658 691
2301 2401 2501 2601 2701 2801 2901 3001	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591 625 658 691 725
2301 2401 2501 2601 2701 2801 2901 3001 3101	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 591 625 658 691 725
2301 2401 2501 2601 2701 2801 2901 3001 3101 3201	Q M F I P L N T A A Y T P N S N G F P Q Q A N R T H N R AGGATTCTTCACCGCACCTGGGCGTATGGTAAATGGACCACTAGTGCGCGGGGGGGG	458 491 525 558 625 658 691 725 729
2301 2401 2501 2601 2701 2801 3001 3101 3201 3301	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 625 658 691 725 729
2301 2401 2501 2601 2701 2801 3001 3101 3201 3301 3401	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 625 625 658 691 725 729
2301 2401 2501 2601 2701 2801 3001 3101 3201 3301 3401 3501 3501	$ \begin{array}{ccccc} & \begin{array}{ccccccccccccccccccccccccccccccccccc$	458 491 525 558 625 625 658 691 725 729
2301 2401 2501 2601 2701 2801 3001 3101 3201 3301 3401 3501 3601 3701	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	458 491 525 558 625 658 691 725 729

FIG. 2. Nucleotide sequence and deduced amino acid sequence of M antigen of H. capsulatum. The TATA element, the CAAG motifs, the T+C-rich block, and the stop codon are in bold, italic, and shaded type. Base positions are shown on the left, and amino acid positions are shown on the right. Introns are indicated by lowercase letters. The leader peptide (1 to 40) is bold and shaded. The potential N-glycosylation sites are in italics, and the 321-bp fragment is underlined.

fragment by PCR, which then was used to screen an *H. cap*sulatum genomic library. Several lines of evidence confirmed that the cloned cDNA and gene described in this research represent the authentic M antigen. The deduced amino acid sequence includes all the peptide sequences obtained from the native protein. The deduced mature M open reading frame encodes a polypeptide of 689 amino acids with a predicted mass of 76,491 Da, which is concordant with the molecular mass of the native glycoprotein described by immunochemical analysis of M antigen (8, 12). In addition, the recombinant M antigen, expressed as a fusion protein in *E. coli*, reacted strongly with MAbs against M antigen as well as with polyclonal mouse M antiserum and with an antiserum sample from a patient with histoplasmosis. No reactivity was observed with

1 100 En CatB ~mralglvgl v v n v mt lg.rrd ..tnp atea . ey ld td y Af CatB -mrltfipsl i v n v mt ln.rrd eisdg raaa ln ndafm s e . qy n 5 H ~~~~~~~ *** ----sdp dq . kf id eq v h n mrhfwllpav a i g q An CatR is maftge gdnag tiev qpidntl vn tg ym ft s t Ď Sc CatA α hllen as r ----- -G-A-A-CPY --GE----- -TE-FL-S--Y --D--S-LTT DVGGPIEDQ- SLKAG-RGPT LLEDFIFROK IORFDHERVP Consensus 200 101 En CatB đ f s e **s** 1 i e Af CatB đ f ate 9 Ì ŧ 1 a M nn W 1 n a s An CatR ۷ ytk a w v đ s nd e m C vf vt hc i f ap v i Sc CatA v kg crle eltds 1 d y apy qnv ykc gl g es tp t vsfk ew hwf t lr k vf sq r Dr Consensus ERAVHARGAG ANGVETSY-D -SNITAASEL -A-GK-TPVF VRFSTVAGSR GS-D-ARDVH GEATREYTDE GNEDIVGNN- PVFFIQDAIQ EPDLIHAVKP 201 300 En CatB kg r а t t f 1 ed st 1 DA qi £ Af CatB rg n ſm h dd αt \$ tm t f 8 1 A Ł м q S t 1 s а f w de n t alg fh W An CatR n nn ts 1 d h sta Sð n mn y S f aan t v rtp a v 886 ٧ ngak f Sc CatA d qshln fq dt.tiy yl tln esi qi tymfgdr t a wasmnays g s im nke kdty q vl dt fet t gdk ael s h Consensus -PD-EIPQAA TAHD-AWDFF SQQPS-LH-L LWAMSGHGIP RSFRHVDG-G VHTFR-VT-- G-SKLVKFHW KSLQG-ASLV WEEAQ--SGK NPDY-RQDL-301 400 En CatB es a n f 1 t r У Af CatB đ a t ik Γ q r í. . a м d gf lvn £ i 5 a skd 1 1 v ks i d Y An CatR n non h k v ca a n 1 v v e a t ad d t m . D Sc CatA tqlqn ek k fncy t tp qatk rysv n l whk ef lrkf ti t te vd tntcip rh qi a s kpsnsva P Consensus -ALE-GRYPE -EL-VQINDE EDQLRFGFDL LDPTKIVPEE -VP-TPLGKM -LNRNP-NYF AETEQVMFQP GHV-VRGIDF T-DPLLQGRL FSYLDTQLNR 401 500 En CatB ...va i pd s tlkg stlk 1 n q ۲ d ta n ak stda v ...v v Af CatB ph v s ktsv s k v d tts k ۵V ss e v v q . . q ...i м 1 t a t ms fq r h nr E۷ p el ps n v y tv g ...k v An CatR . . ť q..q tnw q h qt pt n h t **DS** î YBN yas h am w ٧ Sc CatA la.yq v nlgc yskgdsqyta eqcp kav f qrdg m yy fg epnyiss lpnqtlkfkn ev devsdkf kgivl evte vsvrkqeqdq Consensus HGGPNFEQLP INRPR----P -HNNNRDGAG QM--FIPLN- -AY-PNS--N G-P-QANQT- G-GFFTAP-R --NG-LVR-- S--F-D-WSQ PRLF-NSL-P 501 600 En CatB . k an k.. dv kn v is t dae ny d t sni a hr qs a ia a vd ae fsaata Af CatB id an k.. v kn i id r . r ae e n k adv tk kk d v g vg hpg vegast f м f sh r.. et rk - **i** d r . 1 f nka vpi tn 1r d ial trdd gsftiaeqlr 1 ep s An CatR a n sk n.. hrk vn v . rgl l de s n y ts k snv kp lsie qv f a ns hp ikqqqa mnn Sc CatA irnehi d k inqyyy ygi ldfeqpra lyek y eq kklfvhnvvc hackikdpkv kkrv qyf l lned gkvia egl pwepv dl gyaktws Consensus AEKQFVVNAI RFEN--V--- SP-V--NVI- QLNRV-NDLA -RV-A-AIGV --P-PDPT-Y H-N-T---GT FG--L--L-G LK-GVL-S-- --ES------601 700 lkaelsndnl viv a rf. .sn En CatB p aek gaks..... .a nss f 1 đ q v а Af CatB lrdrlkehg vvl a al. .ad dq t v a aes aassftggsa ns sga s t 1 t tq 1 t.... α đ g i ivl gssld pqr м a fnsannk . ga gsi i va 1....ltsa r r it ayay v d neal a g An CatR m aqfsaag lni t ay. .ad t a 1 d l ii d vqs aspalangmn st t.. fg v pa S V Sc CatA i san-----701 754 En CatBt fd ntave v.ada d s f nnle tvk sde а Af CatBa lr atsrd v.ags tdd f ndlk rtk b٧ h~~ М rdvlma ggd sngldqp i.snd s a yvrsvld t ayr n p rs~ An CatRv lk dssrs tossett k i kevl ytrv Sc CatA ******* -----Consensus -----A--N AGI-----GV Y-----V-E- -A----EGL- -F-FLDRF-L D---

FIG. 3. Comparison of the deduced amino acid sequence of M gene with those of other catalases from eukaryotes: *A. fumigatus* (Af), *E. nidulans* (En), *A. niger* (An), and *S. cerevisiae* (Sc). The GenBank accession numbers for the catalases are u87850, u80672, 115474, and x13028, respectively. Alignments were made with Genetics Computer Group, Inc. (Madison, Wis.) software. Blank spaces indicate identity with the consensus; ~, gaps placed at the ends of sequences; dots indicate internal gaps; and dashes indicate no consensus at a plurality of 3.



FIG. 4. Western blot analysis of recombinant M antigen probed with MAbs EC2-EC7 and EC2-AH12 and M pool anti-M antigen (lanes 1, 2, and 3, respectively); with polyclonal mouse serum anti-M antigen at 1:100 (lane 4), 1:500 (lane 5), and 1:1,000 (lane 6); with a serum sample from a patient with a culture-confirmed case of histoplasmosis at 1:100 and 1:500 (lanes 7 and 8); and with negative controls: a preimmune mouse serum at 1:100 (lane 9), a healthy human serum sample at 1:100 (lane 10), polyclonal mouse serum anti-H antigen (lane 11), and MAb CA1-CB4 anti-C antigen (lane 12).

the anti-C antigen MAb (CA1-CB4) and with serum samples used as controls.

Sequence analysis of cDNA obtained by reverse transcription-PCR demonstrated that the gene encoding M protein has five introns, consistent with data already observed for genes that encode other *H. capsulatum* proteins (4).

A comparison of the deduced M protein sequence to the sequences of catalases from other fungi showed a high degree of homology to catalases from *A. fumigatus*, *A. niger*, and *E. ni-dulans*. Fungal catalases are very characteristic and are distinguished from those produced by other eukaryotic organisms. They are larger molecules than those from the other eukaryotes, being composed of four subunits with molecular masses ranging from 80,000 to 97,000 Da (6, 21). M antigen was previously suggested to be a catalase, based on indirect evidence obtained from an analysis by a panel of monoclonal anti-catalases (8). The high degree of homology between M protein and other fungal catalases, as well as the estimated molecular mass of the M protein, 80,719 Da, which is within the size range of fungal catalases, strongly implies that the biological nature of M antigen is a catalase.

Catalases play a protective role for aerobic microbes, neutralizing the toxic effects of hydrogen peroxide (H_2O_2) via its hydrolysis to O_2 and H_2O . Calderon and Shennan (3) have demonstrated that the resistance of *Trichophyton quinckeanum* and *Trichophyton rubrum* to H_2O_2 -mediated killing is due to endogenous catalase. *H. capsulatum* is capable of replicating within polymorphonuclear cells and macrophages despite the generation of a respiratory burst (2, 5, 19). Therefore, M protein may aid in the intracellular survival of this fungus within polymorphonuclear cells by conferring resistance to oxidative fungicidal mechanisms.

The cloning and sequencing of the M gene and its expression as a recombinant protein open the possibility of mapping individual peptide epitopes to permit an analysis of B- and T-cell reactive sites within the protein. Additional knowledge gained in this area would enhance the understanding of the fungushost interaction and would broach an exploration of the role of M antigen in the protective immune response. Moreover, these data may contribute to the development of antifungal therapy directed against fungal catalases. Finally, the recombinant protein and/or its selected peptide epitopes will be useful in the development of more sensitive and specific diagnostic tests. Studies are already in progress in our laboratory to explore this last possibility.

ACKNOWLEDGMENTS

The research of R. M. Zancopé-Oliveira at the Mycotic Diseases Branch, Centers for Disease Control and Prevention, was supported by a fellowship from the American Society for Microbiology, and the work was supported in part by grant Proc. 522249/94-4 from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil. G. S. Deepe was supported by grants AI-34361, AI-42747, and HL-55949.

We thank José Mauro Peralta for review of the manuscript.

REFERENCES

- Boehringer Mannheim Biochemicals. 1995. Genius system user's guide for membrane hybridization. Boehringer Mannheim Biochemicals, Indianapolis, Ind.
- 1a.Bradley, G., L. Pine, M. W. Reeves, and C. W. Moss. 1974. Purification, composition and serological characterization of histoplasmin H and M antigens. Infect. Immun. 9:870–880.
- Bullock, W. E., and S. D. Wright. 1987. Role of the adherence-promoting receptors, CR3, LFA-1, and p150,95, in binding of *Histoplasma capsulatum* by human macrophages. J. Exp. Med. 165:195–210.
- Calderon, R. A., and G. I. Shennan. 1987. Susceptibility of *Trichophyton quinckeanum* and *Trichophyton rubrum* to products of oxidative metabolism. Immunology 61:283–288.
- Deepe, G. S., Jr., and G. G. Durose. 1995. Immunobiological activity of recombinant antigen from *Histoplasma capsulatum*. Infect. Immun. 63:3151– 3157.
- Eissenberg, L. G., and W. E. Goldman. 1991. *Histoplasma* variation and adaptive strategies for parasitism: new perspectives on histoplasmosis. Clin. Microbiol. Rev. 4:411–421.
- Fowler, T., M. W. Rey, P. Vaha-Vahe, S. D. Power, and R. M. Berka. 1993. The *catR* gene encoding catalase from *Aspergillus niger*: primary structure and elevated expression through increased gene copy number and use of a strong promoter. Mol. Microbiol. 9:989–998.
- Green, J. H., and L. Pine. 1985. Preparation of H and M antigens of *Histoplasma capsulatum* free of heterologous antigens. Curr. Microbiol. 12:209–216.
- Hamilton, A. J., M. A. Bartholomew, J. Figueroa, L. E. Fenelon, and R. J. Hay. 1990. Evidence that the M antigen of *Histoplasma capsulatum* var. *capsulatum* is a catalase, which exhibits cross-reactivity with other dimorphic fungi. J. Med. Vet. Mycol. 28:479–485.
- Harris, J. E., and G. S. Deepe. 1988. Characterization of antigenic determinants in histoplasmin that stimulate *Histoplasma capsulatum*-reactive T cells in vitro. Infect. Immun. 56:2343–2349.
- von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. 14:4683–4690.
- Kaufman, L., J. A. Kovacs, and E. Reiss. 1997. Clinical immunomycology, p. 575–584. *In* N. R. Rose, E. C. de Macario, J. D. Folds, H. C. Lane, and R. M. Nakamura (ed.), Manual of clinical laboratory immunology. American Society for Microbiology, Washington, D.C.
 Knowles, J. B. 1985. Immunochemical characterization of monoclonal anti-
- Knowles, J. B. 1985. Immunochemical characterization of monoclonal antibodies to factors in histoplasmin. Ph.D. thesis. University of North Carolina, Chapel Hill.
- Kozak, M. 1986. Point mutation defines a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44: 283–292.
- Lewin, B. 1994. Genes V, p. 127–159. Oxford University Press Inc., New York, N.Y.
- Pine, L. 1970. Growth of *Histoplasma capsulatum*. VI. Maintenance of the mycelial phase. Appl. Microbiol. 19:413–420.
- Pine, L., H. Gross, G. Bradley, J. R. George, S. B. Gray, and C. W. Moss. 1977. Procedures for the production and separation of H and M antigens in histoplasmin: chemical and serological properties of the isolated products. Mycopathology 61:131–141.
- Reiss, E., J. B. Knowles, S. L. Bragg, and L. Kaufman. 1986. Monoclonal antibodies against the M-protein and carbohydrate antigens of histoplasmin characterized by the enzyme-linked immunotransfer blot. Infect. Immun. 53: 540–546.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Schaffner, A., C. E. Davis, T. Schaffner, M. Markert, F. Douglas, and A. I. Braude. 1986. *In vitro* susceptibility of fungi to killing by neutrophil granulocytes discriminates between primary pathogenicity and opportunism. J. Clin. Investig. **78**:511–524.
- U.S. Department of Health and Human Services. 1993. Biosafety in microbiological and biochemical laboratories, 3rd ed. Health and Human Services publication (CDC) 93-8395, p. 80–81. U.S. Government Printing Office, Washington, D.C.
- Vainshtein, B. K., W. R. Melik-Adamyan, V. V. Barynin, A. A. Vagin, A. I. Grabenko, and V. V. Borisov. 1986. Three-dimensional structure of catalase

from Penicillium vitale at 2.0A resolution. J. Mol. Biol. 188:63-72.

- Wheat, L. J., M. L. V. French, R. B. Kohler, S. E. Zimmerman, W. R. Smith, J. A. Norton, H. E. Eitzen, C. D. Smith, and T. G. Slama. 1982. The diagnostic laboratory tests for histoplasmosis. Ann. Intern. Med. 97:680–685.
- Wheat, L. J. 1989. Diagnosis and management of histoplasmosis. Eur. J. Clin. Microbiol. Infect. Dis. 8:480–490.
- Wheat, L. J. 1996. Histoplasmosis in the acquired immunodeficiency syndrome. Curr. Top. Med. Mycol. 7:7–18.
- 25. Zancopé-Oliveira, R. M., S. L. Bragg, S. F. Hurst, J. M. Peralta, and E. Re-

Editor: T. R. Kozel

iss. 1993. Evaluation of cation exchange chromatography for the isolation of M glycoprotein from histoplasmin. J. Med. Vet. Mycol. 31:29–41.
26. Zancopé-Oliveira, R. M., S. L. Bragg, E. Reiss, and J. M. Peralta. 1994.

- Zancopé-Oliveira, R. M., S. L. Bragg, E. Reiss, and J. M. Peralta. 1994. Immunochemical analysis of the H and M glycoproteins from *Histoplasma capsulatum*. Clin. Diagn. Lab. Immunol. 1:563–568.
- Zancopé-Oliveira, R. M., S. L. Bragg, E. Reiss, B. Wanke, and J. M. Peralta. 1994. Effects of histoplasmin M antigen chemical and enzymatic deglycosylation on cross-reactivity in the enzyme linked immunoelectrotransfer blot method. Clin. Diagn. Lab. Immunol. 1:197–207.