Redundancy, phylogeny and differential expression of *Histoplasma capsulatum* catalases

Clayton H. Johnson,1,2 Martin G. Klotz,5 J. Lyndal York,4 Volker Kruft6 and Joan E. McEwen1,2,3

Author for correspondence: Clayton H. Johnson or Joan E. McEwen. Tel: +1 501 257 5559. Fax: +1 501 257 4820. e-mail: johnsonclaytonh@uams.edu or mcewenjoane@uams.edu

*Histoplasma capsulatum* produces an extracellular catalase termed M antigen, which is similar to catalase B of *Aspergillus* and *Emericella* species. Evidence is presented here for two additional catalase isozymes in *H. capsulatum*. Catalase A is highly similar to a large-subunit catalase in *Aspergillus* and *Emericella* species, while catalase P is a small-subunit catalase protein with greatest similarity to known peroxisomal catalases of animals and *Saccharomycotina* yeasts. Complete cDNAs for the CATA and CATP genes (encoding catalases A and P, respectively) were isolated. The transcriptional expression of the *H. capsulatum* CATA, CATB (M antigen) and CATP genes was assessed by Northern blot hybridizations on total RNA. Results at the transcript levels for these genes are shown for three conditions: cell morphology (mycelial versus yeast phase cells), oxidative stress (in response to a challenge with H₂O₂) and carbon source (glucose vs glycerol). Collectively, these results demonstrated regulation of CATA by both cell morphology and oxidative stress, but not by carbon source, and regulation of CATB and CATP by carbon source but not cell morphology or oxidative stress. A phylogenetic analysis of presently available catalase sequences and intron residences was done. The results support a model for evolution of eukaryotic monofunctional catalase genes from prokaryotic genes.

**Keywords:** antioxidants, isozymes, gene regulation, evolution

INTRODUCTION

*Histoplasma capsulatum* is a dimorphic fungal pathogen capable of causing acute pulmonary disease in otherwise healthy individuals and lethal disease in immunocompromised humans (Ampel, 1996; Eissenberg & Goldman, 1994; Wheat et al., 1985). In its most serious form, the infection disseminates throughout the body. Disseminated histoplasmosis, coinciding with laboratory evidence of HIV infection, is regarded as sufficient for a diagnosis of AIDS (Castro et al., 1992). Although AIDS currently represents the most prevalent immunocompromising disease of humans, a variety of other conditions or medical treatments can impair the human immune system and create susceptibility to diseases caused by the primary pathogen *H. capsulatum* as well as opportunistic pathogens (Goodwin et al., 1981). These predisposing conditions include advanced age, diabetes, cancer chemotherapy or immunosuppression induced to prevent rejection of transplanted organs (Wheat et al., 1982; Davies et al., 1978).

In nature, *H. capsulatum* exists as a mycelium that is well adapted for a saprotrophic mode of growth in soil (Scherr & Weaver, 1953). After entrance of microconidia or mycelial fragments into a mammalian host or in the laboratory after shift of mycelial cells to 37 °C, *H. capsulatum* differentiates into budding yeast (for a

The GenBank accession numbers for the cDNA sequences reported in this paper are AF139985 (CATB), AF189368 (CATA) and AF189369 (CATP).
review, see Maresca et al., 1994). In the animal host, the fungus experiences several environmental stresses, including heat shock, exposure to higher osmolarity, changes in pH, and oxidative stress (for reviews, see Deepe, 1997; Eisenberg & Goldman, 1994; Newman, 1999). The ability to resist or overcome these environmental stresses is likely to be important for continued growth and virulence. In addition, these stresses may be important triggers for changes in gene expression necessary for virulence.

*H. capsulatum* likely encounters oxidative stress during the earliest phases of infection of a mammalian host, for instance as a result of mitochondrial uncoupling and inhibition of mitochondrial respiration that is caused by heat shock (Davidson et al., 1996; Maresca et al., 1979; Patriarca et al., 1992; Scherr, 1957). Later in the infection process, the fungus also encounters oxidative stress due to production of reactive oxygen species during the host inflammatory response or oxidative burst from both polymorphonuclear leukocytes (PMNs) and activated macrophages (Deepe, 1994; Eisenberg & Goldman, 1994). Catalase, which converts hydrogen peroxide to water and oxygen, is a crucial component of oxidative stress tolerance (Dempel, 1999; Marchler et al., 1993; Orr & Sohal, 1994; Storz & Tartaglia, 1992; Storz & Imlay, 1999).

To this date, it is not established whether oxidative stress tolerance plays a role in the pathogenesis of *H. capsulatum*. PMNs produce hydrogen peroxide in response to *H. capsulatum* yeasts, and hydrogen peroxide is in turn a substrate of the PMNs’ antimicrobial myeloperoxidase system, yet they are unable to kill the yeasts (Schnur & Newman, 1990). One group reported that short-term assays revealed a fungicidal effect of the oxidative burst of PMNs or macrophages. However, in longer-term assays, *H. capsulatum* was able to overcome human alveolar and peritoneal macrophage mediated oxidative bursts and continued to grow and replicate intracellularly (Bullock & Wright, 1987; Schnur & Newman, 1990). Finally, in light of the demonstrated virulence roles of surface-located or secreted superoxide dismutase or catalase in certain microbial pathogens (Beaman et al., 1983; De Groote et al., 1997; Fang et al., 1999; Garcia et al., 1999; Jeavons et al., 1998; Lopez-Medrano et al., 1995; Radcliff et al., 1997; Rambukkana et al., 1993; Raynaud et al., 1998; Sonnenberg & Belisle, 1997; Spiegelhalder et al., 1993; Vera-Cabrera et al., 1999), it is intriguing that *H. capsulatum* produces an extracellular catalase (the M antigen: Hamilton et al., 1990; Zancope-Oliveira et al., 1999). Taken together, it is likely that these surface-located or secreted antioxidants function as a ‘first line of defence’ against reactive oxygen species produced by the oxidative burst of PMNs or macrophages.

Many of the above-mentioned results are consistent with the hypothesis that catalase, one of the components of the oxidative defence mechanism of *H. capsulatum*, plays an important role in pathogenicity. In an early effort to test this hypothesis, a comparison of the virulence of several *H. capsulatum* isolates that differed in catalase production revealed no correlation between virulence and amount of catalase (Howard, 1983). However, the isolates of *H. capsulatum* were not congenic and probably differed in expression of numerous genes and proteins that could affect virulence. To resolve this problem, it is our long-term goal to compare the virulence of congenic derivatives of *H. capsulatum* that differ only in catalase production. The characterization of three catalase genes and their expression, reported here, represents an essential first step towards this long-term goal.

**METHODS**

**Strains and culture conditions.** The *Histoplasma capsulatum* virulent strain G-217B (ATCC 26032; generously provided by W. E. Goldman, Washington University, St Louis, MO, USA) was used in all experiments. *H. capsulatum* yeast cultures were grown at 37 °C with gentle shaking in 3% (v/v) glycerol or 2% (w/v) glucose HMM medium (Worsham & Goldman, 1988). Cultures of submerged mycelia were grown in the same media with gentle shaking at room temperature. In all experiments, exponentially growing yeast or mycelial cultures were used. H₂O₂ challenges were done with yeast cells grown to an OD₆₀₀ of 3–10 in glucose HMM medium, with H₂O₂ added to a final concentration of 50 mM. Cells were harvested at regular time intervals, up to 6 h after challenge. Under these conditions of H₂O₂ treatment, less than 20% loss of viability was observed (data not shown).

*Escherichia coli* strains DH5α or SOLAR (Stratagene) were used for plasmid transformations.

**DNA purification and analysis.** *H. capsulatum* genomic DNA was isolated according to Woods & Goldman (1992), except that the lysed spheroplast extract was incubated with proteinase K overnight at 50 °C.

For restriction enzyme digestions, PCR reactions or synthesis of radiolabelled probes, plasmid DNA was purified from *E. coli* by the CTAB method (Del Sal et al., 1989). When needed for sequencing, DNA was purified by Qiagen columns. Automated DNA sequencing was done on an Applied Biosystems model 377 instrument in the University of Arkansas for Medical Sciences DNA sequencing facility. Big Dye Terminator chemistry with the Applied Biosystems Dye Terminator Kit version 2.0 was used according to the manufacturer’s instructions. Templates and primers were present in the reactions at 50 ng μl⁻¹ and 0.16 μM, respectively. This system typically gave sequence reads of over 900 bases, with excellent reliability in the first 700 bases. Both strands of each cDNA insert were sequenced in their entirety. Primers used for sequencing reactions were spaced no more than 600 bases apart. The primer sequences and locations in the cDNA insert sequences are available upon request.

**RNA purification and analysis.** Total RNA was extracted from cultures of strain G-217B yeast or mycelia according to a modification of the acid guanidium thiocyanate extraction procedure of Chomczynski & Sacchi (1987). Fifty grams of guanidium thiocyanate was dissolved in 64 ml diethyl pyrocarbonate (DEPC)-treated H₂O₂; then 3·5 ml 0·75 M sodium acetate pH 5·2 and 0·77 ml 2-mercaptoethanol were added. Finally, equal volumes of H₂O₂-saturated phenol and acid guanidium thiocyanate solution were mixed with each other. *H. capsulatum* yeast cells from 20 ml aliquots were resuspended in 10 ml RNA extraction solution and manually
shaken for 5 min with 7.5 ml glass beads (0.5 mm diameter). The bead/extract mixture was frozen at −70 °C prior to further processing. After thawing, the extract was pipetted away from the glass beads and centrifuged for 10 min at 8000 g to pellet cell debris. The supernatant was transferred to a fresh tube and mixed with 0.2 vol. CHCl₃. After vortex mixing, the phases were separated by centrifugation at 8000 g for 10 min and the upper aqueous phase was collected. One volume of 2-propanol was added and the RNA precipitated at −20 °C. After centrifugation at 12000 g for 20 min, the RNA pellet was resuspended in 0.5 ml acid guanidium thiocyanate solution and then mixed with 0.5 ml H₂O-saturated phenol. The procedure was repeated as described above to separate the phases and precipitate the RNA. The RNA pellet was rinsed once with 1 ml 75% ethanol. The final RNA pellet was resuspended in DEPC-treated H₂O and the concentration determined from its absorbance at 260 nm.

RNA samples were electrophoresed in formaldehyde/1% (w/v) agarose (Sambrook et al., 1989), transferred by capillary blotting to Hybond-N membrane (Amersham Pharmacia) and hybridized to radiolabelled probes in the hybridization solution of Church & Gilbert (1984) according to the procedure described by Johnson & Schmidt (1993). Results were obtained both by autoradiography and by phosphorimaging (Molecular Dynamics Storm Phosphorimager, Amersham Pharmacia Biotech). Band intensities were determined from the Phosphorimager data and normalized to the band intensity of the small-subunit rRNA in the same lane. Imagequant 5.1 software (Molecular Dynamics) was used for these measurements.

Protein purification. H. capsulatum yeast cells were suspended at 0.5 g ml⁻¹ in lysis buffer (50 mM sodium phosphate pH 7.5, 250 mM mannitol and 50 mM sucrose) and mechanically broken with glass beads (0.5 mm diameter Zirconium beads; Biospec Prod.) in a Bead-Beater (Biospec Prod.) for five pulses of 30 s, with 10 min cooling on ice between pulses. The cellular debris was removed by centrifugation at 12000 g at 4 °C and the supernatant clarified by centrifugation at 100000 g. The S-100 supernatant was adjusted to 1.5 M ammonium sulfate and clarified by centrifugation at 100000 g. It was then applied to a cyclohexyl-Sepharose column and washed with buffer A (1.5 M ammonium sulfate, 0.025 M sodium phosphate, pH 7). Proteins bound to the column were eluted with a linear gradient between buffer A and buffer B (0.025 M sodium phosphate pH 7). Fractions containing catalase activity were identified by their ability to degrade H₂O₂ (followed as a decrease in absorbance at 240 nm) and were pooled and dialysed against 2 mM sodium phosphate buffer, pH 6.8. The sample was applied to a column of hydroxyapatite and bound proteins were eluted with a linear gradient of 2 to 400 mM sodium phosphate, pH 6.8. Catalase fractions were pooled and the catalase P polypeptide was purified by SDS-PAGE.

The purified catalase protein (57 kDa) was transferred to PVDF membrane (Waters) and fragmented by cyanogen bromide. Four peptide fragments were purified by C₈ reverse-phase HPLC. Sequencing of the purified peptide fragments from their N-termini was performed with an Applied Biosystems pulsed-liquid phase sequencer (model 477A) equipped with a model 120A PTH-amino acid analyser.

Generation of PCR fragment probes from catalase A, B and P genes and isolation of full-length cDNA clones. Details of the H. capsulatum cDNA library and methods for screening it were provided previously (Johnson & McEwen, 1999). For each screening, a small DNA fragment from the desired gene was generated by PCR, radiolabelled with the Random Primers DNA labelling system (Invitrogen Life Sciences) according to the manufacturer’s instructions, and hybridized to plaque lifts of the library (Sambrook et al., 1989).

The CATA probe was generated with degenerate primers that were designed from three highly conserved regions dispersed along the length of the aligned Aspergillus fumigatus and Emericella nidulans catalase A protein sequences. The sequence of the 5’ primer (CatA 5’) was CNATHGAYAYAAR-GCINTAYC; the sequence of the most 3’ complementary strand primer (CatA 3’) was TCRCTRCARTGTCNRCTNCT- CR; and the sequence of a nested complementary primer (CatA 3’B) was CYTCRAANCKRTTNGGCCART. (The mixed-base codes are N = A + C + G + T, R = A + G, Y = C + T, M = A + C, K = T + G, S = C + G, W = A + T, H = A + C + T, B = C + G + T, D = A + G + T and V = A + C + G.) PCR reactions were done with 1 µg G-217B genomic DNA and 2-5 units Taq polymerase (Invitrogen Life Technologies), in standard PCR buffer (supplied by Invitrogen Life Technologies) adjusted to 2 mM MgCl₂ and 200 µM of each DNA primer. Amplification conditions followed a step-down procedure of one ‘hot-start’ cycle at 94 °C for 5 min followed by holding the temperature at 80 °C in order to add the Taq polymerase. The reaction was then continued with six stages of cycling, with each cycle repeated twice: the cycles included a melting step of 94 °C for 20 s; an annealing step which dropped in the order 65, 60, 58, 56, 55 and 52 °C for 20 s for each of the six stages, respectively, and an extension step of 72 °C for 20 s. These six stages were followed by 15 additional stages done with an annealing temperature of 50 °C. The reaction was terminated with a final cycle with steps of 94 °C for 1 min, 50 °C for 1 min and 72 °C for 3 min. The final fragment used as a probe for the cDNA clones in the library was generated by nested PCR reactions, as follows. First, a 600 bp PCR product was generated with primers CatA 5’ and CatA 3’. By use of this product as template with primers CatA 5’ and CatA 3’, a 400 bp PCR product was generated. This final PCR product was gel-purified and ligated with the pCR 2.1 TOPO vector from Invitrogen Life Technologies. The sequence of the insert coded for a peptide that showed significant similarity to a number of catalases. A number of CATA cDNA clones were isolated from the library but all were 5’ truncated. The missing 5’ sequence of the CATA mRNA was generated with the 5’ RACE system (rapid amplification of cDNA ends), according to the manufacturer’s instructions (Invitrogen Life Technologies). RNA for the 5’ RACE amplification was isolated from H. capsulatum cultures challenged with 50 mM H₂O₂, which induces accumulation of CATA mRNA (see Results). The sequences of the gene-specific primers, used in the reverse transcription step of the 5’ RACE reaction, were 5’-CTTTGGTCCAATAGATATGG-3’ (primer GSP-1) and 5’-CAACGTCGATGAGTTTTC-3’ (primer GSP-2). The final product was cloned in the pCR2.1 TOPO vector. Automated DNA sequencing showed that this 650 bp fragment contained 361 bases of missing 5’ CATA mRNA sequence as well as an overlap with the 5’ end of the sequence contained in the originally isolated truncated clones. A full-length CATA cDNA clone was constructed by ligation of a DNA fragment containing the 5’ RACE sequence, with the CATA cDNA clone at an internal ND1 site located at basepair 712 of the full-length reconstructed clone (basepair 351 of the original 5’ truncated cDNA clone). This resulted in the construction of a 2537 bp full-length CATA cDNA clone (GenBank accession no. AF189368). Subsequent isolation and sequencing of a
genomic clone for CATA (C. H. Johnson, unpublished data) confirmed that the reconstructed CATA cDNA is collinear with the genomic sequence, with the exception of introns.

To generate a PCR probe for CATB, primers from the M antigen sequence (GenBank accession no. AF026268) were obtained and a 200 bp fragment was amplified from genomic DNA. This probe was used to screen the cDNA library. A complete cDNA insert of 2600 bp was sequenced and found to contain the coding region for M antigen.

To generate a PCR probe for CATP, degenerate primers were designed from sequences of internal peptides of the purified catalase P enzyme. The primers CATP5′ (ATTCHCAYCRT- NATGCAYYT), CATP3′ (TCYTRRTTINACNCNDATRAA) and CATP3′ (AANGCNCGTYGTYYCCAT) were used in PCR reactions with genomic DNA according to the ‘hot-start’ method described above (see CATA). Sequencing of this fragment confirmed that it contained sequence encoding a catalase enzyme. A full-length cDNA clone was successfully identified in the library after screening with this probe.

Identification of intron locations in the CATA and CATP genes. The presence and locations of introns within the coding regions of the CATA and CATP genes were determined by comparisons of the sizes of products obtained from PCR amplification of the cDNA and genomic DNA as templates. The primers used originally in DNA sequencing, which are spaced at intervals of several hundred bases along each strand, were used for these PCR reactions. When the genomic PCR reactions yielded a larger PCR product than the cDNA PCR, the genomic products containing introns were gel-purified and reactions yielded a larger PCR product than the cDNA PCR, were used for these PCR reactions. When the genomic PCR reactions yielded a larger PCR product than the cDNA PCR, the genomic products containing introns were gel-purified and submitted for automated sequencing. Alignment of these sequences with the cDNA sequences allowed precise identification of intron positions. These results were subsequently confirmed by sequence analysis of large genomic clones for both CATA and CATP (C. H. Johnson, unpublished data).

Phylogenetic and intron-residence analyses. A recent analysis of the core amino acid sequences from 129 monofunctional catalases, which produced a tree with three separate clades of the catalase gene family (Loewen et al., 2000), was extended by newly available sequences including the H. capsulatum CATA, CATB and CATP proteins. The alignment of 210 full-length peptide sequences was created with the multiple-alignment algorithm of Clustal W (http://www.ddbj.nig. ac.jp/htmls/E-mail/clustaw-e.html), with Kimura correction and maximum fixed-gap and gap extension penalties. The confidence levels of the inferred relationships, the datasets included all fungal catalases. In addition, intron residence was mapped on the aligned amino acid sequences from eukaryotic organisms, which generated a positional matrix. The PAUP* program package was used to process this matrix by a full heuristic search inferring maximum parsimony. To estimate the confidence levels of the inferred relationships, the datasets were bootstrapped by 1000 replications. Each search run was performed with stepwise addition of 10 randomly selected sequences and with tree-bisection-reconnection branch swapping by retaining groups with frequencies = 50%. The following catalase sequences are shown in the pruned phylogenetic tree and the tree constructed from analyses of intron residence (following GenBank accession numbers in parentheses and the source organism of the sequences). Animals: CaecleaA (AL034488), CaecleaB (AL034488), CaecleaC (AL034488) — Caenorhabditis elegans; DromeaA (U00145) — Drosophila melanogaster; DicidaA (AF090443) — Dictyostelium discoideum; HomspA (AL035079) — Homo sapiens; RatnoraA (AH004967) — Rattus norvegicus. Archaea: MetarbA (AJ008388) — Methanobrevibacter arboriphilicus; MetharbA (AJ005399) — Methanosarcina barkeri; MetmazaA (TIGR.com) — Methanosarcina mazei. Bacteria: Bacm-zA (TIGR.com, contig. 1484) — Bacillus anthracis; BacubX (NC.000964) — Bacillus subtilis; DeiradaE (AE001825), DeiradaA (AE002037) — Deinococcus radiodurans; MicutaA (P29422) — Micrococcus luteus; MycavieE (L41246) — Mycobacterium avium; PsierB (U89384), PsierC (AE004642) — Pseudomonas aeruginosa; Pseyrf (AE001355) — Pseudomonas syringae; StrccoaA (AL121853), StrccoB (AE004419) — Streptomycetes coelicolor. Fungi: AsfumA (AFU87630) and AsfumB (AFU87850) — Aspergillus fumigatus; AspinR (Z23138) — Aspergillus niger; BotufkaA (Z54346), BotufkB (AF248353) — Botryotinia fuckeliana; CanabalA (A006327) — Candida albicans; CantoA (M18832) — Candida tropicalis; CanbopiA (AB064338) — Candida boidini; ClafulaB (AF222053), ClafulaB (AF222056) — Cladosporium fulvum; ClapuraA (AJ001385) — Claviceps purpurea; CryneoA (http://www-sequence.stanford.edu) — CrynedoA (TIGR.com) — Penicillium ananatum; PencalB (U75451) — Pleurotus ostreatus; PledjaA (U75450) — Pleurotus djamare; PlesajA (AF286997) — Pleurotus sajor-caju; PodansA (AJ011298), PodansB (AJ011309) — Podospora anserina; SaceraA (X13028), SacereT (X04625) — Saccharomyces cerevisiae; SchpomaA (D89126), SchpomX (D55675) — Schizosaccharomyces pombe. Plants: ArathaA (AF021937), ArathaB (AL022053), ArathaC (AF021937) — Arabidopsis thalama; GlymaxA (AF032525) — Glycine max; OrysatA (X61626), OrysatB (D29464), OrysatC (AB020502) — Orozyza sativum; RapsatA (AF031318, AF248491) — Raphanus sativus; RiccomA (D21161), RiccomB (D21162) — Ricinus communis; SoltubB (U27082) — Solanum tuberosum; ZemayA (X12538), ZemayB (X54819), ZemayC (L05934) — Zea mays.
RESULTS

Catalase enzyme activities in H. capsulatum

At the onset of our studies of catalases in H. capsulatum, there was good evidence for multiple catalase isozymes in two other filamentous fungi (Chary & Natvig, 1989; Kawasaki et al., 1997). In H. capsulatum, catalase enzyme activity assays of the growth media in comparison to washed yeast cells indicated substantial catalase activity in both fractions (data not shown). Purification of the protein with catalase enzyme activity from the growth medium revealed an ~87 kDa glycosylated protein likely to be the previously identified M antigen. In contrast, purification of the catalase enzyme activity from the soluble fraction of washed yeast cells revealed a 57 kDa protein (Fig. 1). These results provided initial evidence for isozymes of catalase in H. capsulatum.

Isolation and characterization of the H. capsulatum CATP cDNA clone

A full-length cDNA clone for CATP was isolated from the H. capsulatum cDNA library (GenBank accession no. AF189369). The gene, and the protein it is predicted to encode, catalase P, were so designated because of homology to known peroxisomal catalases (see below).

Southern blotting experiments with the cDNA as probe confirmed the presence of a single CATP gene in the H. capsulatum genome (data not shown). The predicted catalase P protein has 503 residues, a molecular mass of 57 247 Da and an isoelectric point of 7.29.

When we first obtained the CATP sequence, the only fungal peroxisomal catalase sequences available in GenBank were from several Ascomycota, subphylum Saccharomycotina, including Saccharomyces cerevisiae, several Candida species and Hansenula polymorpha. Since then, a CATP-homologous gene from Emericella nidulans (Ascomycota; subphylum Pezizomycotina), termed CATC, was identified (Kawasaki & Aguirre, 2001). In addition, we searched the unannotated sequence data from the Cryptococcus neoformans (phylum Basidiomycota) and Neurospora crassa (Ascomycota; subphylum Pezizomycotina) genome projects. A sequence corresponding to a full-length CATP-homologous gene from C. neoformans, with eight exons and seven introns, was assembled from several sequencing traces obtained from the Cryptococcus genome project server (http://www-sequence.stanford.edu/group/C.neoformans/index.html). The N. crassa gene homologous to CATP is within contig 1.1240, obtained from the N. crassa genome project server (http://www-genome.wi.mit.edu/annotation/fungi/neurospora/). The N. crassa CATP gene contains two exons and one intron. A multiple sequence alignment of the H. capsulatum catalase P with the newly identified fungal catalases, as well as peroxisomal catalases of several Saccharomycotina yeasts, is shown in Fig. 2. The phylogenetic relationships between known or suspected peroxisomal catalases of fungi, yeasts and animals were analysed as described below; these confirm their placement in a separate clade from the CATB and CATA catalases.

Isolation and characterization of the H. capsulatum CATA cDNA clone

The predicted protein sequence of the M antigen revealed it to be more similar to CATB than to CATA of Aspergillus fumigatus. We therefore sought and found an H. capsulatum cDNA clone for the CATA catalase (GenBank accession no. AF189368). CATA encodes a product of 749 residues with a predicted molecular mass of 84643 Da and a predicted isoelectric point of 6.38. The presence of a single CATA gene in the H. capsulatum genome was confirmed by Southern blot analysis (data not shown).

Pairwise comparisons of the H. capsulatum CATA amino acid sequence with CATA sequences of A. fumigatus and Emericella nidulans demonstrated a high level of identity with each (70·2% and 65·6%, respectively). A multiple alignment of seven known fungal CATA sequences confirmed the overall similarity of the enzymes in the CATA group and their placement in a phylogenetic group distinct from the CATB catalases (see below).
Fig. 2. Multiple alignment of the H. capsulatum catalase P enzyme with other small-subunit catalases from fungi. Small-subunit catalases from H. capsulatum (Hc), Emericella nidulans (En), Schizosaccharomyces pombe (Sp), Hansenula polymorpha (Hp), Saccharomyces cerevisiae (Sc) and Candida albicans (Ca) were globally aligned by CLUSTAL W, as described in Methods. Boxshade 3.1 (www.ch.embnet.org) was then used to shade the residues with more than 50% conservation between the examples. Exact identities between residues in the proteins are indicated by black boxes, and similar residues are shaded in grey. Numbering of the residues for each protein is indicated to the left of the protein sequences.
The catalases of *H. capsulatum*

**Fig. 3.** Phylogenetic analysis of the monofunctional catalase gene family: distance-neighbour-joining tree derived from a CLUSTAL W-alignment of the 210 currently available catalase amino acid sequences. The tree was pruned to emphasize residence of fungal catalase family member groups in the major clades. The node that connects the three major clades is indicated with a filled circle. Significant bootstrap values are shown for critical branch points. Fungal branch annotations are given in bold type. The abbreviations used for catalases are based on organism names and catalase isozyme names from each organism and are defined in Methods.

**Phylogenetic and intron residence analyses**

The phylogenetic trees constructed by distance-neighbour-joining or maximum-parsimony inferences yielded nearly identical trees. A distance-neighbour-joining tree that included 210 monofunctional catalase sequences was pruned as described in Methods (shown in Fig. 3) with the aim of emphasizing the phylogenetic relationships between fungal catalases. All catalases from plants and algae are in clade 1 together with a subset of small-subunit catalases from Gram-positive and Gram-negative bacteria. Clade 3 contains only small-subunit monofunctional catalases that are found in diverse phylogenetic groups including bacteria, animals and fungi. Clade 2 contains exclusively large-subunit monofunctional catalases found in bacteria and fungi. Catalases from *Saccharomyces cerevisiae* and other *Saccharomycotina* representatives (such as *Candida*) reside in clade 3 of the phylogenetic tree, closely grouping with animal and the majority of known bacterial catalases (Fig. 3) (Loewen et al., 2000). *H. capsulatum* CATP (HiscapP) and its peroxisomal homologues from other *Pezizomycotina* fungi, including *Cladosporium fulvum*, *Botryotinia fuckeliana* and *Emeriella nidulans*, as well as catalases from *Cryptococcus neoformans* and *Pleurotus* species (of the *Basidiomycota*), also group in clade 3. The fungal CATA and CATB catalases group in separate subclades of clade 2 (Fig. 3). The *H. capsulatum* CATB (HiscapB) protein groups with large-subunit CATB-like sequences in a subclade of clade 2 with predominantly fungal catalases. The *H. capsulatum* CATA sequence (HiscapA) groups with other CATA-like sequences in a subclade with predominantly bacterial catalases.

To refine the relative groupings of the fungal catalases in clades 2 and 3, the number and locations of introns in each gene were analysed. In order to include *H. capsulatum* catalases P and A in these analyses, it was necessary to identify the intron sequences within the CATP and CATA genes. The CATP coding region was found to contain three introns, of 137, 136 and 98 bp (data not shown). The CATA coding region is interrupted by two introns, of 84 and 73 bp (data not shown). The intron positions and sequences for *H. capsulatum* CATB were obtained from GenBank (accession no. AF026268). Intron positions and sequences of other eukaryotic catalase genes were obtained from pertinent GenBank deposits or unfinished genome sequencing projects, and they were mapped on the CLUSTAL W-aligned amino acid sequences. The derived intron matrix was processed with *PAUP* and the results were used to construct a phylogenetic tree inferring maximum-parsimony (Fig. 4). While the grouping of plant (clade 1) and animal (clade 3) catalases implies the existence of single-rooted intron lineages, multiple independent intron lineages are evident for fungal catalase genes in clades 2 and 3 (Figs 3 and 4). Intron lineage 2.4 is congruent with the respective branch of large-subunit fungal catalases in clade 2 (Figs 3 and 4). The three other fungal intron lineages in clade 2 represent catalases that subgroup within the branch of large-subunit bacterial catalases. Analysis of introns of the catalases in clade 3 allowed the definition of two main subgroups of fungal small-subunit catalases. One contains only intronless peroxisomal catalases from *Saccharomycotina* yeasts (*Saccharomyces*, *Hansenula*, *Candida*) and the other subgroup contains intronless catalases (*Schizosaccharomyces*, *Saccharomyces*) and...
catalases from diverse fungi (Cladosporium, Botryotina, Cryptococcus, Emericella, Histoplasma and Neurospora) whose catalase genes have introns. The Cryptococcus catalase gene has seven unique introns, whereas the H. capsulatum CATP and E. nidulans CATC catalase genes, with three and two introns, respectively, belong to the same intron lineage (Fig. 4). The intron in the Neurospora crassa CATP gene is incorporated at

**Fig. 4.** Maximum-parsimony tree constructed from analyses of intron positions mapped onto the alignments of the catalase amino acid sequences. The resulting clades reflect the acquisition (light shades) and loss (dark shades) of introns in an individual intron lineage. The individual intron lineages were grouped according to catalase family structure established by the phylogenetic analysis (see Fig. 3). Abbreviations for catalases are defined in Methods.
position 4 as found in animal catalases, and the Cladosporium fulvum CAT1 catalase gene (Bussink & Oliver, 2001) contains a single intron in a unique position (Fig. 4).

**Regulation of H. capsulatum catalase genes in response to oxidative stress, cell type and carbon source**

The multiple catalase genes of both Neurospora crassa and Emericella nidulans are differentially expressed during oxidative stress and development (Chary & Natvig, 1989; Kawasaki et al., 1997; Kawasaki & Aguirre, 2001). To determine if the H. capsulatum catalase genes are differentially regulated during oxidative stress, we examined the abundance of transcripts for each of the H. capsulatum catalase genes, CATA, CATB and CATP, in response to H₂O₂ challenge. As shown in Fig. 5, CATA RNA abundance showed a marked increase after H₂O₂ challenge. The effect was most pronounced within 1-5 h and nearly abated by 3 h after the challenge. (The amount of CATA RNA was restored to basal levels within 4–6 h after challenge; data not shown.) The CATB and CATP panels of Fig. 5 show that the challenge with H₂O₂ had little or no effect on the abundance of transcripts from either of these genes during yeast-phase growth. This experiment (same strain and growth conditions) was conducted at least five times, with consistent results (data not shown).

We also examined whether the H. capsulatum catalase genes are differentially expressed in yeast versus mycelial cultures. In addition, the effect of glucose versus glycerol as carbon source was examined for both yeast and mycelial cultures, in view of the report that overall catalase activity of H. capsulatum yeast-phase cultures was repressed by glucose (Howard, 1983). As shown in Fig. 6, the CATA gene is differentially regulated according to cell type. CATA RNA is much more

---

**Fig. 5.** Expression of the H. capsulatum CATA, CATB and CATP genes after challenge with H₂O₂. The three sets of panels are Northern blots of H. capsulatum total RNA (25 µg per lane), isolated from cultures at various time points (indicated in hours, above the lanes of each panel) after challenging the cultures with 50 mM H₂O₂. The blots were probed with CATA, CATB or CATP probes, as indicated below the blots. The panels below the CATA and CATB blots are the same blots after they were reprobed with an 18S rRNA probe. The panel below the CATP blot shows 18S rRNA from the ethidium-bromide-stained agarose gel used for the blot. RNA sizes are indicated to the right of the upper panels and were determined from comigration of markers (RNA ladder from Invitrogen Life Technologies).

**Fig. 6.** Cell type and carbon source control of abundance of catalase RNAs. The panels show Northern blots of total RNA (5 µg per lane) hybridized with radiolabelled probes from each of the three catalase cDNA clones and 18S rDNA. The identities of the probes used for each blot are indicated to the left of each panel. The carbon source used during growth (Glu, glucose; Glyc, glycerol) and the phase of growth from which RNA was isolated are indicated at the top. The bottom panel is of total RNA in each lane hybridized with a probe for 18S rRNA and was used for normalization of results.
abundant in mycelia than in yeast, and is not particularly affected by carbon source in either phase of growth. The difference in CATA expression in yeast and mycelia is not a direct consequence of incubation temperature but instead correlates closely with changes in cell morphology between the yeast and mycelial states (J. E. McEwen, unpublished data). The CATB and CATP genes were expressed in both the mycelial and yeast phases, and did not appear to be regulated according to cell type. However, we did not examine the possibility of developmental regulation during asexual conidiation or sexual spore development. Both CATB and CATP showed modest regulation by carbon source. In the yeast cultures, the amounts of both CATB and CATP RNAs were decreased by glucose (2–5-fold and 1.5–3-fold, respectively, depending on the experiment). In mycelial cultures, CATB RNA abundance was increased 2–3-fold by glucose, relative to glycerol, whereas CATA and CATP RNA levels were not significantly changed. These experiments were conducted on three independent sets of RNAs from the same strain and culture conditions, with consistent results (data not shown).

DISCUSSION

Our studies, together with the earlier report of Zancop-Oliveira et al. (1999), demonstrated the existence of three catalase genes in *H. capsulatum* – CATA, CATB (encoding the M antigen) and CATP. Southern blot hybridization experiments indicated that each catalase gene is present in a single copy in the *H. capsulatum* genome. The presence of one or more additional catalase genes in *H. capsulatum*, distinct from CATA, CATB or CATP, is possible, because a recent study (Kawasaki & Aguirre, 2001) of *Emericella nidulans* catalases, for which catalases A and B had already been identified (Kawasaki et al., 1997) revealed not only a homologue to *H. capsulatum* CATP, but enzymic evidence for a fourth protein with catalase activity. The gene for that protein has not yet been reported.

Phylogenetic and intron analyses support a model for a prokaryotic origin for each of the three catalase genes of *H. capsulatum*

In bacteria with multiple catalase genes, the catalases usually group in different clades (Klotz et al., 1997; Loewen et al., 2000). This is fundamentally different from the case of catalases in animals and plants. In animals with multiple catalase genes (e.g. *Caenorhabditis elegans*), the genes group tightly in clade 3 and have likely arisen by an initial duplication of one clade-3-type catalase gene after which each of the loci have evolved independently. All plant catalases reside in clade 1, and multiple genes within an organism are likely to have arisen by an initial duplication of one clade-1-type catalase gene. Plant genomes harbour small catalase gene subfamilies of up to four member genes, and their phylogenetic relationships (Frugoli et al., 1996; Guan & Scandalios, 1996) and gene structures (Frugoli et al., 1998) have been analysed in the recent past. The expression products of these genes can assemble to various different tetrameric isozymes and the regulation of this allows the synthesis of organ-specific catalases (Frugoli et al., 1996). Fungi generally have multiple catalase genes per genome. Some fungi encode catalases that group in a single clade (clade 2 or 3) and others show distribution of their catalases into clades 2 and 3 (Fig. 3). Clade-1-type catalases have not yet been found in fungi.

The three catalases of *H. capsulatum* are classified as monofunctional catalases, which typically display minor peroxidase activity and act upon small organic substrates (Loewen et al., 2000). *H. capsulatum* catalases A and B are large-subunit monofunctional catalases in clade 2 of the phylogenetic tree (Fig. 3). With regard to fungi, the large-subunit monofunctional catalases are, so far, restricted to several representatives of fungi within the Ascomycota subphylum Pezizomycotina, and to a single representative (*Pleurotus ostreatus*) of the Basidiomycota. A BLAST search against available contigs in the *Cryptococcus neoformans* genome project indicated the presence of CATA- and CATB-like sequences. However, due to the presence of numerous introns and incompleteness of the sequences, CATA and CATB sequences for *C. neoformans* were not available for this analysis. Nevertheless, this preliminary information suggests that additional sequences from representatives of all fungal classes will soon be identified when more genome data become available.

Catalase P is a small-subunit monofunctional catalase that groups in clade 3 with animal peroxisomal catalases as well as bacterial and archaeal catalases. While small-subunit monofunctional catalases have been reported previously for *Saccharomyces cerevisiae* and some of its close relatives in the Saccharomycotina (including Candida), the CATP gene from *H. capsulatum* was the first representative of clade 3 to be made public for a fungus of a different subphylum than Saccharomycotina. Subsequently, a gene for a small-subunit monofunctional catalase (CATC) was reported for *Emericella nidulans* (Kawasaki & Aguirre, 2001), and additional small-subunit monofunctional catalase gene sequences are now available in sequence databases (i.e. from Botryotinia, Cladosporium, Cryptococcus and Neurospora). It is likely that this group of catalases will prove to be ubiquitous in fungi, as it is in animals, where these catalases serve as housekeeping hydroperoxidases.

The eukaryotic genes of small-subunit catalases (clade 3) are likely to have evolved from an ancestral, clade-3-type, prokaryotic gene in a primordial eukaryotic cell ancestral to all animals, protozoa and fungi, as articulated in the hydrogenosome hypothesis (Martin & Muller, 1998). In contrast to most of the clade 3 catalases, many of the A- and B-type catalases are secreted and/or inducible upon environmental cues. These properties suggest that these catalases have been acquired as supplements to constitutive housekeeping enzymes. We hypothesize that large-subunit catalase genes (clade 2) were acquired by fungi from prokaryotic
donor(s) by at least two independent lateral gene transfers (representing the A-type and the B-type catalases). Indeed, the amino acid sequences of the CATA catalases [EmenidA (E. nidulans), AspfunA (F. fumigatus), NeucraA (N. crassa), PodansA (P. anserina) and HiscapA (H. capsulatum)] are most similar to those of large-subunit catalases from Deinococcus and Gram-positive bacteria in the genera Mycobacterium, Streptomyces and Bacillus (Figs 3 and 4; Klotz et al., 1997; Loewen et al., 2000). These Gram-positive bacteria are abundant in aerobic soil habitats that overlap with those of the fungi. Furthermore, evidence for lateral transfer of genes from bacteria to fungi has been documented in recent laboratory experiments (Sprague, 1991; Hayman & Bolen, 1993), indicating that molecular mechanisms are in place for such transfer between at least a few modern fungal and bacterial species.

We additionally hypothesize that the independent lateral catalase gene transfers from prokaryote to fungus were followed by intron acquisition. The CATA-like sequences in clade 2 comprise, so far, three independent intron lineages (Fig. 4). One of these three lineages, 2.3, reflects the presence of only one or two small introns that flank the approximately 276 bp large highly conserved catalytic-cleft-encoding region. There are currently two other unique fungal intron lineages among clade 2 catalases with one representative each, which supports our hypothesis that intron acquisition has occurred after lateral transfer and progressed dependent on genome evolution in related taxa. In contrast, the CATB-like sequences form a tight phylogenetic cluster (Fig. 3) and are descendants of one Ascomycota intron lineage (Fig. 4).

The data of Fig. 4 also support the hypothesis that the clade 3 catalases evolved by intron acquisition. The small number of intron loss versus gain events (only three of the plant, two of the fungal and two of the animal catalase introns were lost) supports the proposed polarity of the process: invasion of introns into intronless genes. This is in contrast to the model of intron loss proposed by Frugoli et al. (1998) based on the analysis of selected plant catalase sequences. In addition to analyses of evolutionary distance and intron residence, emerging structural/crystallographic information on catalases from all three clades (clade 3, bovine liver Kat, Micrococcus luteus KatA, Proteus mirabilis KatA, Saccharomyces cerevisiae CatA; clade 2, Escherichia coli KatE, Penicillium vitale CatA; clade 1, Pseudomonas syringae CatF, Listeria seeligeri Kat) supports a model of divergent rather than convergent evolution of the monofunctional catalase gene family (Carpena et al., 2001, and references therein). Similarly, it is unlikely that lateral transfer of catalase genes to prokaryotes has occurred several times from eukaryotic fungi (the large-subunit catalases) and plants and animals or fungi (the small-subunit catalases) as was proposed by Mayfield & Duvall (1996) based on analysis of a small subset of selected sequences.

In addition to their use in phylogenetic analyses, the evolutionary conservation of intron positions provides a useful tool for gene annotation of new genome sequences. A case in point is provided by the analyses of the clade 3 catalases. The original report of CATA from E. nidulans indicated an uninterrupted coding region (Kawasaki & Aguirre, 2001). Our phylogenetic studies (Fig. 3) revealed that intronless CATP homologues were present only in the Saccharomyces; however, the reportedly intronless E. nidulans CATC sequence grouped tightly with the H. capsulatum CATP sequence (Fig. 3). As a result of consultations between one of us (M.G.K.) and Kawasaki & Aguirre (2001), a correction was made to the GenBank deposit of E. nidulans CATC to indicate the presence of two small exons and introns upstream of the originally reported sequence. As shown in Fig. 4, the two introns in E. nidulans CATC are in the same locations in the coding region as the first two introns of the H. capsulatum CATP gene.

Implications of redundant catalase genes and isozymes

The role of catalases in virulence of pathogenic fungi is not currently clear. Disruption of the CATI gene (the CATB homologue) of Aspergillus fumigatus had no apparent effect on H2O2 sensitivity or virulence of the fungus (Calera et al., 1997). However, more recent data on A. fumigatus revealed the existence of at least three catalase isozymes (Takasuka et al., 1999). Similarly, recent studies of Emericella nidulans mutants deficient in both CATA and CATB indicated no role for these catalases in virulence in p47phox−/− mice (a model for chronic granulomatous disease, which creates unusual susceptibility to catalase-positive pathogens) (Chang et al., 1998). However, the recent data of Kawasaki & Aguirre (2001) demonstrated the existence of a third catalase gene CATC, and presented evidence for a fourth catalase activity that remained active in catA catB catC triple mutants of E. nidulans. Hence the redundancy in catalase genes and isozymes may itself be a virulence characteristic and leaves open the question about the role of catalases in fungal virulence. Direct evidence showing that catalases are virulence factors for H. capsulatum is not yet available and awaits the construction of genetic variants or mutants that show altered catalase levels or function.

Expression patterns of the three catalase isozymes of H. capsulatum and implications for fungal virulence

Aside from the additional protection afforded by redundant genes and isozymes, the presence of at least three catalase isozymes in H. capsulatum and other filamentous fungi raises the possibility that differential gene expression and/or function helps to optimize the organisms’ responses to environmental stresses either inside or outside of an animal host. We characterized catalase gene expression of H. capsulatum by Northern blot hybridization analyses, which revealed the amounts of catalase gene transcripts under various environmental conditions. The conclusions therefore refer to catalase gene expression in terms of steady-state levels of RNA. During the mycelial phase of growth under standard conditions, the catalases of H. capsulatum
conditions (i.e., in the absence of any deliberate oxidative challenge), all three catalase genes are expressed. During yeast-phase growth, the CATB and CATP genes are constitutively expressed, but CATA RNA is barely detectable unless there has been a recent oxidative challenge (see below). Although expression of CATB and CATP is 'constitutive' in both mycelia and yeast, we observed modest regulation of the abundance of transcripts from each of these genes by carbon source. CATA RNA abundance showed the greatest change (2–5-fold) according to carbon source, and reproducibly showed a decrease during growth of yeast cells in glucose medium, but an increase during growth of mycelial cells in glucose medium. CATP RNA abundance was affected 1.5–2-fold, showing a decrease in glucose medium in yeast cells. Because abundance of both CATB and CATP RNAs is decreased in yeast cells grown with glucose as carbon source, relative to glycerol, these results are consistent with the observations of Howard (1983), of decreased overall catalase expression in yeast after oxidative stress, seem to be at odds with earlier results of Kamei et al. (1992), who found no increase in catalase activity of \textit{H. capsulatum} yeast cells exposed to H$_2$O$_2$. However, the discrepancy might be explained in several ways, including differences in the amounts of H$_2$O$_2$ used in the two sets of studies, potential ineffectiveness of the native gel zymogram method (Kamei et al., 1992) to detect catalase A activity, the possibility that catalase A is not present in a soluble fraction of the cell, as was assayed by Kamei et al. (1992), and the possibility that catalase A enzyme activity is not proportional to CATA RNA abundance (indicating post-transcriptional control of gene expression).

In addition to differences in RNA levels and patterns of expression of the three catalase genes, it is likely that each isozyme exhibits unique functional properties. It is already known that catalase B from \textit{H. capsulatum} is a secreted enzyme (Zancope-Oliveira et al., 1999). In addition, our preliminary data on catalytic properties of catalase B indicate that it and \textit{Aspergillus niger} catalase R (of the catalase B group) are remarkably stable to substrate inactivation relative to \textit{Saccharomyces cerevisiae} or bovine peroxisomal catalases (J. L. York, unpublished data). These characteristics of the catalase B enzyme are consistent with the hypothesis that it exerts a 'first line of defence' for \textit{H. capsulatum} in animal hosts. Catalase P is likely to be a peroxisomal enzyme, based on its phylogenetic grouping with known peroxisomal enzymes. The location of catalase A in \textit{H. capsulatum} cells is not yet known, although our preliminary data indicate that it is in a particulate (insoluble) fraction of cells (C. H. Johnson, unpublished data).

Other evidence consistent with a virulence function for catalases is provided by the patterns of expression of the catalase genes that we observed \textit{in vitro}, which indicate that all three enzymes are likely to be upregulated in the animal host. Oxidative stress is encountered by \textit{H. capsulatum} yeast cells during exposure to the respiratory bursts of neutrophils and macrophages, and our \textit{in vitro} results suggest that all three catalase enzymes will be synthesized by yeast cells under these conditions. Furthermore, we hypothesize that expression of both CATB and CATP will be elevated in yeast cells that are resident in macrophages. Both genes are upregulated under conditions of limiting glucose \textit{in vitro}, and the fact that genes for glyoxylate cycle enzymes of \textit{Saccharomyces} and \textit{Candida} are induced when these yeasts are inside macrophages suggests that the phagolysosomal compartment of macrophages contains little or no glucose (Lorenz & Fink, 2001).

The results reported here lay the foundation for direct tests of the roles of the \textit{H. capsulatum} catalases in virulence. Antibodies for each catalase enzyme are either already available (Zancope-Oliveira et al., 1999) or under development (C. H. Johnson, unpublished data). The antibodies, together with gene probes and construction of appropriate \textit{H. capsulatum} null mutants or gain-of-function variants, provide the necessary tools to measure and manipulate the expression and function of each protein under pathogenic conditions.

**ACKNOWLEDGEMENTS**

This work was supported by a Merit Review to J.E.M., administered through the Office of Research and Development, Medical Research Service, Department of Veterans Affairs. In addition, aspects of this work were supported by a Walker Research Endowment Award and Pilot Study Grant from the University of Arkansas for Medical Sciences to C.H.J. M.G.K. was supported in part by Research Incentive Funds from the University of Louisville, Office VP Research. We are grateful to Dr Donald Deluca for assistance with catalase assays. The expert technical assistance of Ms Xiaojin Wang is also gratefully acknowledged.
The catalases of *H. capsulatum*

**REFERENCES**


