Identification, molecular characterization, and expression analysis of a DOMON-like type 9 carbohydrate-binding module domain-containing protein of *Coccidioides posadasii*

Jennine M. Lunetta* and Demosthenes Pappagianis

Department of Medical Microbiology and Immunology, School of Medicine, University of California, Davis, California, USA

*To whom correspondence should be addressed. E-mail: jmlunetta@ucdavis.edu

Received 31 October 2013; Revised 3 February 2014; Accepted 7 March 2014

Abstract

Previously, we investigated the effect of N-acetylglucosamine (GlcNAc) on *Coccidioides posadasii* chitinolytic enzymes during *in vitro* spherule-endospore (S/E) phase culture. During those studies, sodium dodecyl sulfate-polyacrylamide gel electrophoresis analysis of supernatants from S/E phase cultures grown in Converse medium with or without added GlcNAc revealed a ~28-kDa band (CFP28), whose abundance was increased by GlcNAc in parallel with the chitinolytic enzymes. Mass spectrometry (MS) of the CFP28 band revealed peptides that matched an open reading frame found in the tentative consensus sequence, TC20325, retrieved from the Dana Farber Cancer Institute *C. posadasii* Gene Index Database. The TC20325 cDNA sequence was used to design internal primers based on MS peptides and a full-length cDNA was isolated using a combination of rapid amplification of cDNA ends and reverse transcription-polymerase chain reaction. The deduced amino acid sequence of the full-length cDNA consists of 231 amino acid residues with a 19 aa signal peptide. The mature protein has a calculated molecular mass of ~24.5 kDa, a theoretical pl of 6.09, and consists of a single DOMON-like type 9 carbohydrate-binding module (CBM9-like-3) conserved domain. The protein shares the highest sequence similarity (~57%) to hypothetical proteins from fungi within the Pezizomycotina subphylum of Ascomycota. Antiserum against a recombinant version of CFP28 recognized native CFP28 in S/E phase cells and culture supernatants. CFP28 mRNA and protein expression were detectable in S/E phase in Converse medium, but were increased in the presence of added GlcNAc. Purified native CFP28 reacted with pooled sera from patients with coccidioidomycosis.

Key words: *Coccidioides*, N-acetylglucosamine, DOMON, carbohydrate-binding module, seroreactive.
Introduction

The amino sugar, N-acetylglucosamine (GlcNAc), is an abundant biomolecule that is involved in many cellular processes [1–4]. GlcNAc can be used as an energy source through its catabolism to fructose-6-phosphate, which can be used for glycolysis [1–4]. Alternatively, through the anabolic pathway, GlcNAc becomes a constituent of macromolecules such as structural polysaccharides, glycoproteins, and glycolipids [1–4]. GlcNAc is also a regulatory molecule that can alter gene expression through various mechanisms [1–4]. These include O-GlcNAc modification of intracellular proteins and induction of genes through cell signaling pathways that induce the enzymes involved in GlcNAc metabolism [1–4].

The in vitro alteration of gene expression by exogenous GlcNAc has been reported for a number of microorganisms, including fungi [3,4]. In particular, the effect of exogenous GlcNAc on gene expression in the opportunistic yeast pathogen, Candida albicans, has been studied in detail [3,4]. The addition of GlcNAc to the extracellular medium of C. albicans cultures induces genes that encode the proteins involved in the uptake, transport, and metabolism of GlcNAc [5,6]. These include the GlcNAc catabolic pathway genes of the nag regulon that encode for the enzymes GlcNAc kinase (HXX1/NAG5), GlcNAc-6-phosphate deacetylase (DAC1/NAG2), and glucosamine-6-phosphate deaminase (NAG1) [7,8]. Together, these enzymes carry out the complete conversion of GlcNAc to fructose-6-phosphate, ammonia, and acetate. Other GlcNAc-induced genes in C. albicans include those that encode for a GlcNAc transporter (NGT1), an amino acid permease (GAP1), and a protein that is thought to be involved in GlcNAc metabolism (GIG1) [9–11].

Some fungal genes that encode for chitinolytic enzymes, such as β-N-acetylhexosaminidase and chitinase, are also inducible by GlcNAc, which is the terminal product of chitin degradation [12–14]. In C. albicans, β-N-acetylglucosaminidase gene (Hex1) expression and enzyme activity are induced by GlcNAc [15–17]. N-acetylglucosamine induction of chitinolytic gene expression and enzyme activity has also been demonstrated in the mycoparasitic fungi Trichoderma species. For example, the expression of the chitinase (Chit33) and the β-N-acetylglucosaminidase (Nag1) genes are increased by GlcNAc in the fungus T. harzianum [18,19].

The study of GlcNAc-induced proteins in pathogenic fungi is of particular interest because they may contribute to pathogenesis [1–4]. For example, a C. albicans mutant defective in the production of the GlcNAc-inducible β-N-acetylglucosaminidase was pathogenic but less virulent than the wild type in a mouse model [20]. Candida albicans mutants deficient in the GlcNAc catabolic pathway genes and, therefore, incapable of using GlcNAc, were highly attenuated in a murine model of systemic candidiasis and showed less adherence to human buccal epithelial cells in vitro [21].

Little is known about the effect of exogenous GlcNAc on the dimorphic pathogenic fungus Coccidioides posadasii. Previously, we observed that the addition of GlcNAc to the extracellular medium of spherule-endospore (S/E) phase cultures of C. posadasii increased chitinolytic enzyme activities, protein levels, and mRNA expression [22,23]. The added GlcNAc increased the concentration of a number of culture filtrate (CF) proteins in parallel with β-N-acetylhexosaminidase 1 (HEX1) and endochitinase 1 (CTS1). These studies have been extended to identify and characterize these GlcNAc-induced proteins. Here, we describe the identification and molecular characterization of one of those proteins [24].

A previously uncharacterized seroreactive ~28-kDa band (CFP28) was observed after sodium dodecyl sulfate-polycrylamide gel electrophoresis (SDS–PAGE) and immunoblot analysis of proteins from S/E phase CFs following supplementation with GlcNAc. The band was sequenced using mass spectrometry, and peptides matched an open reading frame (ORF) of the tentative consensus sequence TC20325. A full-length cDNA encoding CFP28 was cloned and the nucleotide sequence determined. The protein contains a single DOMON-like type 9 carbohydrate-binding module conserved domain, and it has predicted structural features similar to those of the C-terminal carbohydrate-binding module type 9 of xylanase 10A from the bacterium Thermotoga maritima [25–28]. The protein shares the most similarity with putative hypothetical proteins of no characterized function from fungi of the Ascomycota subphylum Pezizomycotina. A recombinant version of the protein (rCFP28) was purified, and antiserum was raised in rabbits. The anti-rCFP28 antiserum reacted with both recombinant CFP28 and the native GlcNAc-induced protein present during in vitro S/E phase growth. We also report on the expression pattern of the CFP28 transcript during the S/E phase under growth conditions in Converse medium with and without GlcNAc supplementation. Purification of native CFP28 from S/E phase CF is also described.

Materials and methods

Preparation of S/E phase cells, culture supernatants, and extracts

The S/E phase was maintained through serial culture of the C. posadasii Silveira strain [29–31]. For synchronized
cultures, endospores of a similar size (small endospores) were used as inoculum [29]. Small endospores were isolated by differential centrifugation (400 × g for 5 min) of 7-day S/E phase cultures. The small endospores that remained in the supernatant were used as inoculum. Converse medium (100 ml) was then inoculated with the small endospores (2 × 10^6 cells/ml) and incubated at 37°C with rotary shaking [30,32,33]. For induction experiments, sterile-filtered GlcNAc (Sigma-Aldrich, St. Louis, MO, USA) was added directly to the culture medium 24 h after inoculation to achieve a final concentration of 0.25% (w/v). Control cultures (Converse medium without added GlcNAc) received an equal volume of sterile Cellgro water (Mediatech, Manassas, VA, USA). S/E phase cells and culture supernatants were collected at 24 h, 48 h, 72 h, and 96 h after inoculation. The 24-h samples were collected before the addition of GlcNAc or water. S/E phase cells were harvested by centrifugation and the supernatants collected. The cells were washed once with sterile water and resuspended in either phosphate-buffered saline containing thimerosal (0.05% w/v) or Buffer RLT (Qiagen, Valencia, CA, USA) containing β-mercaptoethanol (1% v/v). The cell suspensions in Buffer RLT were immediately placed on dry ice and stored at −80°C until used for RNA isolation. Culture supernatants were sterile-filtered using a Durapore (0.22 μm) filtration unit (Millipore, Billerica, MA, USA) and then concentrated using an Amicon ultrafiltration device (Millipore). Protein concentrations were determined using a bicinchoninic acid assay (Thermo Scientific, Rockford, IL, USA). Cellular extracts were prepared as previously described [22].

Protein electrophoresis and immunoblot analysis

Proteins were separated using standard SDS–PAGE methods and then stained with GelCode Blue (Thermo Scientific). Apparent molecular weights of proteins were determined with AlphaView SA software (Protein Simple, Santa Clara, CA, USA). Immunoblot analysis was performed as described previously [34]. Protein samples (25 μg) were run on an SDS–PAGE gel and then transferred overnight at 30 V to an immunoblot polyvinylidene difluoride membrane (Bio-Rad Laboratories, Hercules, CA, USA) in transfer buffer (25 mM Tris, pH 8.3; 192 mM glycine; and 20% methanol). Serum samples from anonymous patients were obtained from the Coccidioidomycosis Serology Laboratory (Davis, CA, USA) in accordance with the policies of the University of California–Davis (UC–Davis) Institutional Review Board. Serum samples from four patients were pooled for each group, that is, negative, coccidioidal complement fixation (immunoglobulin G [IgG]) antibody positive and tube precipitin (IgM) antibody positive.

RNA isolation

Total RNA was isolated with an RNase Plant Mini kit using the method described in the manual for the purification of total RNA from yeast (Qiagen) [22,30]. The S/E phase cell suspension in Buffer RLT was thawed at 37°C, added to an equal volume of glass beads (0.5 mm; BioSpec Products, Bartlesville, OK, USA), and then vortexed repeatedly for 4 min (30-s intervals followed by 1 min on ice). The cell lysate was cleared by centrifugation and added to an equal volume of 70% (v/v) ethanol. The remaining steps, including an on-column DNase digestion, were performed exactly as described in the Qiagen manual. RNA integrity was assessed by electrophoresis through a denaturing formaldehyde (2% v/v) agarose gel (1% w/v). RNA purity and concentration were estimated by measuring the absorbance at 260 nm and 280 nm on a NanoVue spectrophotometer (GE Healthcare, Piscataway, NJ, USA). Poly (A)^+ RNA was prepared using an Oligotex mRNA Mini kit (Qiagen).

Rapid amplification of cDNA ends

The 5′ and 3′ ends of the cDNA encoding CFP28 were mapped by full-length, RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) using a GeneRacer kit with a Superscript III module (Invitrogen, Carlsbad, CA, USA) [22]. Total RNA from 24 h S/E phase cells was used to prepare RACE-ready cDNA as described in the GeneRacer manual. Oligonucleotide primers for RACE were designed based on the peptide fragments generated from mass spectrometric analysis of the ~28-kDa eluted band. In order to obtain the nucleotide sequence encoding the peptides, a tblASTn search was performed with the peptide sequences obtained from mass spectrometry (MS) against the C. posadasii Silveira gene index database at the Dana Farber Cancer Institute (http://www.danafarber.org; accessed November 2010). The tblASTn search retrieved a 1196-bp tentative consensus sequence TC20325 that contained a 231 amino acid ORF (+1695 bp). For 5′ RACE polymerase chain reaction (PCR), the RACE-ready cDNA was amplified with the GeneRacer 5′ primer (supplied with the kit) and a reverse primer (5′-GCTGGCTACGTCAGGTGAAT-3′) designed using the nucleotide sequence of TC20325 that encoded for the amino acid residues in the partial peptide sequence “NSPDVAS”. For 3′ RACE PCR, the RACE-ready cDNA was amplified using the GeneRacer 3′ primer (supplied with the kit) and a forward primer (5′-AAAAACGGCCCTTCTTCTCG-3′) designed using the
nucleotide sequence from TC20325 that encoded for the amino acid residues in the partial peptide sequence “QKRPSLA”. The RACE PCR amplicons were cloned using a TOPO TA Cloning Kit for Sequencing (Invitrogen). Plasmid DNA was isolated from positive Escherichia coli transformants using a QIAprep Spin miniprep kit (Qiagen); the presence and size of insert DNA was determined using restriction enzyme digestion by EcoRI followed by agarose gel electrophoresis. The insert DNA was sequenced using the vector-specific primers M13 forward (-20) and M13 reverse.

**Isolation of full-length cDNA**

A full-length cDNA encoding CFP28 was isolated using a SuperScript III First-Strand Synthesis System (Invitrogen) [22]. The gene-specific primers for reverse-transcription (RT)-PCR were prepared using the nucleotide sequence in the predicted 5′ and 3′ untranslated regions obtained from RACE mapping of the cDNA ends. First-strand cDNA was prepared by RT of 24-h S/E phase poly (A)+ RNA with SuperScript III reverse transcriptase as described in the manual. First-strand cDNA (1 µl) was amplified in a 50-µl reaction mix with 1 µM of each forward (5′-TCTCTCCAAGAAGGTAGCAGC-3′) and reverse (5′-AGGCATCCAATCTGAGAGA-3′) primer. The RT-PCR amplicon was gel-purified using a S.N.A.P. kit (Invitrogen) and cloned into a TOPO-XL vector (Invitrogen). Plasmid DNA was isolated from positive transformants and evaluated for insert DNA by restriction enzyme digestion. Clones that contained insert DNA of the expected size were sequenced using the TOPO-XL vector-specific primers M13 forward (-20) and M13 reverse as well as gene-specific primers based on the partial cDNA sequence obtained from RACE.

**Protein and nucleotide sequence determination and analysis**

SDS–PAGE gels were submitted to the UC–Davis Genome Center Proteomics Facility (http://proteomics.ucdavis.edu) for tandem mass spectrometry (liquid chromatography coupled with mass spectrometry [LC-MS/MS]) based protein identification [34]. All MS/MS samples were analyzed using the X! Tandem search engine (Global Proteome Machine Organization; thegpm.org); Scaffold software (Proteome Software Inc., Portland, OR, USA) was used to validate MS/MS–based peptide and protein identifications. Plasmid DNA was sequenced at the UC–Davis Division of Biological Sciences DNA Sequencing Laboratory (http://dnaseq.ucdavis.edu). Reference nucleotide and protein sequences were obtained from the National Center for Biotechnology Information (NCBI) GenBank database (www.ncbi.nlm.nih.gov), the gene index database at the Dana Farber Cancer Institute (Boston, MA, USA [http://www.danafarber.org]), and the Coccidioides Group Sequencing Project, Broad Institute of Harvard and Massachusetts Institute of Technology (www.broadinstitute.org) [35–38]. Nucleotide and protein sequences were analyzed and manipulated with S.E.R.I.A.L. 2.5 (http://serialbasics.free.fr). The Swiss Institute of Bioinformatics SIB ExPaSy Bioinformatics Resources Portal (http://www.expasy.org) was used for molecular mass and theoretical isoelectric point calculations (Compute pI/MW tool), signal peptide prediction (SignalP 4.0), and O-linked (NetOGlyc 4.0) glycosylation site prediction [39–41]. Subcellular localization was predicted using PSORT II (http://psort.ims.u-tokyo.ac.jp) [42]. The Structure Feature Analysis Tool (http://hive.biochemistry.gwu.edu/tools/sfat) was used for N-linked glycosylation prediction [43]. DIANNA software (http://clavius.bc.edu/~colotelab/DIANNA) was used for disulfide prediction [44]. Conserved domains (CD) were identified with the CD database and CDART tool at the NCBI web site and at the Superfamily web site (http://supfam.cs.bris.ac.uk) [45–47]. A motif scan was performed with MyHits software using all available databases of motifs (http://myhits.isb-sib.ch) [48]. The ClustalW2 program at the European Bioinformatics Institute web site (http://www.ebi.ac.uk) with BLOSUM default settings was used to generate a multiple sequence alignment (MSA) [49]. An MSA image was created with ClustalX2 software [49]. Sequences were only shaded when 70% were similar or identical. Sequence Manipulation Suite software (http://www.bioinformatics.org/sms2/ident_sim.html) was used to calculate percent identity and similarity [50]. Secondary structure and fold recognition predictions were performed with Phyre2 software [51]. Swiss-Model (http://swissmodel.expasy.org) was used for template-based homology modeling [52–54]. Chimera v1.6rc (http://www.cgl.ucsf.edu/chimera) was used for 3D structure file viewing and image output [55]. The NCBI GenBank database accession numbers for the CFP28 cDNA nucleotide and deduced amino acid sequences are JF298211 and AEB21190, respectively.

**Analysis of CFP28 mRNA expression by conventional and real-time RT-PCR**

Temporal expression analysis of the CFP28 transcript in the presence or absence of GlcNAc was performed using conventional and real-time RT-PCR. Conventional
RT-PCR reverse transcription was performed as previously described [22]. The following primers were used for detection of the CFP28 transcripts: forward, 5′-CAAAGACGACATCTGGAAA-3′, and reverse, 5′-GCTGGCTACGTCAGGTGAAT-3′. The ribosomal RNA housekeeping gene was amplified in parallel as an internal amplification standard [22]. The RT-PCR product sizes were 389 bp and 186 bp for the CFP28 and rRNA transcript, respectively. For real-time RT-PCR ([quantitative] qRT-PCR), cDNA was synthesized using a QuantiTect kit (Qiagen) and RNA (1 μg) isolated from S/E phase cells cultured with or without GlcNac. The following primers were used: CFP28: forward, 5′-GCCAGTGAAAGACCCAAGAA-3′, and reverse, 5′-GGCTGGAAGGGTAAGCTGGAA-3′; rRNA: forward, 5′-CGCTTACACCACATCCAGGAA-3′, and reverse, 5′-GCTGGAATTACCGCGGCT-3′. The CFP28 primers generated a 161-bp amplicon. Real-time PCR reactions were performed with Platinum SYBR Green qPCR Supermix (Invitrogen) and an Applied Biosystems 7900HT Fast Real-Time PCR System (Life Technologies). Real-time PCR data were analyzed using the comparative Ct method. The levels of expression of the CFP28 mRNA with and without GlcNac at 48 h, 72 h, and 96 h were normalized to the internal standard, rRNA mRNA, and then the expression levels of treated (GlcNac) cultures vs. untreated (Converse only) were compared.

Expression and purification of a recombinant CFP28 fusion protein

A recombinant version of CFP28 was prepared using a pEXP5-NT/TOPO TA Expression kit (Invitrogen) as described previously [22,34]. The cDNA was amplified with the following forward and reverse primers, respectively: 5′-AGCCCAGCCGCTGAAAGAC-3′ and 5′-CTAAACAAGAATAACATCGCCAAA-3′. The forward primer was designed so that the amplified cDNA excluded the region encoding the predicted signal peptide. An expression plasmid with the correct cDNA sequence in frame with the N-terminal tag was selected for expression experiments (pEXP5-NT_20325exp10). Recombinant protein expression was carried out in chemically competent BL21 E. coli cells (Invitrogen) as previously described [34]. Cultures were analyzed for expression by SDS–PAGE and Western blot analysis with an anti-HisG horseradish peroxidase-conjugated antibody (Invitrogen). The protein was produced in two rabbits using a 57-day protocol at Antibodies Incorporated (Davis, CA, USA). S/E phase CF, spherule cell extract, or spherule cell debris proteins (25 μg) were separated using standard SDS–PAGE. The gel was then transferred overnight at 30 V to a nitrocellulose membrane (Bio-Rad Laboratories) in transfer buffer (25 mM Tris, pH 8.3; 192 mM glycine; and 20% methanol). The membrane was incubated with 5% nonfat dried milk (NFDM) in phosphate-buffered saline (PBS) containing 0.05% (v/v) Tween-20 (PBST) for 1 h. The membrane was incubated for 1 h with anti-rCFP28 antiserum or corresponding preimmune serum diluted 1:100 in PBST containing 5% NFDM. The membrane was washed three times with PBST for 10 min and then incubated for 1 h with HRP-conjugated goat anti-rabbit IgG (H + L) F(ab′)2 antibody (Invitrogen) diluted 1:2000 in PBST containing 5% NFDM. The membrane was washed three times with PBST for 10 min and then developed with diaminobenzidine (ImmPact DAB; Vector Laboratories, Burlingame, CA, USA).

Purification of native CFP28 from culture supernatant

Native CFP28 was purified from supernatant collected from 7-day S/E phase cultures grown in Converse medium supplemented with GlcNac at 24 h after inoculation. The culture supernatant was sterile-filtered using a Durapore (0.22-μm) filtration unit (Millipore). The CF was concentrated approximately six-fold using an Amicon ultrafiltration device (Millipore) and then applied directly to a gel filtration (GF) column (Superdex 200 HiLoad 26/60 Prep Grade; GE Healthcare) attached to an ÄKTA FPLC System (GE Healthcare). The column was equilibrated and the proteins eluted with Cellgro phosphate-buffered saline (Mediatech). Fractions (5 ml) were collected, and an aliquot of each fraction was precipitated with three volumes of acetone and analyzed for the presence of native CFP28 by SDS–PAGE and Western blot analysis with anti-rCFP28 antiserum. Fractions enriched for CFP28 were pooled, dialyzed
against water, lyophilized, and then resuspended in buffer A (50 mM Tris, pH 7.0, containing 50 mM NaCl). The solution enriched for CFP28 was applied to an anion-exchange column (Resource Q; GE Healthcare) preequilibrated with buffer A. The column was washed with buffer A, and a linear gradient of increasing NaCl concentration (50 mM to 1 M NaCl) was used to elute the bound proteins. An aliquot of each fraction was analyzed as described earlier for GF column fractions. Fractions containing a single ∼28-kDa band that reacted with anti-rCFP28 antiserum were pooled. The purified protein was run on an SDS–PAGE gel, and the identity of the purified protein was determined using mass spectrometry.

### Results

**SDS–PAGE analysis of supernatant proteins from S/E phase cultures grown in Converse medium alone or Converse medium supplemented with GlcNAc**

CFs from S/E phase cells grown in Converse medium alone or Converse medium supplemented with GlcNAc were separated by SDS–PAGE, and the protein banding patterns were compared (Fig. 1A). Similar protein banding patterns were observed for each time point analyzed (without or with GlcNAc); however, a number of bands were more prominent in CFs from S/E phase cultures with added GlcNAc compared with CFs without GlcNAc. For example, at 96 h, three protein bands were more abundant in the CFs from GlcNAc-supplemented cultures (96hG CF) compared with Converse medium only (96h CF). The two bands with apparent migrations between the 37-kDa and 75-kDa molecular weight markers were identified as β-1,4-N-acetylhexosaminidase (HEX1) and endochitinase 1 (CTS1) [22,56]. The increased expression in the presence of exogenous GlcNAc for the *C. posadasii* chitinolytic enzymes β-1,4-N-acetylhexosaminidase and endochitinase 1 has been reported previously [22,23]. In addition, immunoblot analysis of 96h and 96hG CF proteins with pooled sera from patients positive for the coccidioidal complement fixation (IgG) antibody showed that two of the three bands were seroreactive (Fig. 1B). The seroreactivity of chitinase with sera from patients with coccidioidal mycosis has also been described [34,57]. Here, our focus was to determine the identity of the protein(s) in the prominent seroreactive band migrating between the 25-kDa and 37-kDa protein standards. The apparent molecular weight of the band was determined to be ∼28 kDa. The protein will be referred to as CFP28 for “culture filtrate protein, ∼28 kDa.”

**Identification of CFP28 using proteomic methods**

Supernatant from S/E phase cell culture with GlcNAc added to the extracellular medium at 24 h after inoculation was collected 72 h later (96hG CF). The 96hG CF proteins (96hG CFP) were separated by SDS–PAGE; the ∼28-kDa band was excised from the gel and then sequenced using mass spectrometry (LC-MS/MS). Thirteen unique peptide sequences were obtained from MS analysis and used to search the *C. posadasii* Silveira gene index database at the Dana Farber Cancer Institute (http://www.danafarber.org; accessed November 2010). A match was made to a 231 amino acid ORF from a tentative consensus sequence, TC20325 (Fig. 1C). The 13 peptides represented 50% coverage of the 231 amino acid–predicted protein. The predicted size of the protein (26.4 kDa) was similar to the apparent molecular weight of CFP28 that was estimated by SDS–PAGE. The predicted protein sequence was not similar to that of any proteins of defined function. Molecular studies were initiated to isolate a full-length cDNA encoding CFP28.

**Isolation of the full-length cDNA encoding CFP28**

The 5′ and 3′ ends of the transcript encoding CFP28 were first mapped by RACE. A full-length cDNA was then isolated using RT-PCR. The 5′ RACE PCR generated three fragments of *Coccidioides*-specific DNA of varying lengths (688 bp, 696 bp, and 714 bp). Each fragment contained a partial ORF (654 bp) that consisted of the amino acid sequence (218 aa) that matched the TC20325-predicted protein. The fragments contained 5′ untranslated regions of varying lengths, which accounted for the size differences of the fragments. The 3′ RACE PCR yielded two fragments (830 bp and 887 bp) of *Coccidioides*-specific cDNA that were identical through 830 bp. Both fragments contained a partial ORF (615 bp) that consisted of 205 amino acid residue, followed by a stop codon and a 3′ untranslated region. The 5′ and 3′ ends contained overlapping sequence, so the largest *Coccidioides*-specific DNA fragments obtained from the 5′ and 3′ RACE PCR were merged to collectively form a 1026-bp cDNA sequence. The merged sequence contained a 60-bp 5′ untranslated region, a 3′ untranslated region of 270 bp, and the largest ORF of 696 bp (Fig. 2A). The cDNA sequence obtained from RACE PCR included the entire 695-bp cDNA sequence encoding TC20325. The merged cDNA sequence was then verified with RT-PCR. A full-length cDNA was then isolated using RT-PCR with primers complementary to sequence in the 5′ and 3′ untranslated regions of the mapped transcript.
Figure 1. Identification of a ∼28-kDa seroreactive protein in culture supernatant from spherule-endospore (S/E) phase cells grown in Converse medium supplemented with GlcNAc. (A) Sodium dodecyl sulfate–polyacrylamide gel electrophoresis gel of culture supernatants collected at 24 h, 48 h, 72 h, and 96 h after inoculation from S/E phase cells grown in the presence or absence of GlcNAc. Lanes 1, 2, 4, and 6, culture filtrates (CFs) from S/E phase grown in Converse medium only; lanes 3, 5, and 7, CFs from S/E phase grown in Converse medium supplemented with GlcNAc (G) at 24 h after inoculation; M, Precision Plus protein standard (Bio-Rad). (B) Immunoblot using pooled sera from patients positive for the complement fixation (immunoglobulin G) antibody and CF collected at 96 h from S/E phase cultures grown in Converse medium alone or Converse medium supplemented with GlcNAc. The numbers to the left of the images indicate the molecular weight in kDa of the adjacent protein standards. The arrows indicate the positions of CFP28, β-1,4-N-acetylhexosaminidase (HEX1), and endochitinase 1 (CTS1). (C) The protein match obtained from mass spectrometric analysis of the ∼28-kDa band, outlined in panel A, is shown. The amino acids that matched the peptides obtained from mass spectrometry are highlighted. This Figure is reproduced in color in the online version of *Medical Mycology*.

A single RT-PCR product that was identical to the sequence produced by RACE was generated. The nucleotide sequence and deduced amino acid sequence were identical to the tentative consensus sequence (TC20325) and translated ORF, respectively. The full-length cDNA lacked a TATA box and a polyadenylation consensus sequence. Comparison of the cDNA sequence to genomic DNA revealed no introns.
Figure 2. The nucleotide and deduced amino acid sequences of the CFP28 full-length cDNA, schematic representation of the conserved domain architecture of CFP28 and sequence alignment of CFP28 with consensus sequences of the DOMON-like conserved domains. (A) The nucleotide and deduced amino acid sequences of the CFP28 full-length cDNA. The numbers to the left of each line indicate either nucleotide or amino acid positions (bold type). The asterisk indicates the stop codon. The 5′ and 3′ untranslated regions are indicated by the lower case letters. The italicized sequence is the predicted cleavable signal peptide (amino acids 1–19). One putative N-glycosylation sequon is double-underlined. The letters in bold type indicate amino acids that matched the peptide sequences generated from mass spectrometry of the CFP28 band shown in Figure 1. The primer sequences are underlined. (B) Schematic representation of the conserved domain architecture of CFP28. CFP28 contains a single DOMON-like type 9 carbohydrate-binding module (CBM9-like-3) conserved domain (40-226 amino acids) and regions similar to CBM9-like 2 (108-212 aa) and DOMON_murB-like (91-121 aa) conserved domains. The triangles indicate the putative ligand-binding sites (ligand may be carbohydrate). (C) Sequence alignment of CFP28 with consensus sequences of the DOMON-like conserved domains. Amino acid sequence alignments of CFP28 with consensus sequences of the DOMON-like conserved domains, CBM9-like-3 (cd09620), CBM9-like 2 (cd09618), and murB-like (cd09627). The hash symbol over the amino acid residues denote putative ligand-binding sites (carbohydrate or, in the case of DOMON_murB-like, heme).
Sequence analysis of the amino acid sequence of the full-length cDNA encoding CFP28

The amino acid sequence deduced from the largest ORF of the full-length cDNA (696 bp) was 231 residues (Fig. 2A). SignalP predicted the first 19 amino acids of the CFP28 to be a cleavable signal peptide. The mature protein has a theoretical molecular weight of 24c495 Da and pI of 6.09. Analysis using NetNGlyc showed one predicted N-glycosylation site at amino acid position 122 (NVTY). NetOGly predicted two O-glycosylation sites at Thr-6 and Ser-31. Dianna software predicted a disulfide bond between Cys-37 and Cys-63. PSORT predicted an extracellular localization for CFP28. One specific hit in the deduced amino acid sequence to the DOMON-like superfamily, specifically, a cbm9-like-3 (cd09620, 40–226 aa, e-value = 4.43e-30) with putative ligand-binding sites at 100, 120, 122, and 190, was revealed by the CDD server (Fig. 2B and 2C). Partial hits to cbm9-like-2 (cd09627, 108–212 aa, e-value = 5.49e-03) and DOMON-murB-like (cd09618, 91–121 aa, e-value = 3.06e-03) domains were also obtained and were present within the larger domains (Fig. 2B and 2C). Motif scans did not yield any matches.

Sequence similarity

A BLASTp search of the NCBI GenBank nonredundant database with the CFP28 amino acid sequence retrieved 124 sequences from 72 organisms (expect value ≤ 10). The 72 organisms included 41 eukaryotes and 31 bacteria. The eukaryotes included 35 fungi (34 ascomycetes, 1 basidiomycete), 4 cellular slime moulds, 1 Capsaspora sp., and 1 Nematostella sp.; the bacteria included 7 firmicutes, 14 betaproteobacteria, and 10 proteobacteria. A select group of sequences retrieved from the BLASTp search are listed in Table 1 along with percentage sequence identity and similarity with the CFP28-deduced amino acid sequence. The sequences retrieved were hypothetical or putative proteins without significant similarity to any characterized proteins of known function. The fungal sequences ranged in percentage identity from 100% with the C. posadasii Silveira sequence from the Broad Institute database (CPSEG_09825) to 42.37% with the hypothetical protein from Magnaporthe oryzae 70-15 (MGG_00281). The proteins that share the most sequence similarity with CFP28 were from fungi of Ascomycota, specifically, the subphylum Pezizomycotina, and, with the exception of the protein from Uncinocarpus reessii (UREG_00568), were from pathogenic fungi. The percentage similarity of CFP28 to proteins from slime moulds and bacteria ranged from 19.25% for Paenibacillus vortex V453 (PVOR_04448) to 12.70% to Franciscella sp. TX077308 (F7308_0777). The CFP28 sequence was aligned with select fungal proteins retrieved by the BLASTp search that shared >42% sequence identity (Table 1, Fig. 3). ClustalW sequence alignment of the similar proteins revealed a number of highly conserved regions (Fig. 3).

Secondary structure and three-dimensional homology modeling

Since the BLASTp search did not retrieve any characterized proteins with significant similarity to CFP28, secondary structure prediction and homology modeling were used to provide some insight into its function. Phyre2 software was used for secondary structure prediction and fold recognition. The secondary structure is shown in the multiple sequence alignment above the amino acid residues (Fig. 3). The secondary structure prediction consisted of 13 β-strands and a single α-helix (excludes predicted signal peptide). Fold recognition with Phyre2 software retrieved the immunoglobulin-like beta sandwich fold of the Family 9 carbohydrate-binding module (CBD9) of the CBD9-like superfamily as the top hit with 100% confidence, 87% sequence coverage (Phe-45 to Val-231), and 15% identity [26]. The second best match retrieved by Phyre2 was chain C of the oxidoreductase, ethylbenzene dehydrogenase (EDH) of the proteobacterium Aromatoleum aromaticum, with 80.6% confidence, 90% sequence coverage (Pro-21 to Ser-213), and 8% identity [58]. Homology modeling with Swiss modeling generated a match with the PDB:1i82a three-dimensional (3D) structure that was used as a template for template-based homology modeling (Fig. 4). The predicted 3D model represented residue 56-231 aa with 13.63% sequence had a QMEAN Z-score of −6.362 and an e-value of 8.9E-24. The PDB:1i82a is the 3D structure of the second C-terminal CBM (CBM9-2) of the endo-1,4-beta-xylanase 10A from the bacterium Thermotoga maritima [26].

Expression analysis of the CFP28 transcript during in vitro S/E phase growth in the presence or absence of GlcNAc

The temporal expression of the transcript encoding CFP28 was evaluated during in vitro S/E phase growth in Converse medium only or Converse medium supplemented with GlcNAc using conventional RT-PCR (Fig. 5A). During S/E phase growth in Converse medium only, the CFP28 transcript was detected at 24 h, 48 h, and 72 h after inoculation but not at 96 h, and the transcript was most abundant at 48 h. After the addition of GlcNAc to the culture at 24 h, the transcript was detected at 48 h and 72 h after
Table 1. Sequence similarities of CFP28 with a select group of putative proteins.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Protein ID</th>
<th>Accession No.</th>
<th>Sequence Length</th>
<th>Identity (%)</th>
<th>Similarity (%)</th>
<th>Class&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eukaryota</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coccidioides posadassi Silveira</td>
<td>CPSG_09826</td>
<td>EFW13611</td>
<td>231</td>
<td>100</td>
<td>100</td>
<td>E</td>
<td>Animal pathogen, dimorphic</td>
</tr>
<tr>
<td>Coccidioides immittis RS</td>
<td>CIMG_00036</td>
<td>EAS35214</td>
<td>231</td>
<td>99.13</td>
<td>100</td>
<td>E</td>
<td>Animal pathogen, dimorphic</td>
</tr>
<tr>
<td>Uncinocarpus reri 1704</td>
<td>UREG_00068</td>
<td>EEP35721</td>
<td>227</td>
<td>67.53</td>
<td>74.46</td>
<td>E</td>
<td>Nonpathogenic saprophyte</td>
</tr>
<tr>
<td>Cordyceps militaris CM01</td>
<td>CCM_00957</td>
<td>EGX87956</td>
<td>227</td>
<td>60.94</td>
<td>72.10</td>
<td>S</td>
<td>Insect pathogen</td>
</tr>
<tr>
<td>Neocallimastix tapetorum varpVI 77-13-4</td>
<td>NECHADRAFT_45731</td>
<td>EEU38858</td>
<td>225</td>
<td>59.66</td>
<td>71.24</td>
<td>S</td>
<td>Animal/Plant pathogen</td>
</tr>
<tr>
<td>Arthrobotrya gyneum CBS 118893</td>
<td>MGYG_03315</td>
<td>EFR00134</td>
<td>222</td>
<td>58.01</td>
<td>70.13</td>
<td>E</td>
<td>Dermatophyte</td>
</tr>
<tr>
<td>Trichophyton rubrum CBS 118892</td>
<td>TERG_05151</td>
<td>EGD88900</td>
<td>222</td>
<td>58.01</td>
<td>69.26</td>
<td>E</td>
<td>Dermatophyte, anthropophilic</td>
</tr>
<tr>
<td>Metarhizium anisopliae ARSEF 23</td>
<td>MAA_04960</td>
<td>EFZ00032</td>
<td>230</td>
<td>57.63</td>
<td>70.76</td>
<td>S</td>
<td>Insect pathogen</td>
</tr>
<tr>
<td>Gibberella zeae PH-1</td>
<td>FG07921.1</td>
<td>XP_388097</td>
<td>221</td>
<td>57.08</td>
<td>70.39</td>
<td>S</td>
<td>Plant pathogen</td>
</tr>
<tr>
<td>Fusarium oxysporum Fo5176</td>
<td>FOXB_07441</td>
<td>EGU82038</td>
<td>221</td>
<td>57.08</td>
<td>68.24</td>
<td>S</td>
<td>Plant pathogen</td>
</tr>
<tr>
<td>Penicillium marneffei ATCC 18222</td>
<td>PMAA_00948</td>
<td>EEA23291</td>
<td>227</td>
<td>56.60</td>
<td>68.09</td>
<td>E</td>
<td>Animal pathogen, dimorphic</td>
</tr>
<tr>
<td>Trichoderma atroviride IMI 206040</td>
<td>TRIATDRAFT_296843</td>
<td>EHK51130</td>
<td>227</td>
<td>56.22</td>
<td>69.96</td>
<td>S</td>
<td>Mycorrhiza</td>
</tr>
<tr>
<td>Panascoccioides brasiliensis Ph03</td>
<td>PABG_05543</td>
<td>EKE23332</td>
<td>243</td>
<td>56.15</td>
<td>68.03</td>
<td>E</td>
<td>Animal pathogen, dimorphic</td>
</tr>
<tr>
<td>Acremonium capsulatus NAa1</td>
<td>HCAG_03027</td>
<td>EDN06424</td>
<td>238</td>
<td>55.88</td>
<td>66.81</td>
<td>E</td>
<td>Animal pathogen, dimorphic</td>
</tr>
<tr>
<td>Magnaporthe oryzae 70-15</td>
<td>MGG_00281</td>
<td>XP_368963</td>
<td>229</td>
<td>42.37</td>
<td>57.20</td>
<td>S</td>
<td>Plant pathogen</td>
</tr>
<tr>
<td><strong>Fungi/Basidiomycota</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melampsora larici-populina 98AG31</td>
<td>MELLADRAFT_109437</td>
<td>EGG03282</td>
<td>238</td>
<td>43.28</td>
<td>60.08</td>
<td>Plant pathogen</td>
<td></td>
</tr>
<tr>
<td><strong>Amoeboza/Dictyostelida</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polysphondyllum pallidum PN500</td>
<td>PPL_07946</td>
<td>EFA79121</td>
<td>234</td>
<td>16.22</td>
<td>27.03</td>
<td>S</td>
<td>Slime mould</td>
</tr>
<tr>
<td>Dictyostelium fasciculatum</td>
<td>DFA_11379</td>
<td>EGG13618</td>
<td>236</td>
<td>15.83</td>
<td>25.35</td>
<td>S</td>
<td>Slime mould</td>
</tr>
<tr>
<td><strong>Prokaryota/Bacteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Firmicutes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paenibacillus vortex V453</td>
<td>PVOR_04448</td>
<td>EFU43203</td>
<td>211</td>
<td>19.25</td>
<td>30.54</td>
<td>Endospore-forming, swarming motility</td>
<td></td>
</tr>
<tr>
<td>Chloroflexi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roseflexus sp. RS-1</td>
<td>RoseRS_4252</td>
<td>ABQ92595</td>
<td>221</td>
<td>19.25</td>
<td>32.64</td>
<td>Green, nonsulfur, gliding, filamentous</td>
<td></td>
</tr>
<tr>
<td>Acidobacteria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candidatus Solibacter usitatus Ellin076</td>
<td>Acid_6151</td>
<td>221</td>
<td>18.75</td>
<td>30.00</td>
<td></td>
<td>Terrestrial</td>
<td></td>
</tr>
<tr>
<td><strong>Bacteroidetes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclobacterium marinum DSM745</td>
<td>Cycoma_1187</td>
<td>YP_004773190</td>
<td>243</td>
<td>17.27</td>
<td>34.14</td>
<td>Marine</td>
<td></td>
</tr>
<tr>
<td>Halococcus brescia DSM1100</td>
<td>Halby_0157</td>
<td>AEE48070</td>
<td>216</td>
<td>16.33</td>
<td>31.02</td>
<td>Green sulfur, rod shaped, aquatic, nonmotile</td>
<td></td>
</tr>
<tr>
<td>Rumella slabiformis DSM19594</td>
<td>Ruml_2250</td>
<td>AEH8662</td>
<td>220</td>
<td>15.92</td>
<td>26.12</td>
<td>Nonmotile, aquatic, ring shaped</td>
<td></td>
</tr>
<tr>
<td><strong>Proteobacteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Franciscella sp. TX777308</td>
<td>F7308_0777</td>
<td>AEB3704</td>
<td>210</td>
<td>12.70</td>
<td>27.05</td>
<td>Pathogen, purple, nonsulfur, nonmotile</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>The Sequence Manipulation Suite website (http://www.bioinformatics.org/sms2/ident_sim.html) was used to determine percentage identity and similarity.

<sup>b</sup>E, eurotiomycetes; S, sordariomycetes.
Figure 3. Multiple sequence alignment of CFP28 (Cp) with representative conserved fungal hypothetical proteins of *Coccidioides immitis* (Ci), *Uncinocarpus reesii* (Ur), *Paracoccidioides brasiliensis* (Pb), *Ajellomyces capsulatus* (Ac), *Trichophyton rubrum* (Tr), *Arthroderma gypseum* (Ag), *Cordyceps militaris* (Cm), *Metarhizium anisopliae* (Ma), *Trichoderma atroviride* (Ta), *Nectria haematococca* (Nh), *Fusarium oxysporum* (Fo), *Penicillium marneffei* (Pm), *Melampsora larici-populina* (Ml), and *Magnaporthe oryzae* (Mo). The GenBank accession numbers for the proteins are listed in Table 1. An asterisk (*) indicates an amino acid identity; a colon (:) indicates a conserved substitution; and a dot (.) indicates a semiconserved substitution. The predicted secondary structure of CFP28 is shown above the corresponding amino acid sequence. An arrow (→) indicates a β sheet and “hhh” indicates an alpha helix.
Figure 4. Comparative three-dimensional (3D) model of CFP28 and alignment of CFP28 with the C-terminal CBM9–2 of Thermotoga maritima xylanase 10A. The homology model was generated by the Swiss model program using PDB:1i82a as a template (A). PDB:1i82a is the C-terminal CBM9–2 of T. maritima xylanase 10A [26]. The alignment of CFP28 with the template is shown below the 3D structures. The secondary structure is indicated by “s” for β sheets and “h” for alpha helices (B). This Figure is reproduced in color in the online version of Medical Mycology.
Figure 5. Expression analysis of CFP28 mRNA and protein during spherule-endospore (S/E) phase growth in the presence or absence of supplemental GlcNAc. (A) Ethidium bromide-stained gels of RT-polymerase chain reaction (PCR) products from reactions to detect CFP28 and rRNA transcripts using conventional RT-PCR. Lanes 1–4, mRNA isolated at 24 h, 48 h, 72 h, and 96 h after inoculation from S/E phase cultures grown in Converse medium only at 24 h, 48 h, 72 h, and 96 h after inoculation; lanes 5–7, mRNA isolated at 48 h, 72 h, and 96 h after inoculation from S/E phase cells grown in Converse medium supplemented with GlcNAc at 24 h. The numbers to the right of the ethidium bromide-stained gels indicate the size in bp of the RT-PCR product. (B) Analysis of relative abundance of the CFP28 transcript using quantitative RT-PCR. The chart shows the relative fold change in CFP28 transcript levels in treated (GlcNAc) vs. untreated (Converse only medium) cultures at 48 h, 72 h, and 96 h. (C) Western blots of S/E phase culture supernatants, spherule cell extracts, and cell debris with anti-rCFP28 antiserum. Lanes 1–4, culture supernatants, spherule extracts, and debris from S/E phase cultures grown in Converse medium only at 24 h, 48 h, 72 h, and 96 h after inoculation; lanes 5–7, culture supernatants, spherule extracts, and debris from S/E phase cells grown in Converse medium supplemented with GlcNAc at 24 h after inoculation.
inoculation but not at 96 h, and the transcript was most abundant at 72 h. The level of the transcript was similar at 48 h in control and treated cultures, while at 72 h, the transcript was more abundant in the presence of GlcNAc. The expression of the internal housekeeping gene rRNA remained relatively constant over all time points evaluated. The relative fold change of the transcript expression levels was measured using qRT-PCR (Fig. 5B). The expression of the transcript in Converse medium supplemented with GlcNAc was compared with expression in Converse medium only, which was given an arbitrary value of 1. The 48-h and 48-h G time points showed no difference in the transcript levels, while at 72 h, a large relative fold (∼30-fold) increase in the transcript was observed. At 96 h, a half-fold decrease was observed for the transcript generated in the presence of GlcNAc.

Preparation and purification of a recombinant version of CFP28

The cDNA encoding the mature CFP28 protein (without the signal peptide) was cloned into a pEXP5-NT/TOPO vector and expressed in BL21 E. coli cells as an N-terminal polyhistidine-tagged fusion protein. The recombinant CFP28 fusion protein (rCFP28) was localized to the bacterial inclusion bodies and not the soluble fraction. A partially purified fraction of the rCFP28 was obtained by isolation of inclusion bodies from bacterial cells followed by solubilization under denaturing conditions. The recombinant protein fraction was further purified under denaturing conditions with a nickel-chelating column attached to an AKTA FPLC system. A single Coomassie blue-stained band with an apparent migration on SDS–PAGE gel of ∼27 kDa (includes a 2-kDa N-terminal tag) that reacted with the anti-HisG antibody was observed after FPLC purification of rCFP28 (Supplementary Fig. 1A and 1B). MS analysis of rCFP28 generated 71 unique peptides with 94% coverage (200/212 aa [CFP28 without 19 aa signal peptide]; Fig. 1C).

Production and reactivity of antiserum against rCFP28

The FPLC-purified recombinant version of CFP28 was used to immunize rabbits for antiserum production. The reactivity of the rCFP28 antiserum was evaluated by Western blot analysis (Supplementary Fig. 1 and 5). The anti-rCFP28 antiserum reacted with rCFP28, while the corresponding preimmune serum showed no reactivity (Supplementary Fig. 1B). The anti-rCFP28 antiserum also reacted with a ∼28-kDa band in S/E phase cells and culture supernatant (Fig. 5). The corresponding preimmune serum only showed reactivity with a ∼50-kDa band in S/E phase fractions (data not shown). Anti-rCFP28 antiserum did not react with the ∼50-kDa band (data not shown). S/E phase CFPs or cells were separated by SDS–PAGE, and a Western blot was performed with antiserum raised against CFP28 (Fig. 5). A strong reaction was observed for a ∼28-kDa band with the CFP28 antiserum that was not seen with the paired preimmune serum. These results showed that the antiserum raised to the recombinant CFP28 also reacted with the native protein from S/E phase cells and the extracellular medium.

Expression of CFP28 during S/E phase in the presence or absence of GlcNAc

Western blot analysis with the anti-rCFP28 antiserum was used to evaluate the temporal expression of the CFP28 protein during in vitro S/E phase growth in Converse medium alone or Converse medium supplemented with GlcNAc (Fig. 5). CFP28 was detected at each time point measured in both filtrate and cells from S/E phase cultures with or without GlcNAc, although the protein was more abundant in the presence of GlcNAc. In the extracellular medium, CFP28 steadily increased over 72 h and then dropped slightly at 96 h in the absence of GlcNAc; in the presence of GlcNAc, CFP28 increased steadily over 96 h. In S/E phase cell extracts and debris, very little protein was detected in the absence of GlcNAc, and the abundance was highest at 48 h after inoculation. The temporal pattern of CFP28 in S/E phase cells was similar to that seen in CF; however, the protein was more abundant in the presence of GlcNAc.

Purification of native CFP28 from CF collected from S/E phase growth in the presence of GlcNAc

Native CFP28 was purified from 7-day culture supernatant collected from S/E phase cells grown in Converse medium supplemented with GlcNAc. The CFPs were separated using size-exclusion chromatography, followed by anion-exchange chromatography of fractions enriched with CFP28. The CFPs eluted in two large peaks from the column, with CFP28 eluting in both peaks in a trailing pattern during the salt gradient. CFP28 co-eluted in both peaks with a number of other proteins. In the first peak, CFP28 co-eluted with β-N-acetylhexosaminidase; in the second peak, it co-eluted with 1,2-α-mannosidase and endochitinase 1. The fractions enriched for CFP28 were grouped into two peaks, pooled into enriched fractions, and then separated by charge on an anion-exchange column. An SDS–PAGE gel of purified CFP28 is shown in Figure 6. The native protein was purified to a single band as judged by SDS–PAGE.
Figure 6. Purification of native CFP28 from culture supernatant collected from spherule-endospore phase cultures supplemented with GlcNAc and serological response. (A) Coomassie blue-stained sodium dodecyl sulfate–polyacrylamide gel electrophoresis gel of purified native CFP28. (B) Western blots of purified native CFP28 with anti-rCFP28 antiserum (left) and preimmune serum (right). (C) Serological response to purified native CFP28. Immunoblots using sera from patients positive for the coccidioidal complement fixation (immunoglobulin [Ig] G) antibody (left), precipitin (IgM) antibody (center), or pooled negative sera (right). Lane 1, native CFP28; M, Precision Plus protein standard (Bio-Rad). The numbers to the left of the gel or blot indicate the size in kDa of the adjacent protein standard. (D) The protein match obtained from mass spectrometric analysis of the native CFP28 band shown in panel A. The amino acids that matched the peptides obtained from mass spectrometry are highlighted. This Figure is reproduced in color in the online version of *Medical Mycology*. 

HNLASILTGL LAATVYSAG LPEREGKRP GLAVPRCPDK ATASPDKSPF EKAPPETGV DLSCTPFATQ TFKAFDETH FYFDPKHRTH CDIWKCYVCE AFIYHOTNDR QTYFEEFVSP HNVTYOTFVY NPSKVRKEGA WSRDVAFBFN VTPFPPGIVL V
analysis (Fig. 6A). The band was reactive with anti-rCFP28 antiserum (Fig. 6B). The band was sequenced by mass spectrometry, which yielded 20 unique peptides representing 74% coverage of the mature CFP28 protein. These findings confirmed the identity of the CFP28 protein (Fig. 6D).

Serological response to purified native CFP28

The serological response to purified native CFP28 was evaluated with pooled sera from patients with coccidioidomycosis. Purified native CFP28 was run on an SDS–PAGE gel, then an immunoblot analysis was performed with pooled sera from patients positive for the complement-fixation (IgG) antibody or the precipitin (IgM) antibody (Fig. 6). Purified native CFP28 (Fig. 6A) was reactive with pooled sera from both IgG- and IgM-positive patients (Fig. 6C). No reactivity was observed with sera negative for CM (Fig. 6C).

Discussion

A number of studies have shown that fungal chitinolytic enzymes such as β-N-acetylhexosaminidase and chitinase are induced by chitin or chitin degradation products such as GlcNAc [12–14]. Based on those studies, we investigated the effect of GlcNAc on chitinolytic enzymes in C. posadasii. It was determined that supplemental GlcNAc increases the expression of chitinolytic enzymes, such as endochitinase 1 (CTS1) and β-N-acetylhexosaminidase 1 (HEX1), during in vitro S/E phase growth in Converse medium [22,23]. Here, we describe the identification and molecular characterization of a protein (CFP28) that showed increased expression in the presence of added GlcNAc, in parallel with CTS1 and HEX1, under the same in vitro S/E phase growth conditions [24].

SDS–PAGE analysis of culture supernatant proteins from S/E phase cells grown in Converse medium or Converse medium supplemented with GlcNAc revealed a number of bands that showed increased abundance in the presence of added GlcNAc, including a ~28-kDa band (CFP28). MS analysis of the ~28-kDa band yielded peptides that matched partial amino acid sequences of a translated ORF present in the tentative consensus sequence, TC20325, that was retrieved from the C. posadasii gene index database. Using primers based on the peptide sequences and molecular biology techniques, a full-length cDNA encoding the protein was isolated, cloned, and sequenced. The deduced amino acid sequence of the full-length cDNA consists of 231 amino acid residues with a 19 aa residue signal peptide. The mature protein has a calculated molecular mass of ~24.5 kDa, a theoretical pI of 6.09, and was a perfect match with the TC20325 ORF protein sequence. The CFP28-deduced amino acid sequence is very similar to hypothetical proteins of no characterized function from a variety of fungi, including U. reesii, Nectria hematococca, Cordyceps militaris, Metarhizium anisopliae, and Trichophyton rubrum. The deduced protein consists of a signal peptide and a single DOMON-like type 9 carbohydrate-binding module conserved domain (CBM9-like-3). The CFP28 protein is predicted to be structurally related to the C-terminal CBM9 (TmCBM9–2) of the xylanase 10A from the bacterium, Thermotoga maritima [25,26,59,60].

The CFP28 protein has a limited taxonomic distribution. A BLASTp search of the NCBI GenBank nonredundant protein database retrieved similar sequences from only 72 organisms. The sequences with the highest similarity to CFP28 were, with a few exceptions, from fungi that belong to the Pezizomycotina subphylum of Ascomycota, specifically, the classes Eurotiomycetes, Sodariomycetes, and Dothideomycetes of the Leotiomyceta superclass. Most of the highly similar sequences present in the Pezizomycotina fungi belong to the orders Onygenales (class Eurotiomycetes) and Hypocreales (class Sodariomycetes). Interestingly, in Eurotiiales, the sister order of Onygenales, the protein was present in Penicillium marneffei but absent from Aspergilli. Similar proteins were absent from yeast (Saccharomycoitina). The nonpathogenic saprophyte, U. reesii, excluded, the highly similar hypothetical proteins in fungi were from pathogens of plants, insects, animals, and fungi. In addition to the fungal proteins, CFP28 also shares low sequence similarity with hypothetical proteins from a select number of slime moulds and bacteria.

The DOMON-like type 9 carbohydrate-binding module domain present in CFP28 belongs to the DOMON superfamily, which is a diverse group of ligand-binding domains of the immunoglobulin fold that interact with sugars and heme through a common mode [27,28,47]. The DOMON domain proteins have been grouped into families based on shared sequence and structural features [28]. The families include the sugar-binding CBD9-like members such as those found in bacterial xylanases, including TmCBM9–2, heme-binding members such as fungal cellobiose dehydrogenase, and the EDH γ-cytochrome domain and a functionally obscure family called “Gibberella zeae FG07921.1-like” [28]. The G. zeae FG07921.1-like family contains hypothetical proteins that were retrieved by a BLASTp search with CFP28 against the NCBI nonredundant database (Table 1). These include the proteins from the pathogenic filamentous fungi Magnaporthe oryzae (MGG_00281), G. zeae (FG07921), and C. immitis (CIMG_00536) and the bacteria Roseiflexus sp. (RoseRS_4252) and Candidatus Solibacter ustatus (Acid_6151). The domain architectures of the G. zeae FG07921.1-like family of...
proteins are a signal peptide plus a single DOMON domain (SIG+DOMON) or a single DOMON domain [28]. CFP28 shares high sequence similarity (>-57%) and domain architecture (SIG+DOMON) with the fungal members of the G. zeae FG07921.1 family. The predicted ligand-contacting residues in the proteins belonging to the G. zeae FG07921.1-like family of DOMON domains are different from the characterized heme- or sugar-binding DOMON proteins [28].

*Coccidioides posadasii* and *C. immitis* are the causative agents of coccidioidomycosis, also known as San Joaquin valley fever. In addition to studying the basic biology of *Coccidioides*, our laboratory is involved in research to develop a vaccine against *Coccidioides* [61,62]. A number of characteristics of CFP28 make it a good target to investigate as a potential vaccine candidate against coccidioidomycosis. First, it is expressed during the parasitic S/E phase of *Coccidioides*, which has been determined to be the best source of protective antigens [61]. Second, it is a secreted protein; proteins with cell wall or extracellular localizations are the first proteins encountered by the host during natural infection and are more likely to be protective antigens. Next, the protein elicits a serological response, and a number of known protective antigens against *Coccidioides* are seroreactive proteins. Another feature that makes CFP28 a good vaccine candidate is that it does not appear to share significant sequence similarity to human proteins, making it less likely to generate an autoimmune response in humans. Taken together, all of these features make CFP28 a good choice as a vaccine candidate against *Coccidioides*.

**Acknowledgments**

The technical assistance of Jason Sanchez and Tiffany Young is greatly appreciated. Molecular graphics and analyses were performed using the University of California–San Francisco (UCSF) Chimera package. Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at UCSF, with support from the National Institutes of Health (National Center for Research Resources, grant 2P41RR001081; National Institute of General Medical Sciences, grant 9P41GM103311).

**Declaration of interest**

The authors report no conflicts of interest. The authors alone are responsible for the content and the writing of the paper.

**Supporting information**

Supplementary material is available at Medical Mycology online (http://www.mmy.oxfordjournals.org/).

**References**


