In vitro and In vivo Proteome Analysis of Coccidioides posadasii

by

Setu Kaushal

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2015 by the
Graduate Supervisory Committee:

Douglas Lake, Chair
D. Mitchell Magee
Douglas Chandler
Jeffery Rawls

ARIZONA STATE UNIVERSITY

December 2015

ABSTRACT

Coccidioidomycosis (valley fever) is caused by inhalation of arthrospores from soil-dwelling fungi, *Coccidioides immitis* and *C. posadasii*. This dimorphic fungus and disease are endemic to the southwestern United States, central valley in California and Mexico. The Genome of *Coccidioidies* has been sequenced but proteomic studies are absent. To address this gap in knowledge, we generated proteome of Spherulin (lysate of Spherule phase) using LC-MS/MS and identified over 1300 proteins. We also investigated lectin reactivity to spherules in human lung tissue based on the hypothesis that coccidioidal glycosylation is different from mammalian glycosylation, and therefore certain lectins would have differential binding properties to fungal glycoproteins. Lectin-based immunohistochemistry using formalin-fixed paraffin-embedded human lung tissue from human coccidioidomycosis patients demonstrated that Griffonia simplificonia lectin II (GSL II) and succinylated wheat germ agglutinin (sWGA) bound specifically to endospores and spherules in infected lungs, but not to adjacent human tissue. GSL II and sWGA-lectin affinity chromatography using Spherulin, followed by LC-MS/MS was used to isolate and identify 195 proteins that bind to GSL-II lectin and 224 proteins that bind to sWGA lectin. This is the first report that GSL II and sWGA lectins bind specifically to *Coccidioides* endospores and spherules in infected human tissues. Our list of proteins from spherulin (whole and GSL-II and sWGA binding fraction) may also serve as a Coccidioidal Rosetta-Stone generated from mass spectra to identify proteins from 3 different databases: The Broad Institutes *Coccidioides* Genomes project, RefSeq and SwissProt. This also serves as a viable avenue for proteomics based diagnostic test development for valley fever. Using lectin chromatography and LC MS/MS, we identified over 100 proteins in plasma of two patients and six proteins in urine of one patient. We also identified over eighty fungal proteins isolated from spherules from biopsied infected lung tissue. This, to the best of our knowledge, is the first such example of detecting coccidioidal proteins in patient blood and urine and provides a foundation for development of a proteomics based diagnostic test as opposed to presently available but unreliable serologic diagnostic tests reliant on an antibody response in the host.

DEDICATION

This dissertation is dedicated to my parents. I love you dearly.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to the chair of my committee and my mentor Dr. Douglas Lake, for his guidance, encouragement and for his demand for excellence throughout the years. I am also very grateful to the members of my committee, Dr. Mitch Magee, Dr. Alan Rawls and Dr. Douglas Chandler for their time, encouragement, and expertise throughout this project. Dr. Magee also generously provided materials for this study which were pivotal for conducting this research.

I would like to acknowledge the Mayo Clinic Proteomics and Bioinformatics team at Rochester, especially Dr. Surendra Dasari. Dr. Thomas Grys at Mayo Clinic Scottsdale, provided funding and guidance throughout this study.

This work would not have been possible without the help of my friends and former and current members of the lab, Yvette Ruiz, Paul Hanavan, Dr. Kwasi Antwi and Yasmynn Chowdhury. I would also like to thank the staff members of School of Life Sciences for their support. Lastly, I want to acknowledge all my past mentors including Dr. Janet Duerr. I still think fondly of my time as her student.

Finally, my warmest thanks goes to my best friend and husband, Amol, for his unfailing support and love. I am lucky to have my wonderful daughter Anagha, who inspires me and keeps me strong. Lastly, I owe many thanks to my parents and my brother. Their constant love has sustained me throughout my life.

Thank you all, now and always.

TABLE OF CONTENTS

Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Life cycle and epidemiology of *Coccidioides* species

The two closely related species *Coccidioides immitis* and *Coccidioides posadasii* cause Coccidioidomycosis (Valley Fever), a respiratory illness acquired through inhalation of airborne arthroconidia (spores) of this fungus. The genus *Coccidioides* belongs to the order of Onygenaceae and is a member of the phylum *Ascomycota*.

The first reports by Posadas of the fungus infecting humans dates back to the 1890s in parts of South America and Southwestern North America. Those reports described autopsies of patients who had numerous nodules on the lungs and other organs including liver, spleen, and skin. An attribute of this fungus is that it exists in two distinct morphological states depending on the growth conditions (Figure 1). In soil, it occurs in a filamentous mycelial stage and can fragment into nonsexual arthroconidia spores (2-5 µm). Under drying conditions, the arthroconidia disarticulate and may be distributed via aerosols when the soil is disturbed (Nguyen,C. 2013). After hyphal spores are inhaled into lungs, they transform into spherules (up to 120 µm in diameter). (Figure 1). A spherule divides internally into endospores (2-4 µm) and eventually ruptures, each of which can become spherules and the cycle repeats. Endospores can enter the environment and undergo transformation into the saprophytic stage as mycelia once again (Tsang,C.A. 2010; Goto,M. 2007; Fisher,F.S. 2007).

Coccidioides is endemic in desert soils in the southwestern United States, including the California central valley and also certain parts of northern Mexico, Central and South America (Figure 2) (Brown,J. 2013). From 1998-2011, the number of reported cases in Arizona surged from 1,474 to 16,467, which represents an increase in incidence of 16% each year (Centers for Disease Control and Prevention (CDC) 2013). An estimated 150,000 valley fever infections occur annually in the United States. The highest number of coccidioidomycosis cases are reported in

Arizona and California. In 2012, over 17,000 cases were reported in the U.S., about 75% of which were reported in Arizona and 25% of which were reported in California (Sunenshine,R.H. 2007; Hector,R.F. 2011; Centers for Disease Control and Prevention (CDC) 2013).

There is a wide spectrum of clinical illness associated with valley fever and although most cases are self-limiting and do not require medical intervention, disseminated disease can occur in some cases (about 1 %) sometimes leading to death (Galgiani,J.N. 2005). Acute pulmonary coccidioidomycosis is notable for substantial morbidity, lost productivity and overall low quality of life. A 2007 survey conducted by Arizona Department of Health Services showed that patients experienced a median of 120 days of illness, 14 days of missed work, 9 days of missed school and 47 days of missed activities of daily living. Cases of coccidioidomycosis are shown to have a significant impact on the healthcare system as well. The Arizona Department of Health Services reported that hospitalization costs for valley fever in 2012 were more than $100 million (for 1700 hospital admissions).

Clinical presentation and diagnosis in humans

Approximately 60% of infected individuals do not need/seek medical attention and remain asymptomatic, while disease severity in the remaining 40% ranges from mild flu-like respiratory symptoms to severe extra-pulmonary life-threatening disease(Ampel,N.M. 2010). Among patients in endemic areas with community-acquired pneumonia (CAP), an estimated 15-29% have acute coccidioidal infection (Chang,D.C. 2008; Kim,M.M. 2009). Approximately 5% of the infected individuals have dermal hypersensitivity (erythema nodosum or erythema multiforme). Coccidioidomycosis may last a few weeks to several months and patients commonly complain of fever, fatigue, night sweats, cough, chest pain, dyspnea, headache and skin rash. Severity of disease correlates with multiple risk factors, such as age, gender, ethnicity, occupation and immune status. Higher incidence of valley fever is reported in certain ethnic groups, including African-Americans and Filipinos. Individuals working in occupations that involve high exposure to dust, such as agricultural or construction workers, are also at increased risk There is a higher

incidence of disease in males as compared to females and those over 65 years in age (Brown,J. 2013). Immunocompromised individuals such as those with HIV/AIDS or ones taking immunosuppressive drugs are at a greater risk of developing a disseminated disease (Ampel,N.M. 2007). Spherules in the lung expand in size as large as 100-200um and then rupture, releasing endospores, each of which can grow into another spherule. In about 1% cases, endospores can disseminate hematogenously, beyond the lung to other tissues and organs where they may grow and cause disseminated disease.

Necrotizing granulomata in the lung are the hallmark of the immune response to coccidioidomycosis. Lymphocyte clusters in the perigranulomatous region are predominantly B cells with interspersed CD4+ T lymphocytes and fewer CD8+ T lymphocytes (Ampel,N.M. 2007; Ampel,N.M. 2003). Both suppressive and stimulatory cytokines such as Il-10 and interferon- are produced within coccidioidal granuloma. The human immune response during coccidioidomycosis involves the development of delayed-type hypersensitivity and cellular immunity (Ampel,N.M. 2007). Cell-mediated immunity is protective in host and is mediated as dermal hypersensitity in form of erythema nodosum or erythema multiforme. These rashes resolve in a small amount of time. A skin test that measures a cellular response is the newly re-introduced Spherusol, It tests the ability of spherule-derived coccidioidin to induce delayed-type hypersensitivity in patients and is more than 98% accurate (Johnson,R. 2012). Anticoccidioidal humoral antibody in an infected host is believed not to provide protection but indicate level of pathogenicity of the fungus (Saubolle,M.A. 2007). However, in patients with compromised immune systems and even in otherwise healthy individuals with no known immune suppression, immune responses to the fungus may be delayed or aberrant. Therefore, although detection of antibodies may be useful in diagnosis of the disease but absence of it does not rule out the possibility of an infection.

Currently available diagnostic tests for coccidioidomycosis include culture, microscopy and serology. *Coccidioides* grows easily in culture and can be visualized in about 7 days but culturing the fungus is only as good as the clinical specimen. Microscopic detection of the fungal spherules and endospores can be done in fixed biopsied tissue with variety of stains such as

Grocott-methenamine-silver and hematoxylin-eosin .However, this diagnostic technique is not very sensitive and reliable (Saubolle,M.A. 2007; Pappagianis,D. 1990). Other forms of diagnosis include a chest X-ray, computed tomography (CT) and magnetic resonance imaging (MRI). Chest radiography aids in clinical diagnosis and following the progression or resolution of the disease. However, radiography findings are often non-specific and cannot alone reliably diagnose the Coccidioidomycosis. It can be challenging to differentiate coccidioidal lung nodules from other pulmonary nodules caused by lung cancer or tuberculosis by CT scans or MRI. Furthermore, in some patients with disseminated disease, radiographs may not show evidence of previous pulmonary disease (Chang,D.C. 2008; Ampel,N.M. 2003).

Serology tests that detect antibodies to the fungus are the most commonly used. However, diagnosis is difficult even with a serological antibody-based blood test because the immune response to *Coccidioides* may be delayed or even non-existent in some individuals . Several serologic tests have been developed, including the tube precipitin (TP), complement fixation (CF), immunodiffusion tube precipitin (IDTP), immunodiffusion complement fixation (IDCF) and IgM/IgG ELISA immune assays (EIA) (Ampel,N.M. 2010). The tube precipitin (TP) and immunodiffusion tube precipitin (IDTP) assays utilize the TP antigen. These assays detect the presence of IgM (anti-TP) in bodily fluids and demonstrate positive results in early stages of illness (within 1-3 weeks after onset of symptoms). The complement fixation (CF) and immunodiffusion complement fixation (IDCF) assays utilize the CF antigen, which is a heat-labile chitinase. These assays detect the presence of IgG (anti-CF) in bodily fluids and typically demonstrate positive results in later stages of illness (2-3 weeks or more after onset of symptoms). Higher titers indicate severe forms of disease. ID assays are specific but early detection is still a major challenge. IgM and IgG enzyme immunoassays (EIAs) are also widely used for diagnosis although in some studies it has been deemed non-specific (Blair,J.E. 2013). All the above described serologic tests rely on the host immune response. Although antibodies are useful for diagnosis, cellular immunity against the fungus indicates a favorable prognosis, as the host mounts a granulomatous reaction around the spherules (Kirsch,E.J. 2012; Durkin,M.

2008). Cellular immunity is considered protective for life unless the patient becomes immunosuppressed, a condition which increases the risk of re-activation or re-infection (Ampel,N.M. 2003; Ampel,N.M. 2001) . During such prolonged periods between diagnosis and treatment, the likelihood of complications of disease increases.

There is a dire need of diagnostic tests that provide a definitive and early diagnosis. Misdiagnosis has many implications. Symptoms of Coccidioidomycosis are commonly confused with that of pneumonia that is caused by bacterial or viral agents. For this reason, patients are routinely prescribed antibiotics. The patients, hence, receive unnecessary (usually multiple courses) anti-bacterial treatment that may cause disruption of normal flora leading to opportunistic infections such as *Clostridium difficile*. Furthermore, multiple rounds of antibiotics contributes to development of drug resistance and patients also endure side effects of these drugs. In a study done in Arizona in 2007, it was found that patients exhibiting symptoms of coccidioidomycosis waited for more than a month before seeking medical care. On average, it took an additional five months, including three or more visits to a physician, before the patient was correctly diagnosed. More than 50% of these individuals sought care from an emergency room, 40% were admitted to a hospital for at least one night, and in 25%, disease necessitated ten or more visits to a physician (Tsang,C.A. 2010). During prolonged periods of non-diagnosis, additional tests, which can be both invasive and expensive, may be sought out by either the physician or the patient. Finally, delayed diagnosis may increase the likelihood of disease-related complications. With the importance of definitive, early diagnosis in mind, it is evident that a rapid, robust diagnostic test would help patients and their healthcare providers.

Detection of disease biomarkers in Human plasma using Mass spectrometry

The human plasma proteome holds the promise of revolution in disease diagnosis (Marimuthu,A. 2011; Nanjappa,V. 2014). Blood plasma is often the primary clinical specimen, and contains highly abundant proteins like albumin and immunoglobulins, moderate-to-low abundant proteins such as tissue proteins (leakage markers) as well as very low abundance proteins such

as cytokines. These proteins exist in a dynamic range of more than 10 orders of magnitude in concentration. The classical plasma proteins are highly abundant and are primarily secreted by intestines and liver and have an extended residence time in plasma while peptide (and protein) hormones and cytokines have a relatively shorter residence time and abundance in plasma (Valdivia,L. 2006; Anderson,N.L. 2002).

At the high abundance end, albumin secreted by liver has a half-life of 21 days and is present at a concentration of 35-50 mg/ml. On the other hand, low abundant cytokines such as interleukin 6 are present in the range of 0-5 pg/ml. Tissue leakage proteins in plasma may imply serious pathology that can be detected in plasma. Cardiac myoglobin (Mb) is an example of such a protein clinically measured for diagnosis of cardiac infarction. Its concentration in plasma increases from 1-85 ng/ml to 20-1100ng/ml in patients.

More than half of the proteins (total of close to 10000 identified at present) in plasma fall in a molecular weight range of 80kDa to 10 kDa suggesting that they persist in plasma without rapid loss into urine. Urine serves as a subset of plasma and contains approximately 50ug/ml of protein/peptides from at least 75 proteins. It constitutes an impressive proteome and peptidome (Marimuthu,A. 2011). Urine proteome till date consists of 1823 proteins. Peptides below the kidney size cut off are collected in urine and provide a complementary picture of many events at the low mass range of the plasma proteome. Urine is also an easily collectable non-invasive clinical specimen. Both host and pathogen proteins may be broken down by proteases leaving peptide fragments behind that circulate in plasma and/or urine (Schiess,R. 2009; Richter,R. 1999).

Tandem mass spectrometry coupled to liquid chromatography (LC-MS/MS) is a powerful analytical tool that allows identification of proteins in complex biomaterials like plasma and urine (Pisitkun,T. 2007). The first step is the proteolytic digestion of proteins using an enzyme like trypsin which hydrolyzes peptide bonds following lysine or arginine. These trypisnized samples are the subject to high pressure chromatography (HPLC) which stratifies the peptides according to hydrophobicity. This serves the purpose of spreading out the delivery time of the peptide to the

electrospray element of mass spectrometer which leads to identification of more peptides in a complex sample. The first MS spectrum is thus formed which consists of a readout of mass/charge (m/z) of the various peptides. The second MS spectrum is generated when peptides undergo another level of analysis initiated by targeted fragmentation of the selected “parent” peptide ion. This creates complementary b- and y type ions series of fragments. The b-ions contain the NH2 terminus through the cleavage site; the y-ions contain the COOH terminus through the cleavage site. The difference in *m*/*z* between adjacent b- or y-series peaks is exactly the residue mass of the amino acid present in one fragment, but absent in the other. Thus the peptide sequences are determined which can be matched back to the database of choice organism. When multiple peptides are identified from a protein, the confidence of identification increases. However, even when a single peptide is identified from any given protein, other approaches such as target decoy analysis can further the confidence of identification (discussed in-depth in experimental procedures).

LC-MS/MS is hence a great means of discovering unknown proteins in a sample like plasma. In general, a particular peptide ion may be only sporadically detected unless it is derived from one of the most abundant proteins in a given sample. To circumvent this problem, targeted proteomics (multiple reaction monitoring or MRM) can be used to identify specific peptides of interest on a complex sample. Furthermore, initial processing of sample before the proteolytic digestion (including stratification on SDS-PAGE) can be applied to attain the desired resolution in identification of proteins of interest in a complex sample (Pisitkun,T. 2007).

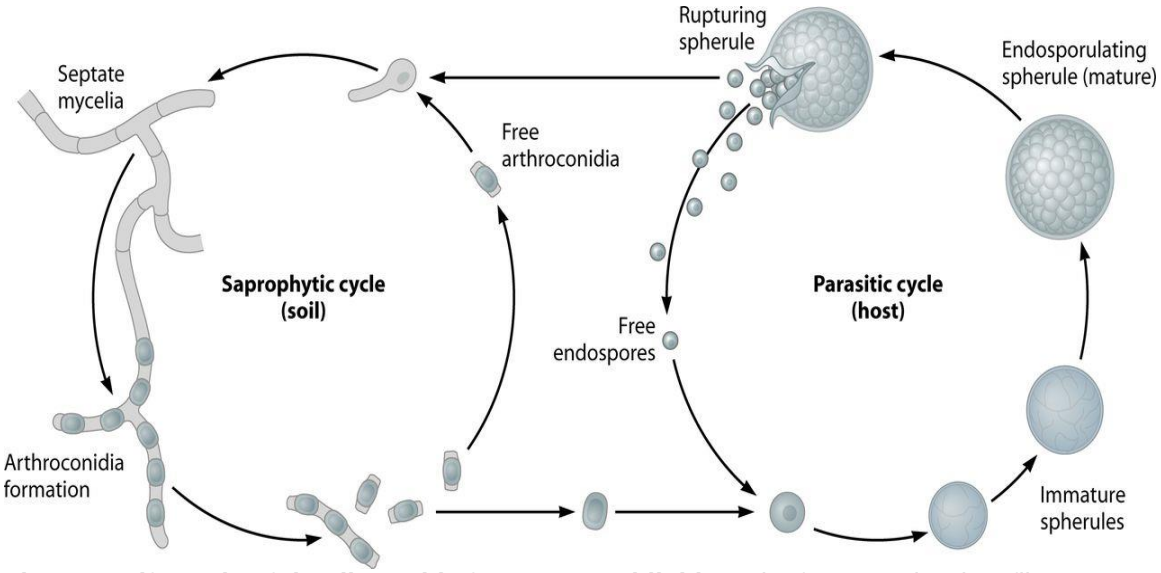

Figure 1. Life cycle of the dimorphic fungus *Coccidioides*. The fungus exists in a filamentous mycelial state in soil. When conditions of growth decline, asexual spores called arthroconidia are formed which can, upon inhalation, cause coccidioidomycosis in a variety of mammalian hosts. In the host, the fungus forms spherules which consist of multiple endospores. These endospores are released when spherules rupture and are dispersed in the lung tissue. These can also undergo transformation back into the saprophytic cycle. (Nguyen,C. 2013).

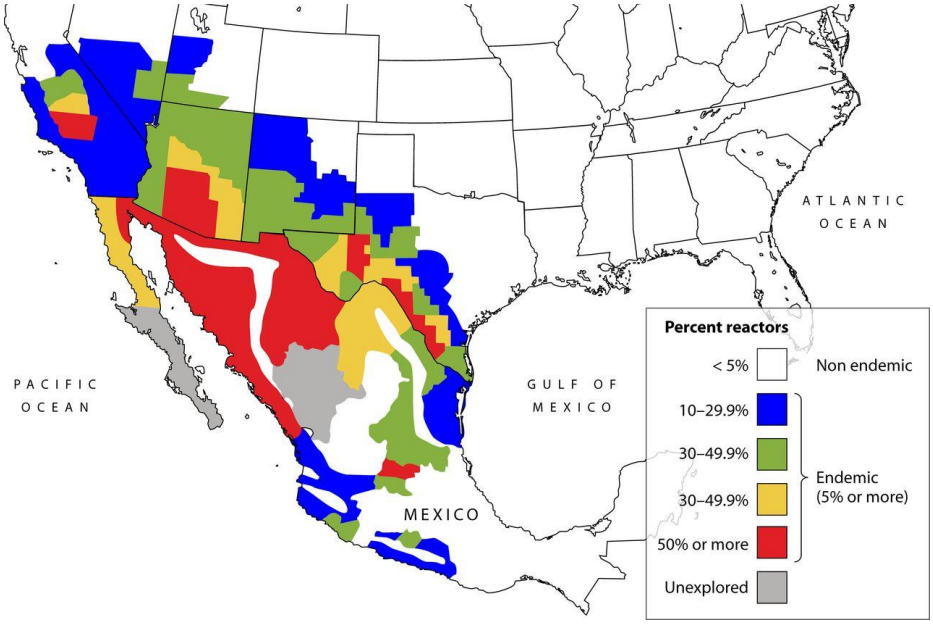

Figure 2. Epidemiology of *Coccidioides*. *Coccidioides* is endemic in the desert soils of southwestern part of the United States including California, Arizona and New Mexico as well as Mexico and Parts of Central America. Most cases are reported in Arizona and California (Nguyen,C. 2013).

CHAPTER 2

SPHERULIN PROTEOME AND LECTIN-BINDING GLYCOPROTEOME OF *COCCIDIOIDES POSADASII*

Abstract

Prior to our studies, there was no available proteomics data for the fungus, *Coccidioides*. From a preparation of Spherulin (parasitic phase protein lysate) we generated a proteome consisting of 1390 protein entries. We also investigated the lectin binding glycoproteome subset of this proteome. We hypothesized that coccidioidal glycosylation patterns may be different from that of mammals and that some lectins would selectively bind to fungal glycoproteins in a background of mixed human and fungal proteins. Using lectin-based immunohistochemistry on formalin-fixed paraffin-embedded (FFPE) human lung tissues, we demonstrated that N-acetylglucosamine (GlcNAc)-binding lectins *Griffonia simplificonia* lectin II (GSL-II) and succinylated wheat germ agglutinin (sWGA) bind to endospores and spherules (parasitic phase) in infected lungs, but not to adjacent human tissue. We then performed GSL-II and sWGA affinity chromatography followed by LC-MS/MS to identify lectin-binding glycoproteins from a coccidioidal spherule preparation (Spherulin). A total of 195 glycoproteins from Spherulin were bound to GSL-II whereas 224 glycoproteins bound to sWGA, with an overlap of 145 glycoproteins between the two lectins. Coccidioidal proteome from Spherulin was generated using the putative sequences present in three *Coccidioides* databases: RefSeq, SwissProt and The Broad Institute's *Coccidioides* Genome project.

Overview

Fungi possess different glycosylation enzymes (glycosyltransferases and glycanases) than mammalian cells (Tsang,C.A. 2010; Goto,M. 2007). Therefore, many of their glycan structures are different from mammalian glycans and can be used to differentiate between host from fungal

glycoproteins (Martin,K. 2007; Goto,M. 2007). Although little is known about glycosylation in *Coccidioides sp.*, *Candida albicans* as well as other fungi have been shown to have mannosyltransferases that are not present in mammalian cells (Mora-Montes,H.M. 2010). We have taken advantage of the differences in fungal and mammalian glycosylation by identifying lectins that specifically bind to the *Coccidioides* spherule (parasitic phase) and not to host (human) proteins.

Lectins have remarkable specificity for various glycan structures and they are often used to capture glycoproteins for analysis. We focused on two lectins, GSL-II and sWGA, in this study. GSL-II is a dimeric glycoprotein that has a unique binding site for α- or β-linked GlcNAc residues on the nonreducing termini of oligosaccharides. sWGA is a succinylated derivative of WGA (derived from *Triticum vulgaris*), a dimeric protein with an affinity to N-Acetylglucosamine dimers and trimers. In this study, we demonstrate that GSL-II and sWGA lectins specifically bind to *Coccidioides* endospores and spherules in infected human tissues. We also studied Spherulin, spherule phase lysate of this fungus. Further, we performed lectin-affinity chromatography and LC-MS/MS to isolate and identify coccidioidal glycoproteins from Spherulin.

Experimental Procedures

Preparation of Spherulin

Coccidioides posadasii (strain Silvera) spherule-phase cells were maintained in continuous culture at 40°C, 20% CO_2, with continuous shaking at 120 RPM in modified Converse medium (Cox,R.A. 1987). In brief, cultures were initiated by seeding flasks of Converse medium with ~ 1-5 X 10^5 arthrospores/ml and at 3-4 day intervals the cells collected by centrifugation, washed in sterile distilled water and stored at 4 °C in 0.5% formalin in water. Prior to fixing in formalin, the cells were checked by microscopy to ensure that the culture was mixed-phase spherules and endospores by morphology and for purity by culture on glucose-yeast extract agar plates. The spent medium was supplemented to 0.5% formalin and stored at 4°C. The spent medium contained antigens elaborated during cellular growth (Spherulin filtrate, SPH-F). The SPH-F antigens were concentrated using a 10,000 MW ultrafiltration membrane. The collected mixed-phase spherules

and endospores were processed to obtain the lysate antigen preparation (Spherulin lysate, SPH-L). To release the internal antigens, the fungal cells were processed in an ice-cooled Beadbeater using 0.5 mm glass beads for 5 minutes. Cellular debris was removed by centrifugation (6,000 X g, 10 minutes) and the supernatant collected. Both SPH-F and SPH-L were lyophilized and stored at -80 °C until use. Spherulin L and F were combined in these studies and will be referred to as "Spherulin."

Spherulin analysis using LC-MS/MS

20 ug of Spherulin was suspended in SDS sample loading buffer (50mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 1% 2-mercaptoethanol, 12.5mM EDTA and 0,02% bromphenol blue) and heated to 95°C. The proteins were then separated by 12% SDS-PAGE and stained with Bio-Safe Coomassie G-250 Stain. Each lane of the SDS-PAGE gel was cut into six equal size slices, placed in a 0.6-ml polypropylene tube, destained twice with 375 µl of 50% acetonitrile (ACN) in 40 mM NH_4HCO_3 and dehydrated with 100% ACN for 10 min. After removal of ACN by aspiration, the gel pieces were dried in a vacuum centrifuge at 60 °C for 30 min. Trypsin (250 ng; Sigma) in 20 µl of 40 mM NH_4HCO_3 was added, and the samples were maintained at 4 °C for 15 min prior to the addition of 50 µl of 40 mM NH_4HCO_3. The digestion was allowed to proceed at 37°C overnight and was terminated by addition of 10 µl of 5% formic acid (FA). After further incubation at 37 °C for 30 min and centrifugation for 1 min, each supernatant was transferred to a clean polypropylene tube. The extraction procedure was repeated using 40 µl of 0.5% FA, and the two extracts were combined. The resulting peptide mixtures were purified by solid phase extraction (C18 ZipTip) after sample loading in 0.05% heptafluorobutyric acid:5% FA (v/v) and elution with 4 µl of 50% ACN:1% FA (v/v) and 4 µl of 80% ACN:1% FA (v/v), respectively. The eluates were combined and dried by vacuum centrifugation and 6 µl of 0.1% FA (v/v) was added followed by sonication for 2 min. The sonicated samples were briefly centrifuged and 2 µl of the sample was subsequently analyzed by mass spectrometry as described below.

Lectin-based Immunohistochemistry (IHC)

Formalin-fixed paraffin embedded (FFPE blocks) were obtained under a human subjects protocol approved by the Mayo Clinic Institutional Review Board. IHC was performed using FFPE tissue blocks from seven patients with Valley Fever. 5µm tissue sections from seven patients who underwent either lobectomy, wedge resection or excisional biopsy from a skin lesion (wrist) were used for IHC. Tissue sections on slides were blocked in Alkaline Phosphatase/Horseradish Peroxidase Block (SurModics, Cat# APHP-0111-01) for 15 minutes followed by Carbo-Free Blocking Solution (Vector Laboratories, Cat# SP-5040) for 1 hour. Biotinylated lectins were obtained from Vector Laboratories (Cat# B-1215, B-1025S). Preliminary experiments were performed to optimize the lowest concentration of lectin that showed positive staining, which was 2ug/ml for both GSL-II and sWGA. Biotinylated lectins bound to tissue sections were detected with streptavidin (SA) coupled to horseradish peroxidase (HRP) using Diaminobenzidine (DAB) as substrate. Sections were washed with 1X PBS (3 times for 5 minutes each) between blocking, incubation with lectin, detection with SA-HRP, and staining with DAB. Tissue was counterstained using hematoxylin (Santa Cruz Cat# SC-24973). GSL-II and sWGA were inhibited with serial dilutions of chitin hydrolysate, a concentrated solution of GlcNAc (Vector Labs, Cat# SP-0090). *Phaseous vulgaris* erythrolectin (PVE), a lectin that binds Galβ4GlcNAcβ2Manα6, was used as a negative lectin control.

Lectin-based Inhibition ELISA

Spherulin was coated onto a 96-well flat-bottom ELISA plate at 1ug/ml in PBS for one hour at room temperature. Wells were blocked with 1% carbo-free BSA in PBS for an additional hour. Two-fold (starting from 1uM) dilutions of non-biotinylated sWGA and GSL-II were used to challenge the binding of biotinylated GSL-II and sWGA, respectively. For instance, non-biotinylated sWGA was incubated with biotinylated GSL-II for 10 minutes prior to adding to the plate. Non-biotinylated lectin dilutions started at 1 uM, and biotinylated lectins were held constant at 10nM. The mixture was then added to the plate for one hour. PVE was used as a negative lectin control. Bound lectins

were detected with a 5000-fold dilution of SA-HRP (Thermo-Pierce, Cat # 21130) in PBS. Plates were washed three times with PBS containing 0.05% Tween-20 (PBST) between coating, blocking, incubation with lectin, detection with SA-HRP, and addition of 3, 3', 5, 5'-Tetramethylbenzidine (TMB) (Becton-Dickinson, Cat# 555214). 1N H_2SO_4 was used to stop the HRP enzyme and the plate was read in a Molecular Diagnostics plate reader at 450nm using SoftmaxPro software. Percent of control was calculated using the following formula: (O.D. biotinylated lectin in presence of non-biotinylated inhibitor) / (O.D. biotinylated lectin in the absence of inhibitor) x 100 = percent of control.

Lectin Affinity Chromatography

GSL-II and sWGA coupled to agarose beads were purchased from Vector Labs (Cat# AL1213 and AL1023S) and used to affinity-purify glycoproteins from Spherulin. 500 ug of Spherulin dissolved in PBS was applied to lectin-agarose columns (0.5ml bed volume). Spherulin starting material and column flow through were saved for subsequent SDS-PAGE analysis. Ten bed volumes of PBS were used to wash the column of unbound lysate. The last 200ul of PBS wash was saved for mass spec analysis to ensure that glycoproteins were not non-specifically washing off the column. Then three bed volumes of "Glycoprotein Eluting Solution for GlcNAc Binding Lectins" (Vector Labs, Cat# ES5100) were used to elute glycoproteins bound to GSL-II and sWGA-Agarose. The elutions were collected, and concentrated by ultrafiltration to 50ul using Amicon Ultra 0.5ml 3 KDa cutoff centrifugal filters (Cat # UFC500396). Protein content in the concentrated eluates were measured using the Micro BCA Protein Assay Kit (Thermo Pierce, Cat# 23235) according to the manufacturer's directions.

Deglycosylation

Spherulin was deglycosylated using a PNGaseF kit according to the manufacturer's instructions (New England Biolabs, Cat# P0704L). Briefly, 500ug of Spherulin was denatured using

10X glycoprotein denaturation buffer at 95°C for 5 minutes followed by 5 minutes on ice. For deglycosylation, 10X G7 Reaction Buffer, 10% NP40 and PNGase were added and allowed to incubate for 6 hours at 37°C. Deglycosylated Spherulin was applied to lectin bound agarose beads in columns (same as above) and eluates were analyzed by SDS-PAGE and digested with trypsin as described below.

SDS-PAGE and in-gel trypsin digestion

The SDS-PAGE gel bands are prepared for mass spectrometry analysis using the following procedures. Colloidal blue stained gel bands were destained in 50% acetonitrile/50mM Tris pH 8.1 until clear, and the proteins reduced with 50mM TCEP/50mM Tris pH8.1 at 55ºC for 30 minutes, followed with alkylation using 20mM iodoacetamide/50mM Tris pH 8.1 at room temperature for 30 minutes in the dark. Proteins were digested in-situ with 0.15ug trypsin (Promega Corporation, Madison WI) in 25 mM Tris pH 8.1 / 0.0002% Zwittergent 3-16, at 37°C overnight, followed by peptide extraction with 2% trifluoroacetic acid and acetonitrile. The pooled extracts were concentrated and the proteins identified by nano-flow liquid chromatography electrospray tandem mass spectrometry (nanoLC-ESI-MS/MS) using a Thermo Scientific Q-Exactive Plus Mass Spectrometer (Thermo Fisher Scientific, Bremen, Germany) coupled to a Thermo Ultimate 3000 RSLCnano HPLC system.

Mass spectrometry

Peptides present in each sample were loaded onto a 0.25uL bed OptiPak trap (Optimize Technologies, Oregon) custom-packed with 5um, 200A Magic C18 stationary phase. Loaded trap was washed for 4 minutes with an aqueous loading buffer of 0.2% FA and 0.05% TFA at 10uL/min. Following the wash, peptides were transferred onto a 35cmx100□m PicoFrit column, self-packed with Agilent Poroshell 120S 2.7um EC-C18 stationary phase, using a Dionex UltiMate® 3000 RSLC liquid chromatography (LC) system (Thermo, San Jose, CA). Peptides were separated using a 400 nL/min LC gradient comprised of 2%-40% B in 0-70 min. Mobile phase A was 2% acetonitrile (ACN)

in water with 0.2% FA and mobile phase B was ACN/isopropanol/water (80/10/10 by volume) with 0.2% FA. Eluting peptides were analyzed using a QExactive Plus mass spectrometer (Thermo-Fisher, Waltham, MA). The instrument was configured to operate in data-dependent mode by collecting MS1 data at 70,000 resolving power (measured at m/z 200) with an AGC value of 1E6 over a m/z range of 360-2000, using lock masses from background polysiloxanes at m/z 371.10123 and 446.12002. Precursors were fragmented with normalized collision energy (NCE) of 28, fragments measured at 17,500 resolving power and a fixed first mass of 140. Resulting tandem mass spectra (MS/MS) were collected on the top 20 precursor masses present in each MS1 using an AGC value of 1E5, max ion fill time of 50ms, and an isolation window of 1.5 Da. All raw data files were transcoded into mzML format using msConvert tool of the ProteoWizard library (PMID: 18606607).

Bioinformatics

We compiled a composite protein sequence database to identify the *Coccidioides* proteins present in the lysate. This database contained *Coccidioides* proteomes obtained from the Broad Institute's *Coccidioides* Genomes project (Neafsey,D.E. 2010; Sharpton,T.J. 2009), SwissProt and RefSeq. RefSeq human and bovine proteomes were added to this database to prevent misidentification of proteins originating from cell culture and other human contamination as *Coccidioides* proteins. Common contaminants (wool, cotton, etc.) were added to the database to account for sample handling artifacts. Reversed protein sequences were appended to the database to estimate protein and peptide identification false discovery rates (FDRs).

MyriMatch (Tabb,D.L. 2007) (version 2.1.38) database search engine matched the MS/MS present in each data file against the composite protein sequence database. The software was configured to use 10ppm m/z tolerance for both precursors and fragments while performing peptide-spectrum matching. The software derived semitryptic peptides from the sequence database while looking for the following variable modifications: carbamidomethylation of cysteine (+57.023 Da.), oxidation of methionine (+15.994 Da.) and formation of n-terminal pyroglutamic

acid (-17.023 Da.). IDPicker (version 3.0.504) software filtered the peptide-spectrum matches at 2% FDR. The software was configured to use an optimal combination of MVH, mzFidelity and XCorr scores for filtering. Protein identifications with at least two unique peptide identifications were considered to be present in the sample. Resulting proteins were clustered into groups of proteins that match the same set of peptides.

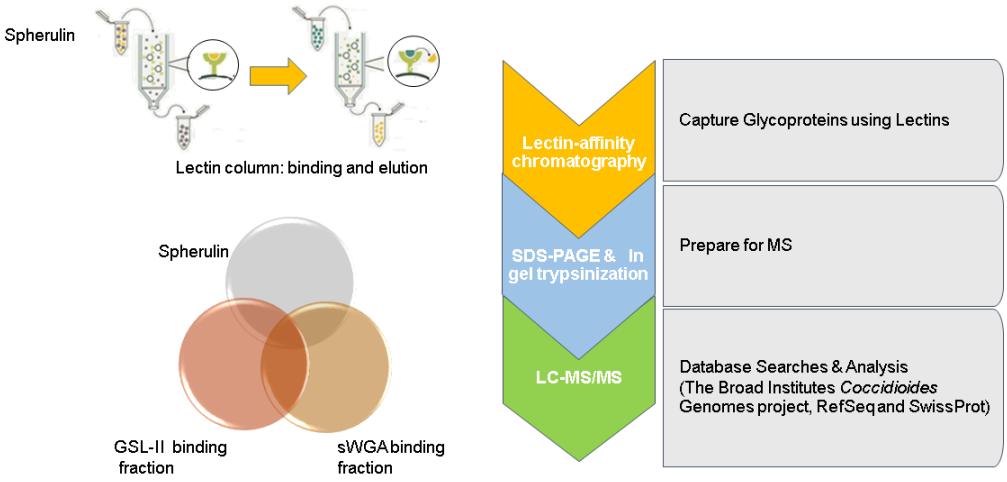

Figure 3. Workflow for generation of Spherulin proteome. To identify proteins present in Spherulin, in-gel trypsin digestion was done following the fractionation with SDS-PAGE. For identifying the fraction of GSL-II and sWGA binding fraction of Spherulin, it was affinity purified on lectin columns and eluates were analyzed with LC-MS/MS.

Results

Lectin-based IHC

Lectin-based IHC with biotinylated GSL-II and sWGA was performed using FFPE lung tissue sections from six patients who underwent wedge resections for Coccidioidomycosis and one tissue section from the wrist of an infected patient. Data from all seven patients is represented in Figure 4. Figure 5 shows GSL-II and sWGA reactivity to the spherules in one patient. Positive and specific binding of GSL-II and sWGA lectins was observed for *Coccidioides* spherules and endospores and not to adjacent lung tissue. PVE lectin did not bind to spherules, but did bind to adjacent host tissue. To support lectin-like binding to spherules and endospores, reactivity of GSL-II and sWGA was inhibited in a concentration-dependent manner with commercial solution of monomeric and oligomeric GlcNAc (100 mM). A high concentration of GlcNAc (1:4 dilution; 25mM) was required to completely inhibit sWGA binding to the spherules while a relatively lower GlcNAc (1:400 dilution; 0.25mM) amount inhibited GSL-II binding suggesting that sWGA binding to spherules and endospores is stronger than GSL-II (Figure 5). When galactose (non-specific sugar) was used to inhibit the binding of these two lectins to the fungal spherules, a complete lack of inhibition was seen.

Lectin-based inhibition ELISA

We wanted to confirm our lectin-based IHC results and assess whether sWGA and GSL-II compete for the same glycan structure. To accomplish this, we performed an inhibition ELISA using biotin-GSL-II and biotin-sWGA as detection agents. As shown in Figure 6, sWGA inhibits binding of biotinylated GSL-II to Spherulin in a concentration dependent manner with a relative IC_{50} of 1.5 uM. In contrast, we were not able to reach 50% inhibition of 1nM biotinylated sWGA with GSL-II even at 20uM, suggesting higher avidity of sWGA for GlcNAc on coccidioidal glycoproteins than GSL-II. PVE, a Galβ4GlcNAcβ2Manα6 binding lectin, did not inhibit either GSL-II or sWGA and served as a “control” lectin.

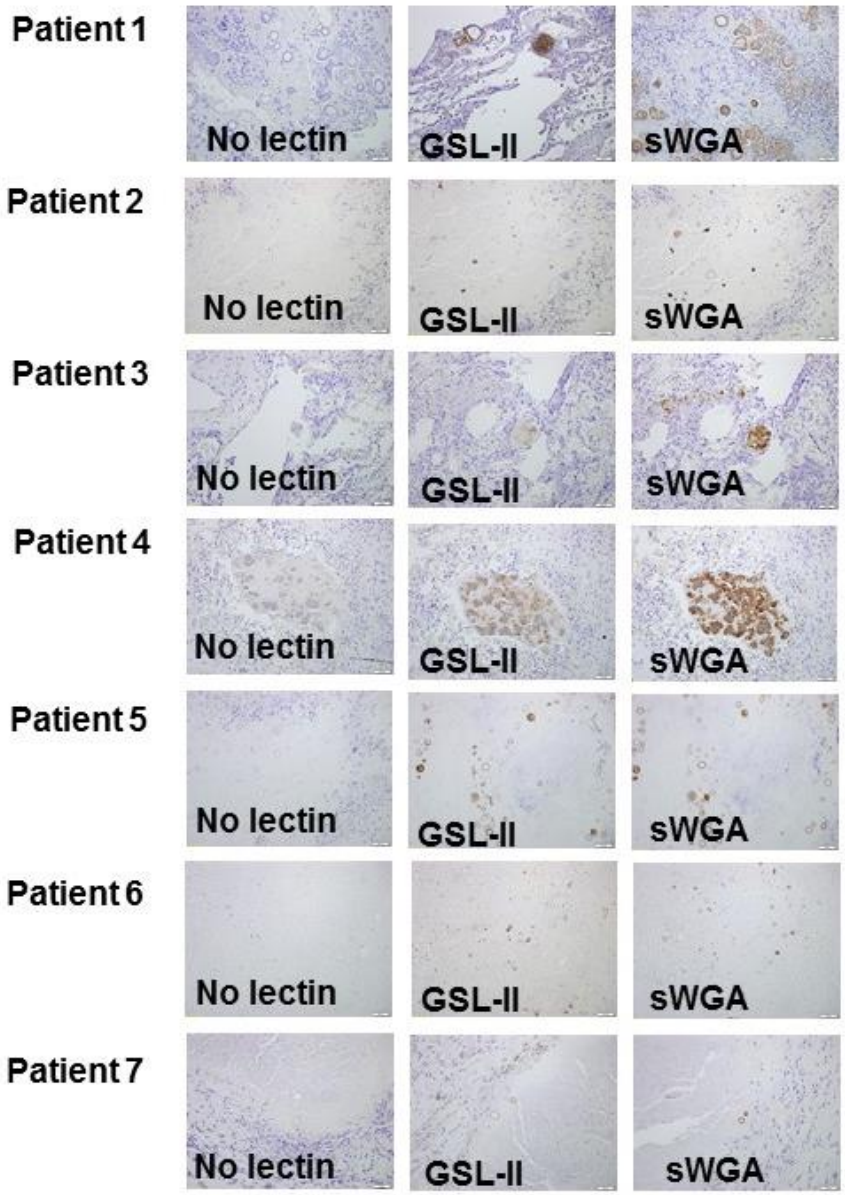

Figure 4. Lectin-based IHC micrographs of seven valley fever patients. Biotinylated lectins were incubated with infected lung tissues from seven patients, as labeled in the picture and as stated in experimental procedures. Brown color indicates reactivity of lectin. GSL-II and sWGA reacted positively and specifically to endospores and spherules, and not to the adjacent human lung tissue.

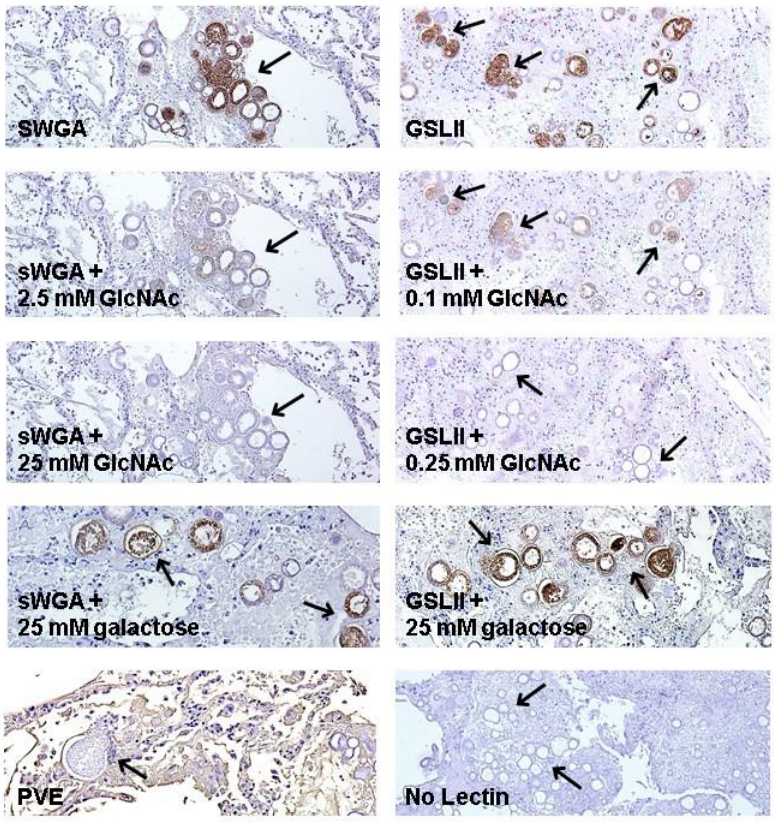

Figure 5. Lectin-based IHC micrographs of lectins. Biotinylated lectins were incubated with infected lung tissues, as labeled in the image and as stated in experimental procedures. Brown color indicates reactivity of lectin. Spherules are round structures with or without endospores inside. Arrows indicate examples of spherules or groups of spherules. GlcNAc-mediated inhibition of sWGA and GSL-II binding to spherules and endospores is shown at different dilutions. Galactose did not inhibit binding of either lectin to spherules.
PVE did not bind the spherules but to the human lung tissue indicating non-specific binding.

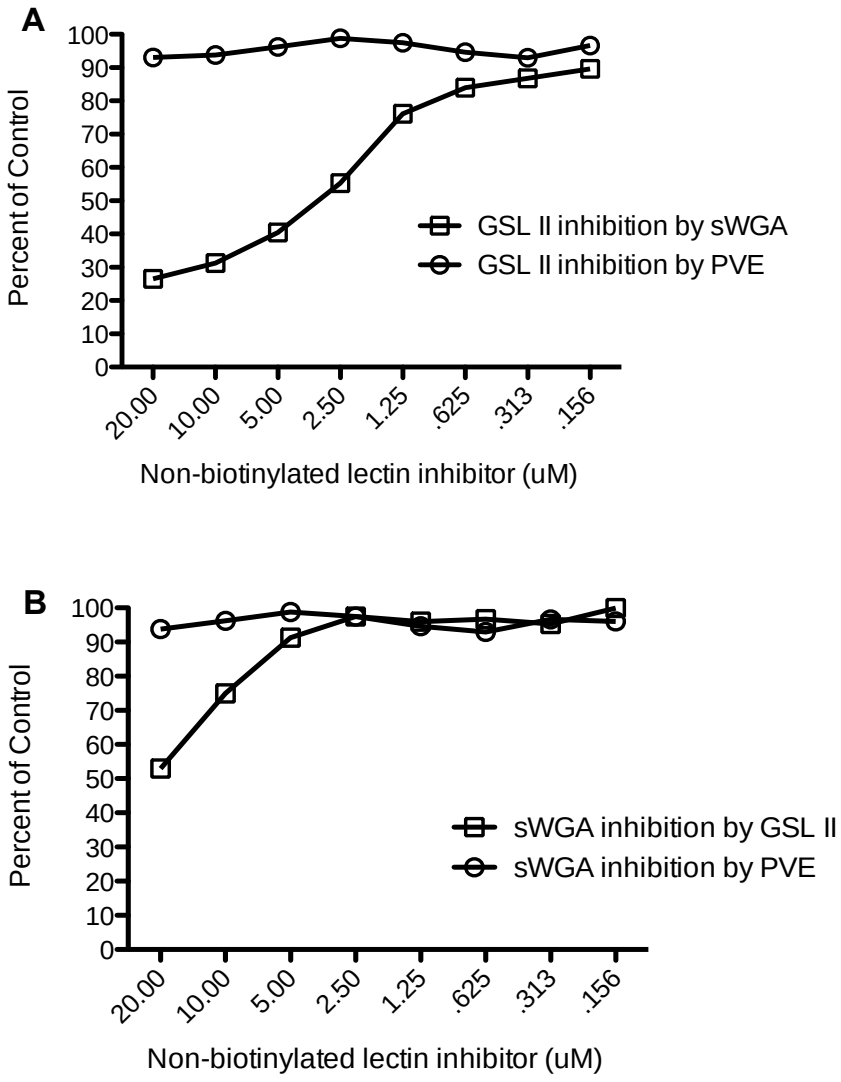

Figure 6. GSL-II and sWGA inhibition ELISAs. Two-fold dilutions of non-biotinylated sWGA and GSL-II lectins starting at 20uM were incubated in an ELISA plate coated with Spherulin for 20 minutes. Then biotinylated GSL-II or sWGA were added and incubated for one hour. After washing the plate, streptavidin-HRP was added to detect biotinylated lectins that were not inhibited from binding to Spherulin. After TMB substrate development the plate was read at 450nm. (O.D. biotinylated lectin in presence of non-biotinylated inhibitor) / (O.D. biotinylated lectin in the absence of inhibitor) x 100 = percent of control.

Mass Spectrometric Identification of Proteins in Unfractionated Spherulin

Given the results from the IHC and ELISA using Spherulin as antigen, our first step in the proteomic analysis of *Coccidioides* was to identify the proteome of Spherulin. Three different *Coccidioides* databases (SwissProt, RefSeq, and Broad Institute's *Coccidioides* Genomes project (Neafsey,D.E. 2010; Schiess,R. 2009) were employed to search the MS/MS spectra derived from Spherulin. A total of 1390 proteins were identified in Spherulin. A listing of the 10 most abundant coccidioidal proteins in Spherulin with high tryptic fragment coverage is shown in Table 1. All of these proteins have an effective FDR of 0.0%. Eight of the top 10 identified proteins in Spherulin are metabolic enzymes important for fungal growth. The second most abundant Spherulin protein identified is a "conserved hypothetical protein" (CPSG_03975) with a pentapeptide (PT) repeat sequence and has high homology with an exoprotein involved in adhesion. This finding indicates that CPSG_03975 is no longer "hypothetical" and is highly abundant during fungal growth in vitro. In total, 434 hypothetical proteins were identified, constituting 31% of the total proteins entries in the proteome of Spherulin.

Other highly abundant proteins found in Spherulin were 5-methyltetra-hydropteroyltriglutamate-homocysteine methyltransferase (MET –E; CPSG_03208), Heat shock protein 90 (CPAG_06539), 3-isopropylmalate dehydrogenase (CPAG_08709), glucose-6-phosphate isomerase (CPAG_05681), enolase (CPAG_04681) and fructose biphosphate aldolase (CPAG_09270).

Protein ID	% sequence coverage	*Unique* tryptic peptides identified from Spherulin
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase (CPSG_03208)	89	188
conserved hypothetical protein (CPSG_03975)	86	89
3-isopropylmalate dehydrogenase (CPAG_08709)	91	89
malate dehydrogenase (CPAG_07192)	89	89
heat shock protein 90 (CPAG_06539)	78	87
enolase (CPAG_04681)	77	86
fructose biphosphate aldolase (CPAG_09270)	79	78
H538.4 glucose-6-phosphate isomerase (CPAG_05681)	80	78
malate synthase (CPAG_07630)	71	73
fumarate reductase Osm1 (CPSG_05536)	83	70

Table 1: Mass spectrometric identification of top 10 coccidioidal proteins in Spherulin. The table lists the top 10 (with highest number of tryptic peptides) proteins identified in Spherulin.

SDS-PAGE of and identification of proteins from GSL-II and sWGA affinity purified Spherulin eluates

Since GSL-II and sWGA lectins bind to spherules and endospores in infected human lung tissue sections, it was possible that the lectins simply bound to chitin since the principle component of chitin is GlcNAc. However, it was also possible that GlcNAc structures are present on coccidioidal glycoproteins and accessible to the lectins. Therefore we employed lectin-affinity chromatography using Spherulin followed by SDS-PAGE to determine if there were any glycoproteins containing GlcNAc that bound to GSL-II and sWGA lectins. After Spherulin was loaded onto GSL-II and sWGA columns, they were washed extensively and eluted with "Glycoprotein Eluting Solution". The eluted material was dialyzed, concentrated and quantified as described in experimental procedures. Figures 7A and 7B show SDS-PAGE gels of Spherulin elution profiles from GSL-II and sWGA columns respectively. Mass spectrometric analysis of the Wash 6 (Figure 7A and 7B) from GSL-II and sWGA columns did not reveal any coccidioidal proteins, indicating that the remaining material bound to the lectin columns was specific for lectin (data not shown).

Our next goal was to learn the identities of lectin-binding glycoproteins in Spherulin. To accomplish this goal we performed lectin affinity chromatography of Spherulin using both sWGA and GSL-II, and then identified the enriched glycoproteins using tandem mass spectrometry and analysis of MS/MS spectra. We also wanted to determine if lectin chromatography enrichment would provide the identities of glycoproteins not identified in whole Spherulin. Searching spectra from tryptic peptides against the *Coccidioides posadasii* (strain Silviera) databases (The Broad Institute's *Coccidioides* Genomes project, RefSeq and SwissProt) revealed that 195 coccidioidal glycoproteins bound to GSL-II (FDR of 1.11%), while 224 glycoproteins bound to sWGA (FDR of 1.01%). A total of 145 fungal glycoproteins bound to both GSL-II and sWGA lectins (Figure 7C). The top 10 most abundant coccidioidal glycoproteins eluted from both GSL-II and sWGA lectin

columns (highest tryptic peptide coverage) are listed in Table 2. The table also lists the percentage peptide coverage for these proteins as well as the *unique* peptide count for each of the replicate runs from GSL-II and sWGA columns. Many hypothetical proteins were also identified in the lectin affinity analysis (36 in GSL-II analysis and 54 in sWGA). A highly abundant hypothetical protein (CPSG_01012), different from the hypothetical protein identified in whole Spherulin (see below) bound to both the lectins. This hypothetical protein has 70% sequence similarity to a ribonuclear protein binding domain protein from Aspergillus and other fungi. Other abundant glycoproteins in the lectin analysis included 5-methyltetrahydropteroyl-triglutamate-homocysteine methyltransferase (CPSG_03208), malate dehydrogenase (CPAG_07192), glucose-6-phosphate isomerase (CPAG_05681, and complement fixation-chitinase (CPSG_08657). A complete list of proteins can be found in appendix A. The lectin-binding glycoproteome constitutes a subset of the proteome derived from Spherulin (Figure 8 and 10).

To further support that the binding of the lectins was due to GlcNAc structures on proteins, Spherulin was deglycosylated with PNGase F. PNGase F cleaves glycans from proteins at asparagine residues (except when there is an α1-3 fucose on the core GlcNAc of the glycoprotein). Deglycosylated Spherulin was run through both GSL-II and sWGA columns followed by SDS-PAGE and trypsin digested using the same methods as before. Mass spectrometry analysis of the PNGase-treated elutions from GSL-II and sWGA lectin columns demonstrated nearly complete loss of binding by deglycosylated coccidioidal glycoproteins to the lectin columns (Table 2 and Figure 9). This deglycosylation experiment suggests that the binding of coccidioidal glycoproteins is lectin-like and specific for GlcNAc. Once the glycoproteins were deglycosylated, they no longer bound to the lectin chromatography columns (eluate profile of deglycosylated Spherulin for sWGA and GSL-II affinity columns is shown in lanes 6 and 8 respectively in Figure 9).

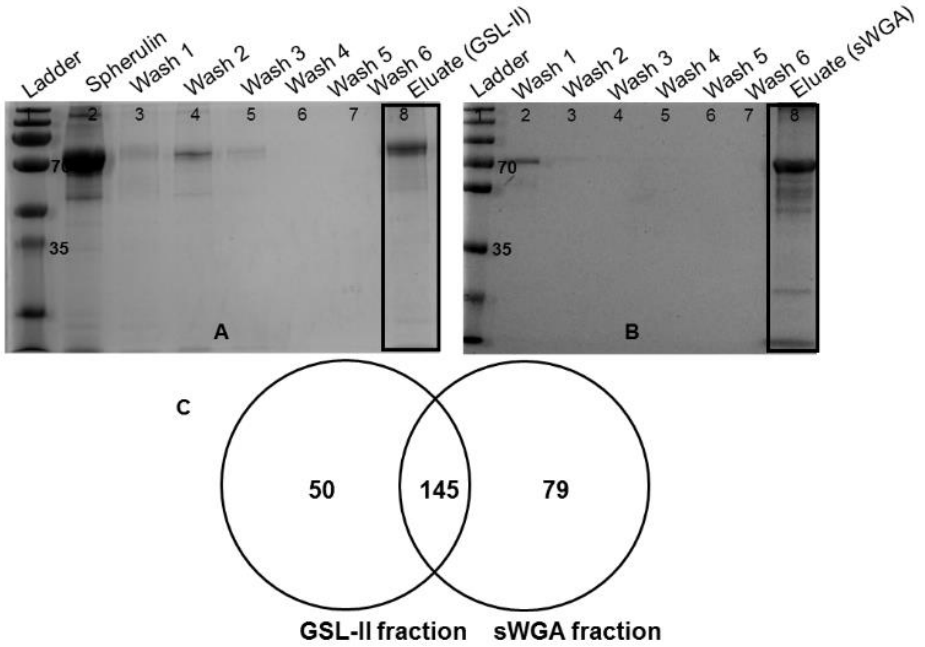

Figure 7. SDS-PAGE from GSL-II (A) and sWGA (B) lectin columns. Lanes are labeled as indicated. Spherulin was starting material prior to running the column. Wash 1-6 for GSL-II and sWGA columns represent six 0.5ml PBS washes collected from the column (column bed volume was 0.2ml). The last lane in each gel represents elution of glycoproteins that bound to the lectin column using "Glycoprotein Eluting Solution" (Vector Labs). GSL-II and sWGA elution lanes (8A and 7B) were cut into slices and processed for trypsin digestion.
Venn diagram (C). The diagram represents total and common proteins identified from GSL-II and sWGA lectin column eluates**.** 145 proteins bound both the lectins. A total of 195 proteins bound to GSL-II and 224 to sWGA.

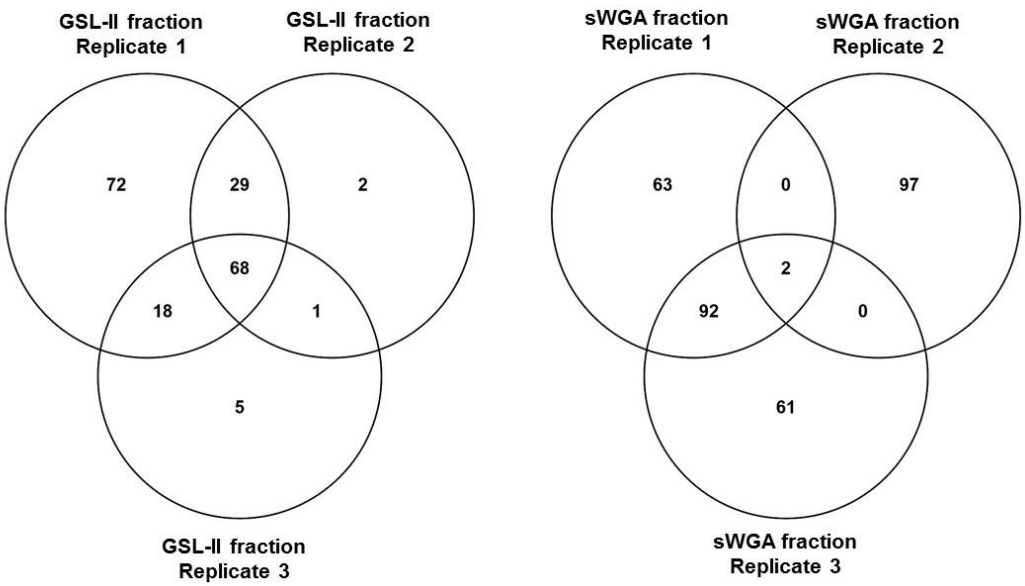

Figure 8. Venn Diagrams of GSL-II and sWGA replicates. Glycoproteins from Spherulin were isolated using GSL-II and sWGA affinity columns in three separate technical replicates (same experiment was conducted thrice). Eluates from each affinity chromatography run were analyzed on mass spectrometer. The Venn diagram shows the number of proteins identified in each replicate and also the common proteins among them.

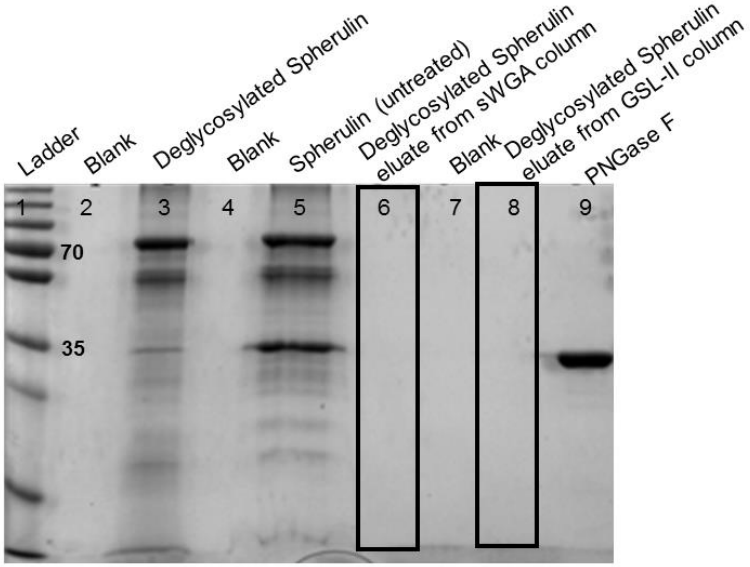

Figure 9. PNGase treatment of Spherulin abrogates the binding of GSL-II and sWGA lectins. SDS-PAGE lanes are numbered and labeled according to treatment. Lanes 5 and 7 (black boxes) were cut into gel slices, treated with trypsin, and run on an Orbitrap QExactive mass spectrometer. Spectra were searched and analyzed using the *Coccidioides posadasii* (Silviera strain) as described in the experimental section.

.

Protein ID	**Max % seq Cove-rage***	*Unique* peptides in **GSL-II** replicates			*Unique* peptides in **sWGA** replicates			*Unique* peptides identified in **PNGase GSL-II** replicates		*Unique* peptides identified in **PNGase sWGA** replicates	
		1	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**1**	**2**
5-methyltetrahydropteroyl-triglutamate-homocysteine methyltransferase (CPSG_03208)	89	23	16	19	28	15	20	0	0	0	0
malate dehydrogenase (CPAG_07192)	89	13	7	10	25	11	11	0	0	0	0
fructose biphosphate aldolase (CPAG_09270)	78	12	7	7	15	8	10	0	0	0	0
enolase (CPAG_04681)	77	13	9	9	9	8	10	0	0	1	0
3-isopropylmalate dehydrogenase (CPAG_08709)	91	13	6	7	12	8	9	0	0	0	0
glucose-6-phosphate isomerase (CPAG_05681)	76	12	0	5	14	2	3	0	0	0	0
aldehyde reductase 1 (CPAG_06394)	68	12	4	6	12	5	6	0	0	0	0

hypothetical protein (CPSG_01012)	60	13	10	8	5	2	10	0	0	0	0
heat shock protein 90 (CPAG_06539)	62	12	5	2	8	6	8	0	0	0	0
complement fixation-chitinase (CPSG_08657)	75	10	5	4	20	8	9	0	0	0	0

Table 2: Mass spectrometric identification of top 10 coccidioidal glycoproteins in Spherulin that bound to both GSL-II and sWGA lectins. Numbers in each column represent technical replicates run on the mass spectrometer. CPAG and CPSG numbers denote the accession numbers. For example, enolase (CPAG_04681) produced 13, 9, 9 tryptic peptides in each of 3 replicates from a GSL-II column chromatography that map back to *C. posadasii* enolase. "PNGase GSL-II replicates" and "PNGase sWGA replicates" indicate that Spherulin was deglycosylated with PNGase prior to lectin affinity chromatography.
*Maximum % sequence coverage from all replicates

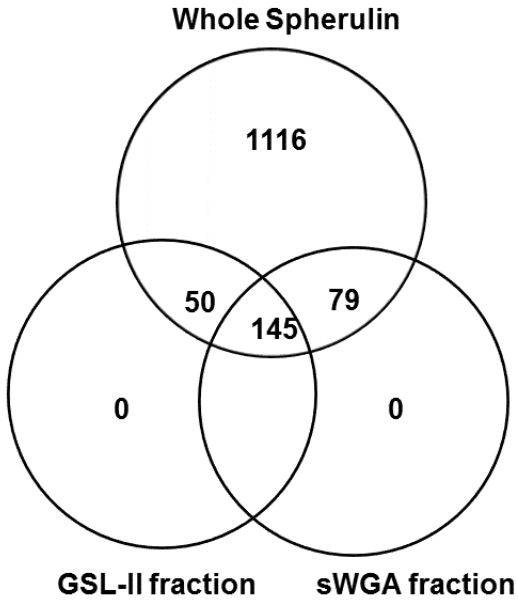

Figure 10. Venn diagram of common proteins among Spherulin, GSL-II affinity purified Spherulin and sWGA affinity purified Spherulin. 145 proteins were common among all three conditions. We identified 1390 total proteins in Spherulin, 195 glycoproteins from Spherulin bound to GSL-II and 224 to sWGA. All glycoproteins in the GSL-II fraction were subsets of whole Spherulin and sWGA. Proteins affinity purified on two lectin columns represent subsets of Spherulin proteome. A complete list of proteins can be found in appendix A.

Conclusions

This study is the first report of the proteome from Spherulin, an antigen preparation from *Coccidioides posadasii*, and a lectin-binding glycoprotein subset of Spherulin. We initially hypothesized that certain lectins could distinguish coccidioidal glycoproteins from human glycoproteins in infected human lungs. This hypothesis is supported by data (Figure 4 and 5) showing that GlcNAc-binding lectins GSL-II and sWGA bound to spherules and endospores in lectin-based immunohistochemistry. Since it was possible that the reactivities of GSL-II and sWGA were due to a non-glycan binding interaction, we inhibited GSL-II and sWGA from binding to the *Coccidioides* fungus in infected lung tissue by pre-incubating the lectins with dilutions of GlcNAc (chitin hydrolysate), (Figure 5). This experiment demonstrated concentration dependent inhibition of GSL-II and sWGA binding to spherules by oligo and monomeric GlcNAc in lung tissue. It also confirmed that the interaction between the lectins and *Coccidioides* was "lectin-like".

An inhibition ELISA using Spherulin as antigen showed that GSL-II and sWGA compete with each other and bind to the same or similar GlcNAc structures. While sWGA inhibited the binding of biotin-GSL-II to Spherulin in a concentration-dependent manner, GSL-II did not inhibit biotin-sWGA from binding to Spherulin as strongly. In fact, we could not reach an IC_{50} at the tested concentrations. Although we did not measure avidities of the lectins, they both have 2 binding sites for GlcNAc. This result suggests that sWGA may have a higher avidity for GlcNAcs on coccidioidal glycoproteins than GSL-II. Further evidence for this is that a 1:400 (0.25mM) dilution of a GlcNAc solution inhibited binding of GSL-II to spherules and endospores, while a 1:4 dilution (25mM) of the same GlcNAc solution was sufficient to inhibit the binding of sWGA to spherules and endospores in the IHC (Figure 5).

Chitin is composed of repeating units of GlcNAc. Therefore, it was possible that the lectins were binding chitin and not coccidioidal glycoproteins in IHC. To address this issue, we performed lectin affinity chromatography followed by SDS-PAGE (Figure 7). GSL-II and sWGA column elutions demonstrated the presence of proteins, suggesting again the presence of

GlcNAc on the glycan moiety of coccidioidal glycoproteins in Spherulin. As a further proof that the lectins were binding to glycans, PNGase F was used to deglycosylate Spherulin prior to lectin affinity chromatography. As shown in the SDS-PAGE in Figure 9 and mass spectrometry results in Table 2 (PNGase GSL-II and sWGA replicates), virtually no tryptic peptides were detected by mass spectrometry when Spherulin was deglycosylated. Taken together, these results strongly suggest that GSL-II and sWGA bind GlcNAc on coccidioidal glycoproteins, not just chitin. The few peptides identified after PNGase treatment could have been due to O-linked glycans containing GlcNAc, as PNGase cleaves only N-linked glycans.

The key to the success of any proteomics experiment is having a protein sequence database with validated sequence entries. Such a database did not exist for coccidioidal proteins. Hence, we constructed a composite database containing predicted proteins derived from the *Coccidioides* genome. These sequences were harvested from RefSeq, SwissProt and The Broad Institute's *Coccidioides* Genome project (Broad, Institute). We chose this composite approach over picking one of the databases due to the complimentary nature of their sequence entries and sequence annotations. A composite database also allows us to conduct proteogenomic studies for novel gene finding (out of scope of this study). As more researchers conduct whole proteome studies of Spherulin, the field would likely advance towards a single validated Spherulin reference proteome.

We investigated the proteome of Spherulin and subsequently examined the GSL-II and sWGA-binding glycoproteome. GSL-II and sWGA were found to be subsets of the Spherulin proteome (Figure 10). Mass spectrometry analysis from GSL-II and sWGA affinity chromatography of Spherulin demonstrated that 145 coccidioidal proteins commonly bound to both GSL-II and sWGA lectins. This was not surprising, as our inhibition experiments demonstrated that GSL-II and sWGA cross-inhibit in a Spherulin ELISA. However, 50 coccidioidal glycoproteins bound to GSL-II that did not bind to sWGA and 79 glycoproteins bound to sWGA that did not bind to GSL-II (Figure 7C). This may suggest some degree of differential specificity between the two GlcNAc-binding lectins. Comparing the total number of glycoproteins,

more (244) bound to sWGA while only 195 bound to the GSL-II lectin. These data corroborate our cross-inhibition data and support of the idea that sWGA has a higher avidity for GlcNAc (hence, more glycoproteins bound to the sWGA column) than GSL-II.

We identified many hypothetical proteins (in whole Spherulin as well as in the lectin eluate fractions (Appendix A) indicating that these "hypothetical proteins" are indeed expressed. In fact, about 30% of the Spherulin proteome consisted of hypothetical proteins. Further investigation of the hypothetical proteins by other researchers may evaluate the sequence similarity to known proteins and elucidate the functions and requirement of hypothetical proteins.

MET-E was found to be a highly abundant protein (see Table 2). It is a zinc-binding vitamin B-12 independent enzyme involved in methionine biosynthesis. This enzyme is found in many other microorganisms and requires folate containing at least 3 glutamic acid residues for its activity. Heat shock protein 90 (CPAG_06539) is an ATP-binding protein with sequence similarity to other fungal HSP90s. It has been suggested as a target for anti-fungal therapies (Burnie,J.P. 2006). Other highly abundant proteins included cytosolic enzymes like 3-isopropylmalate dehydrogenase (leucine synthesis enzyme), glucose-6-phosphate isomerase, enolase and fructose biphosphate aldolase (glycolytic enzymes). We note that each of these proteins are metabolic enzymes (NCBI: protein database).

Complement fixation-chitinase (CPSG_08657) was among the top 10 abundant proteins in both GSL-II and sWGA lectin pull downs of Spherulin. Chitinases hydrolyze chitin, an abundant polymer of GlcNAc, and play a key role in growth regulation and morphogenesis of the fungus. This protein is secreted in to the growth medium as well as localizes in the cytoplasm while retaining the signal peptide that binds to the endoplasmic reticulum (Yang,C. 1996). Complement fixation (CF) is a classical serologic test for diagnosis of valley fever based on the detection of the IgG generated against the CF antigen (complement fixation-chitinase) (Pappagianis,D. 1990).

All of the glycoproteins identified by lectin affinity chromatography were subsets of the Spherulin proteome (Figure 10 and Appendix A). One reason for this might be that mass spectrometry instrumentation has become sensitive and accurate enough to interrogate relatively

complex samples such that enrichment does little to reveal proteins not detected in the original sample. Mass spectrometry may detect multiple thousands of different proteins per run. However, mass spectrometric analysis of our Spherulin preparation identified only 1390 total proteins (including isoforms) which does not represent the entire Coccidioidal predicted proteome, which would include proteins necessary for growth of the fungus in the hyphal phase. The *C. posadasii* genome is 27 Mb and contains 10,355 annotated genes. However, the order (by % sequence coverage) and identities of the top ten lectin-binding glycoprotein fractions are different than the top ten proteins identified from whole Spherulin (Tables 1 and 2). For example, malate dehydrogenase was the second most abundant glycoprotein identified (by spectra) after lectin affinity chromatography, but was the fifth most abundant glycoprotein identified in whole Spherulin. Therefore, some enrichment did occur after lectin affinity chromatography. Lectin affinity chromatography may be more useful to enrich Coccidioidal proteins in other settings with when *Coccidioides* is present in a complex medium, however.

Since many glycoproteins were identified from GSL-II and sWGA lectin affinity chromatography, it appears that GlcNAc is a common glycan structure on coccidioidal proteins. GlcNAc transferases are present in the genomic sequence of *Coccidioides*, but they have not yet been characterized. Preliminary analysis of the GSL-II and sWGA-binding Spherulin glycoproteome suggests that the most abundant proteins are enzymes involved in cellular growth and metabolism. This study is the first to report the proteome of a Spherulin preparation and the GSL-II and sWGA lectin binding glycoproteome as a subset of Spherulin.

Note- The IHC figures used in this dissertation were generated by Yasmynn Chowdhury. These are also a part of her honors thesis.

CHAPTER 3

PROTEOMIC ANALYSIS OF *COCCIDIOIDES* INFECTED HUMAN LUNG TISSUE, URINE AND PLASMA WITH MASS SPECTROMETRY

Abstract

Human plasma is an exceptional pool of about 10,000 proteins. Urine constitutes a subset of approximately 1500 plasma proteins. We previously established that using lectin affinity chromatography, we can isolate coccidioidal glycoproteins from Spherulin lysate. We took this lectin enrichment strategy one step further by determining if our approach could enrich coccidioidal glycoproteins from human plasma and urine. Size exclusion ultrafiltration was used in conjunction with lectin chromatography to capture fungal proteins in these two kinds of samples prior to identification using mass spectrometry. Among hundreds of human proteins, we found over 100 proteins in circulation in two patients and 7 in a third patient. These were absent in three control plasma samples obtained from healthy donors in a non-endemic region. Similarly, in urine from a valley fever patient, five fungal proteins (identified by at least 2 tryptic fragments) were present. Using protein lysate isolated from formalin fixed and paraffin embedded infected lung tissue from a patient with lung nodules, we isolated eighty-seven fungal proteins among hundreds of human proteins.

Overview

Human plasma contains over 10,000 proteins (the Plasma Proteome Database currently holds 10,546 entries) (Anderson,N.L. 2002). These proteins exist in a dynamic range of 10 orders of magnitude in concentration. Since blood flows through all organs in the body and contains many tissue leakage proteins, we hypothesized that patients with pulmonary coccidioidomycosis may have fungal proteins in circulation. We also hypothesized that smaller proteins and/or peptides that are products of proteolytic breakdown may be found in patient plasma. The highly abundant proteins (HAPs) in plasma including albumin, apolipoproteins, immunoglobulins etc., pose a challenge to isolation and identification of low abundance tissue leakage proteins (Anderson,N.L. 2002; Schiess,R. 2009). To circumvent this problem, we employed size exclusion ultrafiltration to remove HAP, many of which are high molecular weight proteins (Figure 11) prior to enrichment with sWGA lectin chromatography. In humans, a large majority of the secreted and cell surface human proteins that are shed into the bloodstream are known to be glycosylated (Schiess,R. 2009). It is well established that fungal proteins are also heavily glycosylated. We previously established the use of sWGA and GSL-II lectin chromatography as a means of isolating glycoproteins from Spherulin (fungal lysate). We choose to use sWGA lectin for enrichment of glycoproteins in plasma and urine to identify fungal proteins in plasma as well as urine. The distinct advantage of mass spectrometry is the ability to identify multiple peptides that map to multiple proteins. This study is the first step towards identifying fungal proteins in circulation in patients (Figure 11).

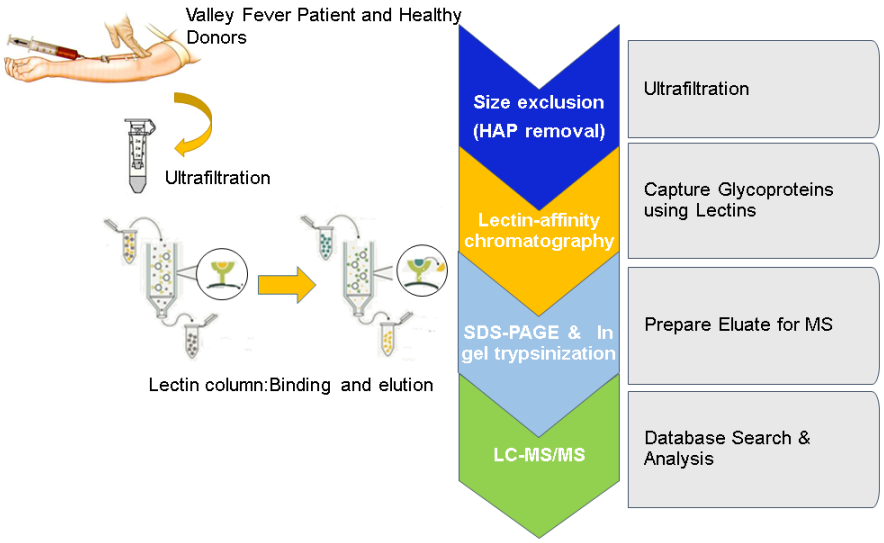

Figure 11. Experimental workflow for coccidiodiodomycosis patient analysis. Plasma urine samples collected from healthy and patient donors were analyzed using lectin affinity columns. These fraction of glycoproteins thus isolated were digested with trypsin and analyzed using LC-MS/MS. The spectra obtained from this analysis were searched against three *Coccidioides* databases.

Experimental Procedures

Collection of patient plasma, urine and FFPE lung tissue

Plasma samples were collected from five valley fever patients. Control plasma was obtained from three healthy donors from non-endemic region (Rochester, MN). The five patients were all positively diagnosed with coccidioidomycosis and the clinical information (in all details available from the attending physician) is listed below.

Patient 1: Presented to Mayo clinic with fatigue, malaise, drenching sweats, non-productive cough, weight loss and dull headache. Left lower lobe of lung showed 2x2 cm nodule and other nodules < 4mm. Diagnostic tests results as follows - CSF serology = 1:2 CF, EIA positive for both IgM and IgG, ID negative.

Patient 2: Long term history of valley fever was reported. At the time of plasma collection for this study, leg lesion and urine cultures were positive for *Coccidioides*. Serology tests around time of sample collection also confirmed the diagnosis: CF test was positive with titer of 1:32, IgG positive by ID and EIA. The patient also suffered from rheumatoid arthritis and was on immunosuppressive agents (leflunomide and prednisone).

Patient 3: Patient was diagnosed with peritoneal coccidioidomycosis. Diagnosis was confirmed with serology testing (CF titer at 1:64 at the time of sample collection).

Patient 4: Patient presented to the clinic with a chronic cough and fatigue with low grade fever. Chest radiography confirmed diagnosis of coccidioidomycosis along with positive serology tests at the time of plasma sample collection (CF titer of 1:32, EIA was also confirmed positive).

Patient 5: Patient was diagnosed with disseminated coccidioidomycosis with meningitis, miliary pneumonia, and dissemination to bone at the time of sample collection.

Urine samples were collected from Patient 1 and a normal volunteer donor (control urine). Due to funding restraints, we were able to only evaluate lysate from one FFPE infected lung.

Lectin affinity chromatography with plasma and urine

50ul of plasma was ultra-filtered with 30 kD filter (Millipore, USA) and filtrate was collected. Gravity packed sWGA lectin columns were made with 400 ul of slurry. The filtered plasma was diluted with 450ul of 1XPBS and allowed to bind to the lectin agarose beads for an hour at room temperature with end on end shaking. The column was then drained and washed with 5-bed volumes of 1X PBS. The glycoproteins were eluted using an N-acetlyglucosamine elution buffer (vector labs). The eluate was dialyzed and concentrated against 1X PBS using a 3kD ultra filter.

Five hundred microliters of urine was centrifuged at 10,000 rpm and supernatant was collected and filtered with 0.22um filter. It was then diluted with 1X PBS and applied to the sWGA lectin column as described above.

FFPE protein lysate

Ten 20 um thick sections were collected in a 1.5 ml eppendorf tube and deparaffinized by incubation in xylene at room temperature(Fisher Scientific, Pittsburgh , PA, USA). The deparaffinized tissue sections were then rehydrated with a graded series of ethanol solutions. The sections were then homogenized in the homogenization buffer (20mM Tris-HCL, pH 9; 2% (w/v) SDS) along with protease inhibitor. These sections were then heated at 100 degrees Celsius for 20 minutes and then maintained at 80 degrees for an additional 120 minutes in a shaking water bath. The tubes were sealed with parafilm to prevent any loss due to evaporation. After protein extraction, any insolubilized material was pelleted and supernatant was filtered through a 0.45 micron filter and quantified using a BCA assay. The protein lysate was then electrophoresed and analyzed with SDS-PAGE as described in the following section.

SDS-PAGE and in-gel trypsin digestion

The SDS-PAGE gel bands were prepared for mass spectrometry analysis using the following procedures. Colloidal blue stained gel bands were destained in 50% acetonitrile/50mM Tris pH 8.1 until clear, and the proteins reduced with 50mM TCEP/50mM Tris pH8.1 at 55ºC for 30 minutes, followed with alkylation using 20mM iodoacetamide/50mM Tris pH 8.1 at room

temperature for 30 minutes in the dark. Proteins were digested in-situ with 0.15ug trypsin (Promega Corporation, Madison WI) in 25 mM Tris pH 8.1 / 0.0002% Zwittergent 3-16, at 37°C overnight, followed by peptide extraction with 2% trifluoroacetic acid and acetonitrile. The pooled extracts were concentrated and the proteins identified by nano-flow liquid chromatography electrospray tandem mass spectrometry (nanoLC-ESI-MS/MS) using a Thermo Scientific Q-Exactive Plus Mass Spectrometer (Thermo Fisher Scientific, Bremen, Germany) coupled to a Thermo Ultimate 3000 RSLCnano HPLC system.

Mass spectrometry

Tryptic peptides present in each sample were loaded onto a 0.25uL bed OptiPak trap (Optimize Technologies, Oregon) custom-packed with 5um, 200A Magic C18 stationary phase. Loaded trap was washed for 4 minutes with an aqueous loading buffer of 0.2% FA and 0.05% TFA at 10□L/min. Following the wash, peptides were transferred onto a 35cmx100um PicoFrit column, self-packed with Agilent Poroshell 120S 2.7um EC-C18 stationary phase, using a Dionex UltiMate® 3000 RSLC liquid chromatography (LC) system (Thermo, San Jose, CA). Peptides were separated using a 400 nL/min LC gradient comprised of 2%-40% B in 0-70 min. Mobile phase A was 2% acetonitrile (ACN) in water with 0.2% FA and mobile phase B was ACN/isopropanol/water (80/10/10 by volume) with 0.2% FA. Eluting peptides were analyzed using a QExactive Plus mass spectrometer (Thermo-Fisher, Waltham, MA). The instrument was configured to operate in data-dependent mode by collecting MS1 data at 70,000 resolving power (measured at m/z 200) with an AGC value of 1E6 over a m/z range of 360-2000, using lock masses from background polysiloxanes at m/z 371.10123 and 446.12002. Precursors were fragmented with normalized collision energy (NCE) of 28, fragments measured at 17,500 resolving power and a fixed first mass of 140. Resulting tandem mass spectra (MS/MS) were collected on the top 20 precursor masses present in each MS1 using an AGC value of 1E5, max ion fill time of 50ms, and an isolation window of 1.5 Da. All raw data files were transcoded into mzML format using msConvert tool of the ProteoWizard library (Kessner,D. 2008).

Bioinformatics

We compiled a composite protein sequence database to identify the *Coccidioides* proteins present in the lysate. This database contained *Coccidioides* proteomes obtained from the Broad Institute's *Coccidioides* Genomes project, SwissProt and RefSeq (Neafsey,D.E. 2010; Sharpton,T.J. 2009). RefSeq human and bovine proteomes were added to this database to prevent misidentification of proteins originating from cell culture and other human contamination as *Coccidioides* proteins. Common contaminants (wool, cotton, etc.) were added to the database to account for sample handling artifacts. Reversed protein sequences were appended to the database to estimate protein and peptide identification false discovery rates (FDRs).

MyriMatch (version 2.1.38) database search engine matched the MS/MS present in each data file against the composite protein sequence database (Tabb,D.L. 2007). The software was configured to use 10ppm m/z tolerance for both precursors and fragments while performing peptide-spectrum matching. The software derived semitryptic peptides from the sequence database while looking for the following variable modifications: carbamidomethylation of cysteine (+57.023 Da.), oxidation of methionine (+15.994 Da.) and formation of n-terminal pyroglutamic acid (-17.023 Da.). IDPicker (version 3.0.504) software filtered the peptide-spectrum matches at 2% FDR (Kessner,D. 2008; Ma,Z.Q. 2009). The software was configured to use an optimal combination of MVH, mzFidelity and XCorr scores for filtering. Protein identifications with at least two unique peptide identifications were considered to be present in the sample. Resulting proteins were clustered into groups of proteins that match the same set of peptides.

Results

***Coccidioides* proteins in patient plasma**

We used both GSL-II and sWGA in our initial experiments. After 30 kD ultrafiltration and sWGA lectin chromatography, we identified 150 proteins in Patient 1 with tryptic peptides equal to or greater than 1. Among these, for 125 proteins at least 2 peptides were identified (tryptic

fragments ranged from 2 to 26). Using sWGA affinity chromatography enrichment, 125 proteins were identified with 2 or more peptides and 24 with a single peptide (total 150 proteins).

Patient 2 also had circulating coccidioidal proteins. Using GSL-II affinity enrichment, we identified a total of 122 proteins and at least 2 peptides were identified from 64 proteins (peptide range was 1 to 17). On the other hand, enrichment of plasma glycoproteins with sWGA affinity enrichment, yielded a total of 137 proteins (peptide range of 1 to 27). Among these, 97 proteins were identified with at least 2 tryptic fragments. Table 3 lists the 10 most abundant proteins (with highest spectral counts) found in patient plasma. Patient 3 had only seven fungal proteins present in plasma which were identified by two or more tryptic fragments. Patient 4 and 5 had only two and three fungal proteins in circulation. In all five patient plasma, actin was identified. This highly conserved protein was the only protein common among all patients.

Among the 3 control plasma investigated with both GSL-II and sWGA chromatography, single peptides were identified from 6 coccidioidal proteins in 3 different 'controls' plasma samples collected from healthy donor. A protein identification can be made by the presence of a single tryptic fragment/peptide ('one hit' protein entries) but if the tryptic fragment is not unique to the sequence of that protein, the confidence in identification by mass spectrometry method like in this study is low. For instance, actin, a highly conserved protein was seen in all controls. Such single hits are considered false positive. Figure 12 shows the proteins identified in patient plasma by presence of at least two tryptic peptides.

Although not a focus in this study, as expected, we identified the presence of many human proteins in patient as well as normal donor plasmas. A total of 56 human proteins were identified in plasma from patient donor 1 and 45 in patient donor 2. This was less than the number of human proteins identified in plasma from patient donors 3, 4 and 5 respectively (338,131 and 222). The total number, including human as well as fungal, proteins identified in all five patient plasma were comparable. In fact, the presence of fewer human and more fungal proteins in patient 1 and 2 suggest that there was a possible enrichment of fungal proteins with the use of lectin column. In our IHC experiments (chapter 2), lectins –GSL-II and sWGA bound

specifically to the fungal spherules (and not surrounding human tissue) which also supports this possibility of enrichment. For a complete list of proteins identified in plasma (patients and healthy donors), please see Appendix B.

Coccidioides proteins in patient urine

A urine sample collected from valley fever patient donor (Patient 1) was enriched for glycoproteins using a sWGA lectin affinity column. The eluate was trypsin digested (in-gel digestion) to reveal presence of 10 coccidioidal proteins. Five proteins were identified by single unique peptides while two or more peptides were identified from three different proteins. Two proteins were identified by the presence of six and five tryptic fragments respectively. Among the total 10 proteins thus identified, four were also identified in a “control” urine sample obtained from a healthy donor. This included highly conserved proteins such as actin and ATP synthase. Three proteins were uniquely present in urine from Patient 1 (and absent in plasma). These proteins included ADP ribosylation factor, a GTP binding protein and ATP synthase beta subunit.

As expected, we identified the presence of many human proteins in patient as well as normal donor urines. Urine from healthy donor had over 400 non fungal proteins (including different isoforms) and patient urine had over 300 non-fungal proteins. These included different isoforms of proteins and each protein was identified by presence of one or more tryptic peptides.

Protein ID	Spectral count	Maximum % coverage	*Unique* tryptic peptides identified from Spherulin	
5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase (CPSG_03208)	283		53	31
malate dehydrogenase (CPAG_07192)	162		52	13
O-acetylhomoserine	117		39	12
enolase (CPAG_04681)	76		55	15
vacuolar protease A	72		40	8
peroxisomal matrix protein	71		64	7
endochitinase 1	70		35	12
superoxide dismutase	59		72	8
heat shock 70 kDa protein	54		24	17
formate dehydrogenase	51		31	9

Table 3. *Coccidioides* proteins identified in patient (n=2) plasma using sWGA lectin chromatography and LC MS/MS. The table shows a partial list of fungal proteins identified in patient plasma (with high spectral and peptide counts) and were absent in any control plasma. Percentage sequence coverage signifies the extent of tryptic fragments identified from the protein sequence.

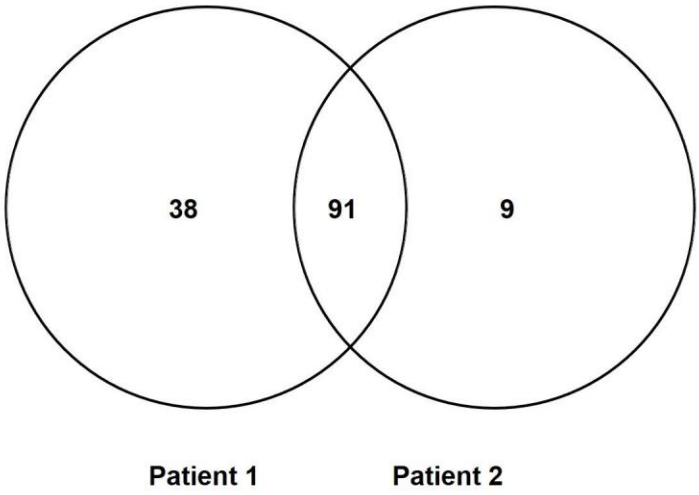

Figure 12. Venn diagram of coccidioidal proteins in patient plasma. The venn diagram shows common coccidioidal proteins in patients' plasma. Ninety one proteins and their isoforms were commonly present in plasma acquired from patient 1 and patient 2 (tryptic peptides ≥ 2).

Coccidioidal proteins in protein lysate of FFPE infected human lung

As described in materials and methods section, protein lysate was extracted from an archived FFPE infeted lung tissue. The lysate was then electrophoresed by SDS-PAGE and in-gel trypsin digested. Mass spectromertic analysis provided evidence of 87 proteins in the lysate. The lung tissue was also stained with hemotoxylin and eosin and showed presence of spherules and endospores. Since we did not eliminate human proteins or enrich for *coccidioidal* proteins, many human proteins were also identified. All proteins extracted from infected lung were also present in the coccidiodin (lysate of mycelial state) (Table 4 and Appendex C). These included many metabolic proteins such as superoxide dismutase, malate dehydrogenase, complement fixation chitinase and 22 hypothetical proteins. Only one FFPE infected lung tissue was available at the time of study.

Protein ID
complement fixation-chitinase CPSG_08657
superoxide dismutase CPSG_07243
exo-beta-1,3-glucanase Exg0 CPSG_02802
alpha-mannosidase CPSG_02648
beta-glucosidase CPSG_03385
hypothetical protein CPSG_02027
alkaline phosphatase CPSG_06408
metalloprotease MEP1 CPSG_05583
FAD-dependent oxygenase CPSG_05770
endo-1,3-beta-glucanase CPSG_08500

Table 4. Ten most abundant coccidioidal proteins identified in protein lysate obtained from FFPE human lung tissue. The table shows a partial list of fungal proteins identified in protein lysate of FFPE infected lung tissue. Please see appendix C for complete list of proteins.

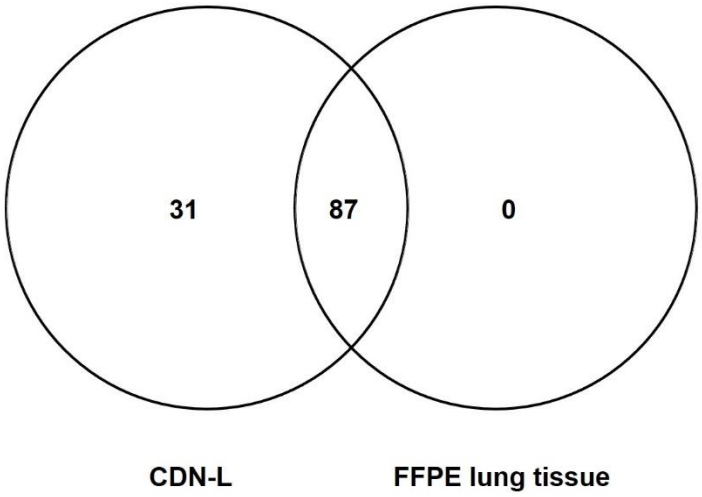

Figure 13. Proteins in *Coccidioides* lysate versus lysate obtained from formalin fixed and paraffin embedded infected lung tissue. The venn diagram shows the number of proteins identified in each lysate. Eighty seven fungal proteins identified from the infected lung lysate were a subset of ones identified in the mycelial lysate (CDN-L).

..

Conclusions

A limitation in the development of an antigen based (as opposed to host immune response) diagnostic assay development is the absence of a *Coccidioides* proteome. If the proteome of *Coccidioides* was thoroughly characterized, techniques such as next generation mass spectrometry would be helpful in diagnosis. To this end, we generated the proteome of Spherulin and the sWGA and GSL-II binding glycoproteome of spherulin (chapter 2). Following the development of a methodology of capturing and identifying fungal glycoproteins using lectin based affinity chromatography, we investigated plasma, urine and FFPE lung tissue lysate for the detection of *Coccidioides* glycoproteins using the same approach. As mentioned earlier, LC-MS/MS combined with pre-sample processing provide an optimal means of discovering proteins in a complex sample such as plasma. This is important for preliminary investigative studies that are discovery-driven like the present study. The ability to observe a given peptide or protein cannot always be controlled in LC-MS/MS studies (Pisitkun,T. 2007). In general, a particular peptide ion may be only be sporadically detected unless it is derived from an abundant protein in the sample (or targeted in a MRM study). This problem is compounded by the fact that tissue leakage proteins (such as in this study) are present in seven degrees less in magnitude (low ng/ml) than most of the highly abundant proteins (HAPs) in human plasma such as albumin and immunoglobulins (5-18 mg/ml). Removal of HAPs is therefore desirable when looking at moderate and low abundant proteins and protein products in plasma. Use of ultrafiltration to exclude higher molecular weight proteins (which also happen to be high abundance) is one of the ways to reduce the complexity of plasma. However, presence of immune complexes is a known challenge in detection of antigenemia in plasma for many infections including coccidioidomycosis (Durkin,M. 2009). Coccidiodal proteins may be present in a complex with antibodies. Ultrafiltration removes the high abundance human proteins but this also risks the removal of fungal entities because antibodies themselves are larger than 150 kDa. Some of the methods to dissociate these complexes such as acid treatments and pretreatment of plasma with EDTA in presence of high heat may improve the chances of detecting coccidioidal proteins and protein products that

may form complexes with HAPs. In addition to size exclusion, the focus on the N-glycoproteome with use of lectin affinity enrichment, not only reduces the complexity further but also helps target an information-rich sub-proteome that may carry disease biomarkers.

The greatest benefits for patients with coccidioidomycosis with use of a blood based antigen test would be early diagnosis to inform clinical management of the disease. To fulfill this aim, the present study was initiated with patient plasma and urine. The process of development of a diagnostic test can be divided into many stages. This study constitutes an early stage of the process. The first step is to learn which proteins are present in the proteome. Then one may conduct discovery-driven investigation and identify the presence of fungal proteins in biosample. A subsequent stage involves verification of the candidate proteins in biosamples from disease-affected donors compared to healthy donors. The latter stages of the process involve the further validation of peptides that reproducibly appear in disease-affected donors and not in control donors.to test the diagnostic capability of the antigen based assay.

In the present study, we have identified multiple fungal proteins recurring in five patients. We identified 91 proteins common in two patients' plasma (Figure 12). Patients 3 and 4 had very few fungal proteins detectable in circulation. One possible reason for the difference in number if fungal proteins identified in the 5 patients may be that patients 3-5 had been taking anti-fungal agents for over 8 weeks and were reported to be responding to this therapy while patient 1 and 2 were documented to have disseminated disease and were unresponsive to the anti-fungal therapy. Antifungal therapy in responsive patients may have arrested growth of the fungus in the host. We also identified 10 proteins in urine of Patient 1. Five proteins out of these 10 proteins were identified on bases of only a single unique peptide. Seven of these ten proteins were common in urine and plasma from patient 1. The most abundant proteins in patient plasma were also identified in Spherulin proteome analysis. These include many proteins involved in growth and metabolism such as enolase, malate dehydrogenase, and formate dehydrogenase. Heat shock protein 90 (hsp90) was also an abundant protein present in patient 1 and 2 plasma. This highly conserved protein is thought to be play a role in pathogenicity of some fungal species and

is expressed on the cell surface (Burnie,J.P. 2006). Endochitinase was also identified in both patients. Many common tryptic peptides were detected among patient 1 and 2 from abundant proteins such as MET-E and malate dehydrogenase.

Due to restraints in funding, we were not able to include more patient and control plasma and urine in the current study. It is important to investigate plasma and urine from more patients and also include controls plasma from endemic regions as well as plasma from patients suffering from community acquired pneumonia (CAP) without a valley fever diagnosis. With a statistically significant number of experimental and control plasmas, confidence in identifying the fungal proteins in circulation in patients may be improved.

When we analyzed lysate from an infected FFPE lung tissue, over 80 fungal proteins were identified. Retrieving protein lysate from archived FFPE tissue poses several challenges but numerous studies have evaluated FFPE tissue lysate with varying degree of success (Maes,E. 2013; Goto,M. 2007; Paulo,J.A. 2013; Sprung,R.W.,Jr 2009; Vincenti,D.C. 2013). For studying *coccidioidal* proteins derived from spherules in lungs, a technique such as laser capture microscope can be very valuable. With aid of a laser capture microscope, spherules can be micro-dissected from the adjacent human lung tissue and lysate from these collected spherules can be valuable in more comprehensive detection and identification of fungal proteins in spherule in lungs. A drastic decrease in the background of the human proteins (similar problem occurs in plasma and urine analysis) can prove to be a major advantage leading to identification of more fungal proteins. It would be ideal to obtain urine and plasma samples from a large set of patients and healthy donors. Then one could learn which proteins in tissues find their way into circulation and which are exclusively present in tissues only.

Following verification with more samples, next steps may include generation of antibodies against a set of most abundant and recurring fungal proteins that can be used in a targeted study. Since we have identified glycoproteins, there is a potential to investigate the roles of specific proteins and protein modifications in pathogenic ability (or dissemination) of this fungus. For

development of a LC-MS/MS based test, a promising solution could be the development of approaches based on the multiple reaction monitoring (MRM) principle.

Some studies have attempted to study antigenuria (antigen in urine) and antigenemia (antigen in blood) in valley fever (Durkin,M. 2009). However, lack of a successful and consistent multi-protein panel was a drawback of such studies (Champer,J. 2012).

CHAPTER 4

DISCUSSION AND CONCLUSIONS

The incidence of Valley Fever is growing in the United States. Although a majority of infected individuals do not seek medical care, new regulations in reporting the incidence of this disease in patients who seek care and are diagnosed, has led to registering of more reported cases today as compared to two decades ago. Less than 5,000 cases were reported in 1995; by 2011, this number had risen to more than 20,000 (Centers for Disease Control and Prevention (CDC) 2013; Centers for Disease Control and Prevention (CDC) 2013). This number is similar to the number of reported Lyme disease cases throughout the USA. Additionally, the CDC estimates that some 150,000 valley fever cases go undiagnosed annually. As the population ages and migrates to the southwest to avoid harsh winter weather, many more people will be at risk of contracting valley fever. There is an urgent need of reliable diagnosis and a therapy (or vaccine) that is currently absent for this disease.

With knowledge of the *Coccidioides* proteome, new antigen based diagnostic tests and possibly an anti-fungal therapy can be developed to combat this disease. If the proteome of *Coccidioides* were thoroughly characterized, techniques such as next generation mass spectrometry could also be applied in diagnosis. To address this gap in knowledge, we generated a Spherulin proteome and sWGA and GSL-II binding glycoproteome of Spherulin (Chapter 2).This proteome may serve as a valuable resource for researchers studying this organism and the disease that it causes. We used three different databases to identify a maximum possible number of total proteins. In future studies, combining or consolidating these different databases into a comprehensive one with all putative sequences will be important. In the proteome of this Spherule phase lysate, Spherulin, we noted that a majority of proteins are ones involved in growth and metabolism of this fungus. Due to funding restraints, we were unable to analyze all the available different antigenic fractions of this fungus. For example, it would be very interesting to study proteome of Coccidioidin (lysate of the mycelial stage of this dimorphic fungus) and identify any differences in the two morphological phases of this organism (Viriyakosol,S. 2013).

In next part of this study, we identified two lectins-GSL-II and sWGA which exhibit specific reactivity to spherules in infected FFPE lung tissues. To identify the proteins that bound to these lectins, we used lectin –chromatography followed by LC-MS/MS. This allowed us to identify a glycoproteome subset of spherulin proteome.

Following the development of a methodology of enrichment and identification of fungal glycoproteins in Spherulin, we investigated plasma, urine and FFPE lung tissue lysate using the same approach. Although LC-MS/MS is a powerful tool for discovery of proteins, the extreme complexity of plasma presents a great challenge. For discovery-based preliminary investigative studies like the present study, use of LC-MS/MS enables sequence-based discovery of novel proteins in a complex sample and allows a comprehensive look at a given proteome. Some studies have indicated antigenemia or antigenuria in valley fever patients via the use of ELISA assays (Galgiani,J.N. 1991). However, positive indication of fungal proteins in circulation present only in a percentage of patients and no protein identifications (from blood) could be made. In the present study, over 100 proteins and their isoforms in plasma acquired from two valley fever patient donors were identified by LC-MS/MS. Analysis of three additional patient plasmas did not demonstrate similar results. Very few coccidioidal proteins were identified in plasma acquired from patients 3, 4 and 5. Actin was the only common protein in all patients. Overall, our findings suggest that many coccidioidal proteins and peptides were in circulation in patient 1 and 2 while very few were present in patient 3, 4 and 5. These results correlate with patient diagnosis and treatment status at the time of plasma collection. Patient 1 and 2 had disseminated form of the disease and were found to be unresponsive to the anti-fungal therapy administered to them while patients 3,4 and 5 had acute (pulmonary infection) form of disease and were responding to anti-fungal agents that were prescribed at least 8 weeks prior to plasma collection. The anti-fungal therapy in these patients may have led to an arrest in the growth of the fungus and concomitantly an absence of shedding of fungal proteins (peptides) in circulation. Resident dendritic cells and macrophages in lungs encounter inhaled arthroconidia and migrate to the lymph nodes where they present fungal antigens and activate lymphocytes. These lymphocytes including

macrophages, B cells and T cells migrate to the site of infection and form granulomata. Endospores from a burst spherule may gain access to the blood stream and disseminate to other non-pulmonary sites in the host. While in circulation, these fungal proteins likely undergo proteolytic degradation due to host serum proteases and immune attack. The presence of antifungal drugs may arrest the growth and possible dissemination and hence may influence the presence of detectable fungal protein products in circulation. The concentrations of these biomarkers (fungal protein products) in blood is unknown. A complete absence or a very low concentration of any possible fungal products in plasma of these three patients may explain the results of LC-MS/MS analysis (Marimuthu,A. 2011; Pisitkun,T. 2007; Schiess,R. 2009).

To develop and validate an assay to detect antigen in circulation in patients with valley fever, it is important to investigate plasma and urine from more patients than was investigated in this study and also include controls samples collected from donors in non-endemic regions as well as plasma from patients suffering from community acquired pneumonia (CAP) without a valley fever diagnosis. Since CAP caused by this fungal infection is frequently misdiagnosed and treated inappropriately with antibiotics because of absence of a reliable diagnostic test, it can be very valuable to study CAP patients (bacterial, viral and fungal origin) (Chang,D.C. 2008). Due to the expense of our mass spectrometric method (>$900 per sample), we were unable to study more patients. In future studies, a panel of a higher number of plasma and urine samples from patient donors, CAP patients and healthy donors should be studied. With a significant number of experimental and control samples, we can identify a set of coccidioidal proteins that circulate in the majority of patients. Ideally, plasma and urine should be collected from each of the patient and healthy donors. This will allow for elimination of false positives. It is worth noting that when MS spectra obtained from tryptic fragments is searched against any given database, it is common to identify highly conserved proteins such as actin, ATP synthase etc. in ‘control’ samples. In order to build a consistent panel of biomarkers (as in study of antigenemia in valley fever), it is important to run many biological (samples from various donors) and technical (multiple runs on MS with a given sample) replicates and identify proteins recurring in experimental groups and

absent in controls (Pisitkun,T. 2007). In this study we provide a preliminary listing of fungal proteins present in patient specimens (and absent in healthy donor bio specimen). However, due to lack of a greater number of different kinds of available bio-specimens (blood and urine) as well as lack of recruitment of additional donors (both patients and healthy individuals) for each kind of bio specimen, the presence of most abundant fungal proteins identified in circulation in patients remains to be confirmed with more MS runs.

In the materials analyzed using MS, the presence of a large number of glycosylated human proteins along with the fungal glycoproteins of interest in this study presents a big challenge. Although use of lectin chromatography is a viable approach to isolate any fungal products in plasma, an initial reduction in human proteins is a necessary step to tackle this hurdle. In this study, size exclusion filtration and lectin chromatography was used. An alternative to use of size exclusion is immunoaffinity depletion of highly abundant proteins (HAP) in plasma. Commercially available kits claim to deplete most abundant human proteins in plasma such as apolipoprotein, albumin, and immunoglobulins and reduce 85% of the plasma proteome burden in a sample and hence introduce the potential for identification of much smaller fraction consisting of disease biomarkers. However, many MS studies provide support for the fact that untargeted proteomic analyses using current LC-MS/MS platforms, even with immunodepletion, cannot be expected to efficiently discover low-abundance, disease-specific biomarkers in plasma (Tu,C. 2010). There is also concern of depletion of non-targeted proteins by such immunodepletion columns. For instance, when using an immunodepletion kit for removing top six most abundant proteins in plasma, non-specific removal of other proteins (non-targeted) may occur. However, in this study we did not explore the use of immunodepletion because the currently available columns allow a scant amount of plasma that decreased the likelihood of detecting fungal proteins in circulation.

Removal of large majority of proteins from the samples (plasma or urine) poses a potential problem of inadvertently removing the peptide components as well (that may or may not be fungal). Peptides in plasma are known to associate with higher molecular weight proteins. If

this were the case with fungal peptides as well, there might be a loss in fungal peptide yield from plasma. It may be helpful to pre-treat plasma with EDTA to dissociate the immune complexes (Durkin,M. 2009). Plasma proteins undergo proteolytic degradation leading to presence of an impressive plasma peptidome. Hence, it may be interesting to study the peptidome (fungal) and determine if fungal peptides (uniquely derived only from *Coccidioides* protein sequence) can be identified in circulation. Some groups have attempted to study the human peptidome but there is much unexplored potential in this field (Tu,C. 2010; Richter,R. 1999).

Another aspect of analysis of patient plasma and urine pertains to sample processing for MS analysis. In this study, protein samples were trypsin digested for analysis on LC-MS/MS. We used in-gel trypsin digestions for LC-MS/MS sample preparation. This technique, though effective and sensitive, is very time consuming and expensive. It takes approximately 3 days to perform the affinity enrichment on lectin columns, and 1-2 days to complete the mass spectrometry analysis, although multiple samples may be processed simultaneously. This discovery-based approach is both too time consuming and expensive to be used as a clinical test. In-solution trypsin digestion may be an alternative approach to identifying the fungal peptides in patient plasma. In-solution trypsin digests of lectin-enriched plasma fractions does not require SDS-PAGE, reducing processing time. However, reduction in sensitivity (maximum total possible number of protein identifications made with LC-MS/MS) as compared to in-gel trypsin digestion is a major concern (Pisitkun,T. 2007).

The goal of this pilot study was to investigate if fungal proteins are circulating in patient plasma. The results indicate that fungal proteins are in fact present in circulation and can be detected. However, this was true only for patients with disseminated form of disease and also ineffective anti-fungal therapy. Following analysis of plasma from other patients and healthy donors, the next steps may include the identification of N-glycosylated peptides and glycosylation sites. To identify the glycosylation sites, PNGase F treatment (an enzyme that removes the N-linked sugar moieties) can be used the presence of heavy (^{18}O) water. This procedure would mark all N-linked glycosylated sites with heavy oxygen. This would help with differentiation of

deamidation (deglycosylation of a glycoprotein induces a mass change of 0.98 Da which is detected by MS) due to sample handling rather PNGase F treatment. The two isotypes of oxygen are differentially detected by LC-MS/MS. After confirmatory studies for the identification of the glycoproteins identified in this study, future studies may investigate the roles of specific proteins and protein modifications with respect to fungal virulence factors. For this purpose, deletion mutants can be generated and fungal growth in vitro can be studied. *Coccidioides* undergoes morphological change induced due to change in growth conditions and disruption of certain genes undoubtedly affect the transition of endosporulating stage to spherule phase. This knowledge may be valuable in developing therapy targeted towards disruption of spherule formation and thereby inhibiting fungal growth in vivo. We identified many proteinases in Spherulin (see Appendix A). A serine proteinase from *Coccidioides* cell wall is thought to be partly responsible for autolysis of the segmentation apparatus of mature spherules, a pivotal morphogenetic process for release of endospores and subsequent proliferation of the pathogen (Yuan,L. 1988).

A majority of valley fever patients that visit a clinic at earlier stages of infection may present with CAP which is frequently misdiagnosed leading to unnecessary anti-bacterial treatments and a delay in treatment of this disease. With the protein identities made in this study, a diagnostic and possible a therapeutic approach can be developed. In addition to the proposed MS based diagnostic test, development of an EIA based assay is a viable option. With the knowledge of proteins (using MS data) present in patients at early stages *Coccidioides* infection, an antibody (against fungal proteins) based assay may be developed. Polyclonal antibodies (or monoclonal) generated in a rabbit or goat immunized against a preparation of Coccidioidin or Spherulin may be used to capture circulating fungal proteins in blood to facilitate early and definite diagnosis. By providing a timely diagnosis and proper management of the initial respiratory infection in patients, healthcare providers can improve patient care and the prognosis of this disease.

References

1. Ampel NM. The diagnosis of coccidioidomycosis. F1000 Med Rep 2010 Jan 18;2:10.3410/M2-2.

2. Ampel NM. Coccidioidomycosis in persons infected with HIV-1. Ann N Y Acad Sci 2007 Sep;1111:336-42.

3. Ampel NM. Measurement of cellular immunity in human coccidioidomycosis. Mycopathologia 2003;156(4):247-62.

4. Anderson NL, Anderson NG. The human plasma proteome: History, character, and diagnostic prospects. Mol Cell Proteomics 2002 Nov;1(11):845-67.

5. Binnicker MJ, Buckwalter SP, Eisberner JJ, Stewart RA, McCullough AE, Wohlfiel SL, Wengenack NL. Detection of *Coccidioides* species in clinical specimens by real-time PCR. J Clin Microbiol 2007 Jan;45(1):173-8.

6. Blair JE, Mendoza N, Force S, Chang YH, Grys TE. Clinical specificity of the enzyme immunoassay test for coccidioidomycosis varies according to the reason for its performance. Clin Vaccine Immunol 2013 Jan;20(1):95-8.

7. Broad I. *Coccidioides* group Sequencing Project, Broad Institute of Harvard and MIT [computer program]. .

8. Brown J, Benedict K, Park BJ, Thompson GR,3rd. Coccidioidomycosis: Epidemiology. Clin Epidemiol 2013 Jun 25;5:185-97.

9. Burnie JP, Carter TL, Hodgetts SJ, Matthews RC. Fungal heat-shock proteins in human disease. FEMS Microbiol Rev 2006 Jan;30(1):53-88.

10. Centers for Disease Control and Prevention (CDC). Notice to readers: Final 2012 reports of nationally notifiable infectious diseases. MMWR Morb Mortal Wkly Rep 2013 Aug 23;62(33):669-82.

11. Centers for Disease Control and Prevention (CDC). Increase in reported coccidioidomycosis--united states, 1998-2011. MMWR Morb Mortal Wkly Rep 2013 Mar 29;62(12):217-21.

12. Champer J, Diaz-Arevalo D, Champer M, Hong TB, Wong M, Shannahoff M, Ito JI, Clemons KV, Stevens DA, Kalkum M. Protein targets for broad-spectrum mycosis vaccines: Quantitative proteomic analysis of aspergillus and *Coccidioides* and comparisons with other fungal pathogens. Ann N Y Acad Sci 2012 Dec;1273:44-51.

13. Chang DC, Anderson S, Wannemuehler K, Engelthaler DM, Erhart L, Sunenshine RH, Burwell LA, Park BJ. Testing for coccidioidomycosis among patients with community-acquired pneumonia. Emerg Infect Dis 2008 Jul;14(7):1053-9.

14. Cole GT, Kruse D, Seshan KR. Antigen complex of *Coccidioides* immitis which elicits a precipitin antibody response in patients. Infect Immun 1991 Jul;59(7):2434-46.

15. Cox RA, Britt LA. Antigenic identity of biologically active antigens in coccidioidin and spherulin. Infect Immun 1987 Nov;55(11):2590-6.

16. Cox RA, Magee DM. Coccidioidomycosis: Host response and vaccine development. Clin Microbiol Rev 2004 Oct;17(4):804,39, table of contents.

17. Dixon DM. *Coccidioides* immitis as a select agent of bioterrorism. J Appl Microbiol 2001 Oct;91(4):602-5.

18. Durkin M, Connolly P, Kuberski T, Myers R, Kubak BM, Bruckner D, Pegues D, Wheat LJ. Diagnosis of coccidioidomycosis with use of the *Coccidioides* antigen enzyme immunoassay. Clin Infect Dis 2008 Oct 15;47(8):e69-73.

19. Durkin M, Estok L, Hospenthal D, Crum-Cianflone N, Swartzentruber S, Hackett E, Wheat LJ. Detection of *Coccidioides* antigenemia following dissociation of immune complexes. Clin Vaccine Immunol 2009 Oct;16(10):1453-6.

20. Fisher FS, Bultman MW, Johnson SM, Pappagianis D, Zaborsky E. *Coccidioides* niches and habitat parameters in the southwestern united states: A matter of scale. Ann N Y Acad Sci 2007 Sep;1111:47-72.

21. Galgiani JN, Ampel NM, Blair JE, Catanzaro A, Johnson RH, Stevens DA, Williams PL, Infectious Diseases Society of America. Coccidioidomycosis. Clin Infect Dis 2005 Nov 1;41(9):1217-23.

22. Galgiani JN, Grace GM, Lundergan LL. New serologic tests for early detection of coccidioidomycosis. J Infect Dis 1991 Mar;163(3):671-4.

23. Galgiani JN, Grace GM, Lundergan LL. New serologic tests for early detection of coccidioidomycosis. J Infect Dis 1991 Mar;163(3):671-4.

24. Goto M. Protein O-glycosylation in fungi: Diverse structures and multiple functions. Biosci Biotechnol Biochem 2007 Jun;71(6):1415-27.

25. Guo H, Liu W, Ju Z, Tamboli P, Jonasch E, Mills GB, Lu Y, Hennessy BT, Tsavachidou D. An efficient procedure for protein extraction from formalin-fixed, paraffin-embedded tissues for reverse phase protein arrays. Proteome Sci 2012 Sep 24;10(1):56,5956-10-56.

26. Hector RF, Rutherford GW, Tsang CA, Erhart LM, McCotter O, Anderson SM, Komatsu K, Tabnak F, Vugia DJ, Yang Y, et al. The public health impact of coccidioidomycosis in arizona and california. Int J Environ Res Public Health 2011 Apr;8(4):1150-73.

27. Hung CY, Ampel NM, Christian L, Seshan KR, Cole GT. A major cell surface antigen of *Coccidioides* immitis which elicits both humoral and cellular immune responses. Infect Immun 2000 Feb;68(2):584-93.

28. Hurtgen BJ, Hung CY, Ostroff GR, Levitz SM, Cole GT. Construction and evaluation of a novel recombinant T cell epitope-based vaccine against coccidioidomycosis. Infect Immun 2012 Nov;80(11):3960-74.

29. Johnson R, Kernerman SM, Sawtelle BG, Rastogi SC, Nielsen HS, Ampel NM. A reformulated spherule-derived coccidioidin (spherusol) to detect delayed-type hypersensitivity in coccidioidomycosis. Mycopathologia 2012 Dec;174(5-6):353-8.

30. Kessner D, Chambers M, Burke R, Agus D, Mallick P. ProteoWizard: Open source software for rapid proteomics tools development. Bioinformatics 2008 Nov 1;24(21):2534-6.

31. Kim MM, Blair JE, Carey EJ, Wu Q, Smilack JD. Coccidioidal pneumonia, phoenix, arizona, USA, 2000-2004. Emerg Infect Dis 2009 Mar;15(3):397-401.

32. Kirsch EJ, Greene RT, Prahl A, Rubin SI, Sykes JE, Durkin MM, Wheat LJ. Evaluation of *Coccidioides* antigen detection in dogs with coccidioidomycosis. Clin Vaccine Immunol 2012 Mar;19(3):343-5.

33. Kruse D, Cole GT. Isolation of tube precipitin antibody-reactive fractions of *Coccidioides* immitis. Infect Immun 1990 Jan;58(1):169-78.

34. Kruse D, Cole GT. Isolation of tube precipitin antibody-reactive fractions of *Coccidioides* immitis. Infect Immun 1990 Jan;58(1):169-78.

35. Laboratories V. Table of lectin properties. 2014.

36. Ma ZQ, Dasari S, Chambers MC, Litton MD, Sobecki SM, Zimmerman LJ, Halvey PJ, Schilling B, Drake PM, Gibson BW, et al. IDPicker 2.0: Improved protein assembly with high discrimination peptide identification filtering. J Proteome Res 2009 Aug;8(8):3872-81.

37. Maes E, Broeckx V, Mertens I, Sagaert X, Prenen H, Landuyt B, Schoofs L. Analysis of the formalin-fixed paraffin-embedded tissue proteome: Pitfalls, challenges, and future prospectives. Amino Acids 2013 Aug;45(2):205-18.

38. Maes E, Valkenborg D, Mertens I, Broeckx V, Baggerman G, Sagaert X, Landuyt B, Prenen H, Schoofs L. Proteomic analysis of formalin-fixed paraffin-embedded colorectal cancer tissue using tandem mass tag protein labeling. Mol Biosyst 2013 Nov;9(11):2686-95.

39. Maes E, Valkenborg D, Mertens I, Broeckx V, Baggerman G, Sagaert X, Landuyt B, Prenen H, Schoofs L. Proteomic analysis of formalin-fixed paraffin-embedded colorectal cancer tissue using tandem mass tag protein labeling. Mol Biosyst 2013 Nov;9(11):2686-95.

40. Mahboob S, Mohamedali A, Ahn SB, Schulz-Knappe P, Nice E, Baker MS. Is isolation of comprehensive human plasma peptidomes an achievable quest? J Proteomics 2015 May 12.

41. Mahboob S, Mohamedali A, Ahn SB, Schulz-Knappe P, Nice E, Baker MS. Is isolation of comprehensive human plasma peptidomes an achievable quest? J Proteomics 2015 May 12.

42. Marimuthu A, O'Meally RN, Chaerkady R, Subbannayya Y, Nanjappa V, Kumar P, Kelkar DS, Pinto SM, Sharma R, Renuse S, et al. A comprehensive map of the human urinary proteome. J Proteome Res 2011 Jun 3;10(6):2734-43.

43. Marsden-Haug N, Hill H, Litvintseva AP, Engelthaler DM, Driebe EM, Roe CC, Ralston C, Hurst S, Goldoft M, Gade L, et al. *Coccidioides* immitis identified in soil outside of its known range - washington, 2013. MMWR Morb Mortal Wkly Rep 2014 May 23;63(20):450.

44. Martin K, McDougall BM, McIlroy S, Chen J, Seviour RJ. Biochemistry and molecular biology of exocellular fungal beta-(1,3)- and beta-(1,6)-glucanases. FEMS Microbiol Rev 2007 Mar;31(2):168-92.

45. Masannat FY, Ampel NM. Coccidioidomycosis in patients with HIV-1 infection in the era of potent antiretroviral therapy. Clin Infect Dis 2010 Jan 1;50(1):1-7.

46. Mora-Montes HM, Bates S, Netea MG, Castillo L, Brand A, Buurman ET, Diaz-Jimenez DF, Jan Kullberg B, Brown AJ, Odds FC, et al. A multifunctional mannosyltransferase family in candida albicans determines cell wall mannan structure and host-fungus interactions. J Biol Chem 2010 Apr 16;285(16):12087-95.

47. Nanjappa V, Thomas JK, Marimuthu A, Muthusamy B, Radhakrishnan A, Sharma R, Ahmad Khan A, Balakrishnan L, Sahasrabuddhe NA, Kumar S, et al. Plasma proteome database as a resource for proteomics research: 2014 update. Nucleic Acids Res 2014 Jan;42(Database issue):D959-65.

48. Neafsey DE, Barker BM, Sharpton TJ, Stajich JE, Park DJ, Whiston E, Hung CY, McMahan C, White J, Sykes S, et al. Population genomic sequencing of *Coccidioides* fungi reveals recent hybridization and transposon control. Genome Res 2010 Jul;20(7):938-46.

49. Nguyen C, Barker BM, Hoover S, Nix DE, Ampel NM, Frelinger JA, Orbach MJ, Galgiani JN. Recent advances in our understanding of the environmental, epidemiological, immunological, and clinical dimensions of coccidioidomycosis. Clin Microbiol Rev 2013 Jul;26(3):505-25.

50. Orsborn KI, Shubitz LF, Peng T, Kellner EM, Orbach MJ, Haynes PA, Galgiani JN. Protein expression profiling of *Coccidioides posadasii* by two-dimensional differential in-gel electrophoresis and evaluation of a newly recognized peroxisomal matrix protein as a recombinant vaccine candidate. Infect Immun 2006 Mar;74(3):1865-72.

51. Pappagianis D, Zimmer BL. Serology of coccidioidomycosis. Clin Microbiol Rev 1990 Jul;3(3):247-68.

52. Paulo JA, Kadiyala V, Brizard S, Banks PA, Steen H, Conwell DL. A proteomic comparison of formalin-fixed paraffin-embedded pancreatic tissue from autoimmune pancreatitis, chronic pancreatitis, and pancreatic cancer. Jop 2013 Jul 10;14(4):405-14.

53. Pisitkun T, Hoffert JD, Yu MJ, Knepper MA. Tandem mass spectrometry in physiology. Physiology (Bethesda) 2007 Dec;22:390-400.

54. Richter R, Schulz-Knappe P, Schrader M, Standker L, Jurgens M, Tammen H, Forssmann WG. Composition of the peptide fraction in human blood plasma: Database of circulating human peptides. J Chromatogr B Biomed Sci Appl 1999 Apr 16;726(1-2):25-35.

55. Saubolle MA. Laboratory aspects in the diagnosis of coccidioidomycosis. Ann N Y Acad Sci 2007 Sep;1111:301-14.

56. Schiess R, Wollscheid B, Aebersold R. Targeted proteomic strategy for clinical biomarker discovery. Mol Oncol 2009 Feb;3(1):33-44.

57. Schiess R, Wollscheid B, Aebersold R. Targeted proteomic strategy for clinical biomarker discovery. Mol Oncol 2009 Feb;3(1):33-44.

58. Sharpton TJ, Stajich JE, Rounsley SD, Gardner MJ, Wortman JR, Jordar VS, Maiti R, Kodira CD, Neafsey DE, Zeng Q, et al. Comparative genomic analyses of the human fungal pathogens *Coccidioides* and their relatives. Genome Res 2009 Oct;19(10):1722-31.

59. Sharpton TJ, Stajich JE, Rounsley SD, Gardner MJ, Wortman JR, Jordar VS, Maiti R, Kodira CD, Neafsey DE, Zeng Q, et al. Comparative genomic analyses of the human fungal pathogens *Coccidioides* and their relatives. Genome Res 2009 Oct;19(10):1722-31.

60. Shi SR, Taylor CR, Fowler CB, Mason JT. Complete solubilization of formalin-fixed, paraffin-embedded tissue may improve proteomic studies. Proteomics Clin Appl 2013 Apr;7(3-4):264-72.

61. Sprung RW,Jr, Brock JW, Tanksley JP, Li M, Washington MK, Slebos RJ, Liebler DC. Equivalence of protein inventories obtained from formalin-fixed paraffin-embedded and frozen tissue in multidimensional liquid chromatography-tandem mass spectrometry shotgun proteomic analysis. Mol Cell Proteomics 2009 Aug;8(8):1988-98.

62. Sunenshine RH, Anderson S, Erhart L, Vossbrink A, Kelly PC, Engelthaler D, Komatsu K. Public health surveillance for coccidioidomycosis in arizona. Ann N Y Acad Sci 2007 Sep;1111:96-102.

63. Sunenshine RH, Anderson S, Erhart L, Vossbrink A, Kelly PC, Engelthaler D, Komatsu K. Public health surveillance for coccidioidomycosis in arizona. Ann N Y Acad Sci 2007 Sep;1111:96-102.

64. Tabb DL, Fernando CG, Chambers MC. MyriMatch: Highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. J Proteome Res 2007 Feb;6(2):654-61.

65. Tamerius JD, Comrie AC. Coccidioidomycosis incidence in arizona predicted by seasonal precipitation. PLoS One 2011;6(6):e21009.

66. Tamerius JD, Comrie AC. Coccidioidomycosis incidence in arizona predicted by seasonal precipitation. PLoS One 2011;6(6):e21009.

67. Tanca A, Abbondio M, Pisanu S, Pagnozzi D, Uzzau S, Addis MF. Critical comparison of sample preparation strategies for shotgun proteomic analysis of formalin-fixed, paraffin-embedded samples: Insights from liver tissue. Clin Proteomics 2014 Jul 8;11(1):28,0275-11-28. eCollection 2014.

68. Thompson GR,3rd, Stevens DA, Clemons KV, Fierer J, Johnson RH, Sykes J, Rutherford G, Peterson M, Taylor JW, Chaturvedi V. Call for a california coccidioidomycosis consortium to

face the top ten challenges posed by a recalcitrant regional disease. Mycopathologia 2015 Feb;179(1-2):1-9.

69. Tsang CA, Anderson SM, Imholte SB, Erhart LM, Chen S, Park BJ, Christ C, Komatsu KK, Chiller T, Sunenshine RH. Enhanced surveillance of coccidioidomycosis, arizona, USA, 2007-2008. Emerg Infect Dis 2010 Nov;16(11):1738-44.

70. Valdivia L, Nix D, Wright M, Lindberg E, Fagan T, Lieberman D, Stoffer T, Ampel NM, Galgiani JN. Coccidioidomycosis as a common cause of community-acquired pneumonia. Emerg Infect Dis 2006 Jun;12(6):958-62.

71. Vincenti DC, Murray GI. The proteomics of formalin-fixed wax-embedded tissue. Clin Biochem 2013 Apr;46(6):546-51.

72. Viriyakosol S, Singhania A, Fierer J, Goldberg J, Kirkland TN, Woelk CH. Gene expression in human fungal pathogen *Coccidioides* immitis changes as arthroconidia differentiate into spherules and mature. BMC Microbiol 2013 May 28;13:121,2180-13-121.

73. Wakabayashi M, Yoshihara H, Masuda T, Tsukahara M, Sugiyama N, Ishihama Y. Phosphoproteome analysis of formalin-fixed and paraffin-embedded tissue sections mounted on microscope slides. J Proteome Res 2014 Feb 7;13(2):915-24.

74. Weiner MH. Antigenemia detected in human coccidioidomycosis. J Clin Microbiol 1983 Jul;18(1):136-42.

75. Whiston E, Taylor JW. Genomics in *Coccidioides*: Insights into evolution, ecology, and pathogenesis. Med Mycol 2014 Feb;52(2):149-55.

76. Yang C, Zhu Y, Magee DM, Cox RA. Molecular cloning and characterization of the *Coccidioides* immitis complement fixation/chitinase antigen. Infect Immun 1996 Jun;64(6):1992-7.

77. Yuan L, Cole GT, Sun SH. Possible role of a proteinase in endosporulation of *Coccidioides* immitis. Infect Immun 1988 Jun;56(6):1551-9.

78. Zhang B, Chambers MC, Tabb DL. Proteomic parsimony through bipartite graph analysis improves accuracy and transparency. J Proteome Res 2007 Sep;6(9):3549-57.

APPENDIX A

COMPLETE LIST OF PROTEINS IDENTIFIED IN SPHERULIN

The complete list of protein can be accessed via an open source resource.

MassIVE Link to Raw Data, mzID files and the FASTA Database	
Web Link	https://massive.ucsd.edu/ProteoSAFe/status.jsp?task=03125bd6c85648488be65c6210e3c855
FTP Link	ftp://MSV000079286@massive.ucsd.edu/
Password	Lake_Grys_Mayo

Database Details	
Name	20150522-*Coccidioides*-Broad-Sprot-RefSeq-20130621-RefSeq59-Human-Bovine-Cntms.fasta

Sequences Extracted From:			Species	Link
	Broad Institute	Downloaded on 05/22/2014	*Coccidioides*	http://www.broadinstitute.org/scientific-community/science/projects/fungal-genome-initiative/*Coccidioides*-genomes
	RefSeq	Downloaded on 05/22/2014	*Coccidioides*	http://www.ncbi.nlm.nih.gov/refseq/
	SwissProt	Downloaded on 05/22/2014	*Coccidioides*	http://www.uniprot.org/uniprot/?query=reviewed%3Ayes
	RefSeq	Downloaded on 06/21/2013	Human	http://www.ncbi.nlm.nih.gov/refseq/
	RefSeq	Downloaded on 05/22/2012	Bovine	http://www.ncbi.nlm.nih.gov/refseq/
Total Number of Sequences = 157600				

APPENDIX B

COMPLETE LIST OF PROTEINS IDENTIFIED IN PLASMA AND URINE

The complete list of proteins identified in plasma and urine can be accessed using a free software- IDpicker (Ma,Z.Q. 2009).

The source code and binaries for the latest version of IDPicker are available from http://fenchurch.mc.vanderbilt.edu/.

File can be accessed via the following link- https://drive.google.com/file/d/0BzqBZK5wIxiXMExUWHlyTGNzR2c/view?usp=sharing

APPENDIX C

COMPLETE LIST OF PROTEINS IDENTIFIED IN PROTEIN LYSATE OF INFECTED LUNG TISSUE AND COCCIDIOIDIN LYSATE

Identified Proteins	Present in CDN-L	Present in lung
| CPSG_08657 | *Coccidioides posadasii* Silveira complement fixation-chitinase (428 aa)	yes	yes
| CPSG_07243 | *Coccidioides posadasii* Silveira superoxide dismutase (155 aa)	yes	yes
| CPSG_02802 | *Coccidioides posadasii* Silveira exo-beta-1,3-glucanase Exg0 (889 aa)	yes	yes
| CPSG_02648 | *Coccidioides posadasii* Silveira alpha-mannosidase (520 aa)	yes	yes
| CPSG_03385 | *Coccidioides posadasii* Silveira beta-glucosidase (859 aa)	yes	yes
| CPSG_02027 | *Coccidioides posadasii* Silveira hypothetical protein (147 aa)	yes	yes
| CPSG_06408 | *Coccidioides posadasii* Silveira alkaline phosphatase (607 aa)	yes	yes
| CPSG_05583 | *Coccidioides posadasii* Silveira metalloprotease MEP1 (277 aa)	yes	yes
| CPSG_05770 | *Coccidioides posadasii* Silveira FAD-dependent oxygenase (515 aa)	yes	yes
| CPSG_08500 | *Coccidioides posadasii* Silveira endo-1,3-beta-glucanase (885 aa)	yes	yes
| CPSG_02111 | *Coccidioides posadasii* Silveira conserved hypothetical protein (285 aa)	yes	yes
| CPSG_05610 | *Coccidioides posadasii* Silveira leucine aminopeptidase (369 aa)	yes	yes
| CPSG_00196 | *Coccidioides posadasii* Silveira conserved hypothetical protein (497 aa)	yes	yes
| CPSG_03152 | *Coccidioides posadasii* Silveira conserved hypothetical protein (391 aa)	yes	yes
| CPSG_03850 | *Coccidioides posadasii* Silveira wall-associated proteinase (310 aa)	yes	yes
| CPSG_08739 | *Coccidioides posadasii* Silveira alkaline protease (400 aa)	yes	yes
| CPSG_07196 | *Coccidioides posadasii* Silveira aspartyl aminopeptidase (504 aa)	yes	yes
| CPSG_05091 | *Coccidioides posadasii* Silveira sialidase (356 aa)	yes	yes
| CPSG_06651 | *Coccidioides posadasii* Silveira conserved hypothetical protein (693 aa)	yes	yes
| CPSG_01581 | *Coccidioides posadasii* Silveira iron/manganese superoxide dismutase (236 aa)	yes	yes
| CPSG_07381 | *Coccidioides posadasii* Silveira rds1 (484 aa)	yes	yes
| CPSG_07353 | *Coccidioides posadasii* Silveira secreted dipeptidyl peptidase (724 aa)	yes	yes
| CPSG_01038 | *Coccidioides posadasii* Silveira beta-glucosidase (307 aa)	yes	yes

| CPSG_08872 | *Coccidioides posadasii* Silveira chitinase (353 aa)	yes	yes
| CPSG_02766 | *Coccidioides posadasii* Silveira conserved hypothetical protein (197 aa)	yes	yes
| CPSG_02828 | *Coccidioides posadasii* Silveira gamma-glutamyltranspeptidase (593 aa)	yes	yes
| CPSG_01482 | *Coccidioides posadasii* Silveira copper amine oxidase (802 aa)	yes	yes
| CPSG_08248 | *Coccidioides posadasii* Silveira acid trehalase (1074 aa)	yes	yes
| CPSG_04449 | *Coccidioides posadasii* Silveira hypothetical protein (225 aa)	yes	yes
| CPSG_07055 | *Coccidioides posadasii* Silveira carboxypeptidase (652 aa)	yes	yes
| CPSG_08841 | *Coccidioides posadasii* Silveira peptidase (453 aa)	yes	yes
| CPSG_04011 | *Coccidioides posadasii* Silveira DUF1237 domain-containing protein (551 aa)	yes	yes
| CPSG_04082 | *Coccidioides posadasii* Silveira serine peptidase, family S28 (556 aa)	yes	yes
| CPSG_04657 | *Coccidioides posadasii* Silveira conserved hypothetical protein (661 aa)	yes	yes
| CPSG_03507 | *Coccidioides posadasii* Silveira glutaminase GtaA (857 aa)	yes	yes
| CPSG_04717 | *Coccidioides posadasii* Silveira serine protease (401 aa)	yes	yes
| CPSG_08888 | *Coccidioides posadasii* Silveira conserved hypothetical protein (504 aa)	yes	yes
| CPSG_06612 | *Coccidioides posadasii* Silveira acetamidase (589 aa)	yes	yes
| CPSG_09845 | *Coccidioides posadasii* Silveira non-hemolytic phospholipase C (708 aa)	yes	yes
| CPSG_02656 | *Coccidioides posadasii* Silveira Ser/Thr protein phosphatase family (632 aa)	yes	yes
| CPSG_07236 | *Coccidioides posadasii* Silveira serine protease (498 aa)	yes	yes
| CPSG_00137 | *Coccidioides posadasii* Silveira carboxypeptidase (601 aa)	yes	yes
| CPSG_07819 | *Coccidioides posadasii* Silveira conserved hypothetical protein (496 aa)	yes	yes
| CPSG_09825 | *Coccidioides posadasii* Silveira conserved hypothetical protein (232 aa)	yes	yes
| CPSG_04840 | *Coccidioides posadasii* Silveira acetylcholinesterase (581 aa)	yes	yes
| CPSG_08492 | *Coccidioides posadasii* Silveira extracelular serine carboxypeptidase (544 aa)	yes	yes
| CPSG_07483 | *Coccidioides posadasii* Silveira conserved hypothetical protein (453 aa)	yes	yes
| CPSG_04742 | *Coccidioides posadasii* Silveira adenosylhomocysteinase (450 aa)	yes	yes
| CPSG_08329 | *Coccidioides posadasii* Silveira aminopeptidase (890 aa)	yes	yes

| CPSG_08627 | *Coccidioides posadasii* Silveira conserved hypothetical protein (393 aa)	yes	yes
| CPSG_04000 | *Coccidioides posadasii* Silveira endonuclease/exonuclease/phosphatase (291 aa)	yes	yes
| CPSG_06347 | *Coccidioides posadasii* Silveira hypothetical protein (504 aa)	yes	yes
| CPSG_09707 | *Coccidioides posadasii* Silveira chitosanase (242 aa)	yes	yes
| CPSG_09885 | *Coccidioides posadasii* Silveira glucose-6-phosphate isomerase (554 aa)	yes	yes
| CPSG_00906 | *Coccidioides posadasii* Silveira aminopeptidase (612 aa)	yes	yes
| CPSG_03388 | *Coccidioides posadasii* Silveira tripeptidyl peptidase SED3 (600 aa)	yes	yes
| CPSG_07314 | *Coccidioides posadasii* Silveira isochorismatase family hydrolase (196 aa)	yes	yes
| CPSG_06733 | *Coccidioides posadasii* Silveira conserved hypothetical protein (165 aa)	yes	yes
| CPSG_08913 | *Coccidioides posadasii* Silveira leucine aminopeptidase (504 aa)	yes	yes
| CPSG_03208 | *Coccidioides posadasii* Silveira 5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase (775 aa)	yes	yes
| CPSG_09790 | *Coccidioides posadasii* Silveira ubiquitin (626 aa)	yes	yes
| CPSG_06520 | *Coccidioides posadasii* Silveira conserved hypothetical protein (1228 aa)	yes	yes
| CPSG_08790 | *Coccidioides posadasii* Silveira stress response protein Rds1 (456 aa)	yes	yes
| CPSG_04302 | *Coccidioides posadasii* Silveira pyridine nucleotide-disulphide oxidoreductase (565 aa)	yes	yes
| CPSG_02467 | *Coccidioides posadasii* Silveira aldose 1-epimerase (433 aa)	yes	yes
| CPSG_08295 | *Coccidioides posadasii* Silveira hypothetical protein (276 aa)	yes	yes
| CPSG_07273 | *Coccidioides posadasii* Silveira conserved hypothetical protein (339 aa)	yes	yes
| CPSG_06448 | *Coccidioides posadasii* Silveira hypothetical protein (217 aa)	yes	yes
| CPSG_04321 | *Coccidioides posadasii* Silveira 3-isopropylmalate dehydrogenase A (364 aa)	yes	yes
| CPSG_08652 | *Coccidioides posadasii* Silveira actin (295 aa)	yes	yes
| CPSG_08451 | *Coccidioides posadasii* Silveira alcohol dehydrogenase (349 aa)	yes	yes
| CPSG_00066 | *Coccidioides posadasii* Silveira hypothetical protein (223 aa)	yes	yes
| CPSG_04381 | *Coccidioides posadasii* Silveira glutamyl-tRNA(Gln) amidotransferase subunit A (484 aa)	yes	yes
| CPSG_08488 | *Coccidioides posadasii* Silveira ribonuclease T2-like (412 aa)	yes	yes

| CPSG_04612 | *Coccidioides posadasii* Silveira extracellular elastinolytic metalloproteinase (640 aa)	yes	yes
| CPSG_09734 | *Coccidioides posadasii* Silveira beta-1,3-glucanosyltransferase (474 aa)	yes	yes
| CPSG_01618 | *Coccidioides posadasii* Silveira glucoamylase (627 aa)	yes	yes
| CPSG_06407 | *Coccidioides posadasii* Silveira cell wall glucanase (439 aa)	yes	yes
| CPSG_02747 | *Coccidioides posadasii* Silveira autophagy protein Atg27 (312 aa)	yes	yes
| CPSG_01885 | *Coccidioides posadasii* Silveira BYS1 domain-containing protein (155 aa)	yes	yes
| CPSG_02334 | *Coccidioides posadasii* Silveira nuclear transport factor (124 aa)	yes	yes
| CPSG_02925 | *Coccidioides posadasii* Silveira chaperonin (103 aa)	yes	yes
| CPSG_09496 | *Coccidioides posadasii* Silveira endoglucanase (215 aa)	yes	yes
| CPSG_02568 | *Coccidioides posadasii* Silveira histone H4 (104 aa)	yes	yes
| CPSG_02793 | *Coccidioides posadasii* Silveira conserved hypothetical protein (234 aa)	yes	yes
| CPSG_06794 | *Coccidioides posadasii* Silveira nicotinate-nucleotide pyrophosphorylase (315 aa)	yes	yes
| CPSG_04555 | *Coccidioides posadasii* Silveira 2-methylcitrate dehydratase (552 aa)	yes	yes
| CPSG_08186 | *Coccidioides posadasii* Silveira conserved hypothetical protein (322 aa)	yes	yes
| CPSG_01960 | *Coccidioides posadasii* Silveira lactoylglutathione lyase (323 aa)	yes	no
| CPSG_05793 | *Coccidioides posadasii* Silveira formamidase (414 aa)	yes	no
| CPSG_00803 | *Coccidioides posadasii* Silveira conserved hypothetical protein (593 aa)	yes	no
| CPSG_05792 | *Coccidioides posadasii* Silveira peptidyl-prolyl cis-trans isomerase (374 aa)	yes	no
| CPSG_08956 | *Coccidioides posadasii* Silveira conserved hypothetical protein (664 aa)	yes	no
| CPSG_06865 | *Coccidioides posadasii* Silveira conserved hypothetical protein (285 aa)	yes	no
| CPSG_09956 | *Coccidioides posadasii* Silveira beta-1,3-glucanosyltransferase (529 aa)	yes	no
| CPSG_00512 | *Coccidioides posadasii* Silveira conserved hypothetical protein (539 aa)	yes	no
| CPSG_03061 | *Coccidioides posadasii* Silveira dipeptidyl peptidase (708 aa)	yes	no
| CPSG_08830 | *Coccidioides posadasii* Silveira tyrosinase (412 aa)	yes	No
| CPSG_00188 | *Coccidioides posadasii* Silveira dipeptidyl aminopeptidase (918 aa)	yes	no
| CPSG_00354 | *Coccidioides posadasii* Silveira hsp70-like protein (651 aa)	yes	no

| CPSG_00752 | *Coccidioides posadasii* Silveira conserved hypothetical protein (822 aa)	yes	no
| CPSG_05591 | *Coccidioides posadasii* Silveira conserved hypothetical protein (752 aa)	yes	no
| CPSG_04907 | *Coccidioides posadasii* Silveira 1,3-b-glucanosyltransferase (448 aa)	yes	no
| CPSG_09883 | *Coccidioides posadasii* Silveira conserved hypothetical protein (758 aa)	yes	no
| CPSG_08837 | *Coccidioides posadasii* Silveira enolase (408 aa)	yes	no
| CPSG_05811 | *Coccidioides posadasii* Silveira alkaline serine protease (398 aa)	yes	no
| CPSG_05817 | *Coccidioides posadasii* Silveira YjgH family protein (165 aa)	yes	no
| CPSG_08319 | *Coccidioides posadasii* Silveira lysophospholipase (630 aa)	yes	no
| CPSG_01055 | *Coccidioides posadasii* Silveira conserved hypothetical protein (249 aa)	yes	no
| CPSG_07165 | *Coccidioides posadasii* Silveira conserved hypothetical protein (193 aa)	yes	no
| CPSG_03569 | *Coccidioides posadasii* Silveira elongation factor 1-alpha (461 aa)	yes	no
| CPSG_07164 | *Coccidioides posadasii* Silveira conserved hypothetical protein (319 aa)	yes	no
| CPSG_06252 | *Coccidioides posadasii* Silveira carboxypeptidase (536 aa)	yes	no
| CPSG_08693 | *Coccidioides posadasii* Silveira cytochrome c peroxidase (319 aa)	yes	no
| CPSG_03642 | *Coccidioides posadasii* Silveira conserved hypothetical protein (399 aa)	yes	no
| CPSG_00802 | *Coccidioides posadasii* Silveira ATP synthase subunit beta (522 aa)	yes	no
| CPSG_08245 | *Coccidioides posadasii* Silveira Ser/Thr protein phosphatase family (604 aa)	yes	no
| CPSG_08571 | *Coccidioides posadasii* Silveira conserved hypothetical protein (306 aa)	yes	no
| CPSG_06381 | *Coccidioides posadasii* Silveira 14-3-3 family protein ArtA (266 aa)	yes	no