Title

Uncovering virulence pathways facilitated by proteolysis in HIV and a HIV associated fungal pathogen, Cryptococcus neoformans

Permalink

Author

Clarke, Starlynn Cascade

Publication Date

2015

Uncovering virulence pathways facilitated by proteolysis in HIV and
a HIV associated fungal pathogen, *Cryptococcus neoformans*

by

Starlynn Clarke

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry and Molecular Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Acknowledgments

I would first like to thank my thesis advisor, Dr. Charles Craik. Throughout my PhD Charly has been unfailingly optimistic and enthusiastic about my projects, but most importantly, he has always had confidence in my abilities. During the course of graduate school I have doubted myself and my capabilities almost daily, but Charly has always believed that I would be successful in my scientific endeavors and I cannot thank him enough for that. Charly is the most enthusiastic scientist that I have ever met and he has the capacity to find a silver lining to almost any event. He also understands the importance of presentation and has the resourcefulness to transform almost any situation into an opportunity. These are all traits that do not come naturally to me, but having Charly as a mentor has helped me to learn some of these important skills.

I would also like to thank my thesis committee members, Dr. Raul Andino and Dr. John Gross who have both been incredibly supportive over the years. Despite the fact that my research ended up veering away from the original focus that was more in line with their expertise, they have continued to provide me with encouragement and thoughtful feedback during my thesis committee meetings as well as at other times that I sought their advice.

Dr. Anthony O'Donoghue was my friend and mentor in the Craik lab for five years. His guidance and knowledge were instrumental in my research success. Besides that, Anthony is one of the most good-natured people that I have ever met and in spite of the long list of demands on his time, he always answered my questions and even went out of his way to check up on me and offer to discuss my project. When I was struggling with a project that was not working he helped me to find a new direction, something that I will always be grateful to him for.

I would like to thank Dr. Hiten Madhani, Dr. Phillip Dumesic and Christina Homer, my collaborators on my research project studying peptidases secreted by *Cryptococcus neoformans*. I first proposed the project to Phillip and despite his extremely busy schedule with

his numerous and successful projects, he made time to plan experiments with me and help me with reagents, experiments, and edits to our manuscript. Likewise, Christina has been a wealth of advice on the project as well as generously contributing her time to helping me with mouse experiments and editing our manuscript. Hiten has been supportive and helpful on the project as well. I am very fortunate to have had an expert in the field to help me construct the story and determine how best to present it for publication.

My favorite aspect of the Craik lab is the wonderful people that have worked in the lab over the years. In addition to Anthony, Dr. Dan Hostetter, Dr. Jungmin Kim and Dr. Michael Winter were my mentors and friends in the Craik lab. All extremely knowledgeable, they have always been more than willing to answer any questions I have had, about science as well as career choices or life in general. However, there has always been a strong sense of camaraderie and support amongst all of the Craik lab members. At some point, I have sought help from almost everyone in the lab and been surprised and touched by how generously everyone has shared their time and expertise. In particular, I would like to thank Dr. Cheryl Tajon for helping me make and test nanosensors for HIV protease and the graduate students of my same year Dr. Jonathan Gable and Dr. Melody Lee for their support as we each achieved the mile stones of graduate school.

I have met so many wonderful people at UCSF, first and foremost, my class mate, room mate and friend, Kelly Nissen. Kelly is the most supportive person that I have ever known and also a wonderful scientist. She has given me pep talks countless times throughout graduate school as well as thoughtful advice on my scientific writings and presentations. Dr. Manon Eckhardt is also a wonderful friend that I have made at UCSF. She is an excellent scientist and a fun and interesting person and has cheered me up many times. Nicole Olsen, a graduate student in the Craik lab has been a great friend to me and made lab a more enjoyable place.

Finally, I would like to thank my mother Felice Clarke and my non-scientist friends for their support and love during this long process. They have always been there for me when things

were not going well and when good things happen they are ready to celebrate with me.

I have had the opportunity to work on several great projects at UCSF and publish with some amazing scientists. I am sixth author on Dr. Stefanie Jager's paper "Global landscape of HIV-human protein complexes" Pubmed ID: PMC3310911 where I contributed studies of the interactions between HIV protease and the eukaryotic translation initiation complex 3 (eIF3). I am fourth author on Christina Homer's paper "Convergent evolution of a peptide-based cell-cell signaling system required for the virulence of a eukaryotic pathogen" (submitted to Cell). For this project I helped to determine that the peptidase secretion of *C. neoformans* is affected by a quorum sensing pathway discovered by Christina. I am second author on Dr. Steven Jones's paper "Evolutionary selection on barrier activity: The aspartyl protease Bar1 is an enzyme with novel substrate specificity" (submitted to mBio) where I contributed to identifying the unique specificity of a peptidase secreted by *Candida albicans* called Bar1. Finally, my findings on the secreted peptidases of *C. neoformans* will be published in collaboration with Dr. Phillip Dumesic, Dr. Christina Homer, Dr. Anthony O'Donoghue, Lenka, Pallova, Dr. Pavel Maher, Dr. Hiten Madhani and Dr. Charles Craik, "Integrated activity and genetic profiling of secreted peptidases in Cryptococcus neoformans reveals an asparty peptidase required for low pH survival and virulence" (submitted to PLOS Pathogens).

Uncovering virulence pathways facilitated by proteolysis in HIV and a HIV associated fungal pathogen, *Cryptococcus neoformans*

By

Starlynn Clarke

Abstract

Peptidases are enzymes that cleave proteins or peptides, irreversibly modifying their substrates. Proteolytic activity is essential for many biological processes, however these enzymes are also associated with many pathological states. In particular, peptidases are frequently used by pathogenic organisms for purposes such as host invasion, immune system evasion and nutrient acquisition. Both HIV and the HIV/AIDS associated fungal pathogen, *Cryptococcus neoformans*, are known to employ peptidases to promote virulence. HIV expresses a protease that has a well characterized role in maturation of infectious viral particles by cleaving viral polyproteins into their individual constituents. However, in this work we present evidence that this enzyme has a dual purpose. Our results indicate that in addition to processing viral proteins, HIV protease cleaves numerous host-proteins to optimize the host environment for viral replication. We conducted a study of HIV protease host-protein interactions using an affinity purification-mass spectrometry based approach with tagged, inactive HIV protease. Over 50 candidate substrate proteins were identified through this technique. We then individually studied over 40 of these proteins, revealing that eight are cleaved by HIV protease *in vitro*. Through experiments assaying viral replication in combination with siRNA knock-down of host proteins, we discovered one

of these host-proteins, the eukaryotic translation initiation factor 3d (eIF3d) restricts viral reverse transcription. Therefore, we hypothesize that cleavage of this protein by HIV protease may promote reverse transcription during viral infection.

In this work we also investigate the secreted peptidases of *C. neoformans*, an opportunistic fungal pathogen responsible for over 600,000 deaths every year. This organism is known to secrete peptidases but the functions, regulation and characteristics of these enzymes have remained mysterious. We used an unbiased activity based approach to detail the specificity and regulation of extracellular peptidase activity produced by this pathogen grown under two different culture conditions. We pinpoint a principle enzyme responsible for secreted peptidase activity and find that this enzyme is important for tolerance to low pH environments as well as virulence in a mouse model of cryptococcus infection. Using knowledge of the substrate specificity of this enzyme, we screen a focused library of small molecule inhibitors, identifying four potent compounds.

Our study of the functions and characteristics of peptidases from viral and fungal pathogens reveals important information regarding the host-pathogen biology of these organisms. Our results may be informative for the development of alternative therapeutic strategies for combating these intractable pathogens. Furthermore, the approaches and techniques we describe are useful for the investigation of proteolytic activity secreted by other disease-causing organisms.

TABLE OF CONTENTS

CHAPTER 7: CONCLUSIONS AND FUTURE DIRECTIONS 201

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction

1.1 Peptidases: an overview

Peptidases are a large and diverse group of enzymes which have in common the ability to hydrolyze peptide bonds, thus permanently altering their protein or peptide substrates. These enzymes are essential to physiological processes including digestion, blood-clotting, cell turnover, and regulation of the immune response, among many others [1-4]. Aberrant peptidase activity is associated with pathological states including cancer, neurodegenerative and autoimmune diseases, thus the regulation of these enzymes is critical to their proper function [5-11]. The importance of peptidases is underscored by the fact that they represent over 2% of human genes, or approximately 500-600 unique proteolytic enzymes [12].

The range of substrates that must be cleaved to facilitate the aforementioned biological processes and the variety of conditions under which peptidase activity is required, has led to extensive enzymatic diversity. Seven separate classes of proteolytic enzymes have been identified based on the mechanisms that they use for catalysis. Further subdivisions determined through homology have grouped all known peptidases into a total of 258 families [2]. Some of these enzymes are capable of highly specific substrate selection such as the tobacco etch virus peptidase which has a strict seven amino acid recognition sequence, while others for example the digestive enzyme trypsin, accept a wide range of substrates with recognition of only a single amino acid required [2].

Peptidase activity is subject to multiple layers of regulation by substrate features, environmental conditions and the presence of inhibitors or activators. Through study of the biochemical characteristics of peptidases, their regulation and their substrate specificity, an abundance of information regarding biological pathways can be obtained. Because proteolytic activity is frequently involved in disease states, characterization of these enzymatic activities is important for both therapeutic and diagnostic purposes.

1.2. Investigating substrate specificity

One of the most important and challenging aspects of peptidase biology is uncovering the substrate specificity of these enzymes and the endogenous proteins that they cleave. In general, peptidase substrate specificity is largely dictated through interactions between four residues on either side of the scissile bond, referred to as the P4-P4' amino acids, and corresponding regions of the enzyme termed the S4-S4' pockets [2,13]. However, the importance of each position can differ greatly between enzymes and exceptions exist, for example an aspartyl peptidase was recently characterized whose substrate selection appears to be entirely determined by the P5'-P8' residues of the substrate (see Chapter 6). The context of the cleavage site within the substrate is also important. The majority of peptidases prefer to cleave at locations within flexible loops of a protein which allow the greatest access to the residues surrounding the cleavage site, although some peptidases recognize substrates within structural elements of proteins. Substrate selection by exopeptidases is affected by similar factors to that of endopeptidases. However these enzymes cleave a single amino acid off of either the amino or carboxyl termini of a protein, meaning that the substrates of aminopeptidases lack P4-P2 residues, while carboxypeptidases cleave substrates lacking P2'-P4' residues.

Historically, the most common approach to characterizing a newly discovered peptidase activity has been through protein hydrolysis assays. In these experiments, proteins such as casein or bovine serum albumin (BSA) are incubated with the peptidase of interest, the sample is run on a gel and protein staining is used to determine whether proteolysis occurred. While this technique can reveal information such as the class of proteolytic enzyme, it is lacking any details regarding substrate features.

Recent technological innovations involving the use of fluorescently labeled peptide substrates, phage display and mass spectrometry have enabled characterization of peptidase

activity at a far more sophisticated level. In the first instance, a peptide library of diverse sequences is labeled at one terminus with a fluorophore and at the other terminus with a quencher. The peptidase under study is then combined with individual substrates from this library and the preference of the enzyme for the substrate can be measured in real-time by following the increase in fluorescence that is produced upon peptide cleavage [14]. This system has many advantages over a traditional protein hydrolysis approach, such as increased speed and the ability to quantify activity. However, it still has a number of drawbacks, perhaps the most significant of which is the large number of substrates that must be designed and tested to comprehensively interrogate substrate specificity. This leads to prohibitively high synthesis costs as well as a requirement for large amounts of enzyme to individually assay all substrates. An alternative approach that avoids some of these issues is the use of phage display to generate the substrates. In this technique, a diverse library of peptides linked to a fluorophore is expressed on the surface of *E. coli*. Upon exposure to the peptidase of interest, substrate cleavage removes the fluorophore from the cells causing a reduction in fluorescence [15]. The cells can then be sorted by fluorescence-activated cell sorting (FACS) to identify those bearing substrate peptides. While this approach is less expensive than *in vitro* synthesis and avoids issues such as solubility of the substrates, it is time intensive and has technical limitations including an inability to detect exopeptidase activity.

A different approach to determining substrate specificity relies on mass spectrometry rather than fluorescence.   In this type of assay, a library of diverse peptide sequences is incubated with the peptidase of interest, followed by mass spectrometry to identify the sequences of cleavage sites. Several variations on this approach exist, but in general the most substantial advantages of these techniques are that greater substrate diversity can be tested than would be practical using fluorescent substrates, much smaller amounts of enzyme can be used and all or most of the substrates can be analyzed simultaneously. The origin of the libraries can be either synthetic or derived from a natural proteome. Libraries derived from

whole proteomes are useful in that they can potentially reveal not only substrate specificity, but also endogenous protein substrates. A disadvantage to this approach is that the complexity of a proteome makes parsing the resulting data extremely difficult [16,17]. The unequal representation of amino acids within different proteomes may also bias the resulting data since all substrates are not present with equal frequency.

The use of a synthetic peptide library in combination with mass spectrometry addresses several drawbacks with the aforementioned approaches [18,19]. Because all sequences present in the library are known and the sample is less complex and lacks contaminating enzymes, determining which substrates are cleaved and the locations of the cleavage sites is far simpler than with the previously described techniques. Additionally, factors such as substrate concentration can be easily optimized for the experimental conditions and the results are more reproducible than with analysis of proteome derived libraries. Algorithms can be used to design unbiased libraries encompassing maximum substrate diversity in a relatively limited number of peptides and analysis by mass spectrometry instead of fluorescence permits the use of much smaller quantities of peptides and enzyme, both factors that reduce costs. Finally, by avoiding the use of fluorophores, the termini of the peptides can remain unmodified, allowing detection of exopeptidases as well as endopeptidases. This strategy for identification of peptidase substrate specificity is discussed in more detail in Chapter 4.

1.3 Pathogens use peptidases to mediate host-protein interactions

The mechanisms used by pathogens to cause disease are as varied as the organisms themselves. However, all pathogens must interact with the proteins of their host to facilitate infection. Many disease-causing organisms including viruses, bacteria, fungi and parasites secrete enzymes such as peptidases that modify host pathways in order to promote survival and virulence. Secreted peptidases have been demonstrated to have multiple functions in

pathogenesis, including adhesion to and invasion of host tissues, immune system evasion and nutrient acquisition. Examples from diverse pathogen types include *Vibrio cholera* which secretes peptidases promoting tissue invasion, the fungal pathogen *Candida albicans* whose secreted aspartyl peptidases (SAPs) function in adhesion, invasion and tissue degradation, and parasitic helminths that express proteases facilitating multiple processes such as migration and host immune depression [20-22].

In addition to general degradative functions that promote processes such as nutrient acquisition and tissue invasion, many pathogens have evolved sophisticated peptidase activities that selectively cleave a limited number of host proteins to facilitate hijacking of specific pathways. Multiple examples of this type of proteolysis have been elucidated in viruses [23]. As obligate intracellular pathogens relying heavily on host proteins for replication, viruses have developed sophisticated strategies for manipulating host cells into producing large quantities of virions. A pathway frequently targeted by viral peptidases is the eukaryotic protein translation initiation machinery. Because viruses often have fewer requirements for protein translation initiation than their hosts, they cleave or inactivate the proteins that they do not need, thus suppressing host protein production and diverting resources to translation of viral proteins. This strategy is used effectively by virus families such as picornaviridae and caliciviridiae [24,25]. For example, the leader protease (L-pro) and 2A protease of picornoviruses including foot and mouth disease virus (FMDV), rhinovirus and polio cleave the host translation proteins eukaryotic translation initiation factor 4GI (eIF4GI) and 4GII (eIF4GII) [26,27]. Whereas caliciviruses have been found to cleave eIF4GI/II and an additional translation initiation protein, the poly(A)-binding protein (PABP) [28,29] .

Through the study of peptidases produced by pathogens, information about pathways facilitating virulence can be revealed. When it is found that an enzyme has an important role in survival or virulence, its direct targeting can be an effective therapeutic strategy. Described below are two instances of pathogenic organisms that express peptidases critical for virulence

and a detailed discussion regarding the functions of these enzymes and the therapeutic

opportunities arising from their study.

1.4 HIV protease, new roles for a canonical virulence factor

One of the most successful examples of therapeutic intervention targeting a pathogen-

specific peptidase activity is that of the human immunodeficiency virus (HIV) protease. The HIV

genome encodes 15 proteins, seven of which are synthesized as two polypeptide chains [30].

However, in order to produce infectious virions these polyproteins must be released into their

constituent components through proteolysis, a process referred to as viral maturation. HIV

protease, a component of one of the two polyproteins, is the enzyme responsible for facilitating

viral maturation through a series of highly regulated proteolytic events [30-32]. Inhibiting the

activity of HIV protease halts this process and causes a dramatic reduction in viremia. Drugs

targeting this enzyme were introduced at the peak of the HIV epidemic and it was the

effectiveness of these compounds that slowed the exponentially increasing HIV/AIDS mortality

rates, eventually helping to transform a certain death sentence into a manageable, although

incurable, condition [33,34].

After the study of HIV protease clearly revealed its essential function in viral protein

processing, it became established within the scientific community that this was the only function

of this enzyme. Indeed, this dogma became so firmly entrenched that few researchers

considered the possibility that this enzyme could have additional roles. However, gradually

evidence for another function of HIV protease has begun to accumulate. The mounting

evidence supports a role for this enzyme in modulating the host environment to facilitate viral

replication, a function that has been established for a range of viral proteases (see section 1.3).

One factor supporting a role for HIV protease host-protein interactions is the discovery of the

many host-protein interactions that are conducted by the other 14 HIV proteins. These include

the findings that the viral accessory proteins modulate several host pathways, HIV integrase

recruits host proteins into the viral particle and gag interacts with multiple host proteins to facilitate virion assembly and release [35-40]. As the list of verified host-protein interactions of other HIV proteins continues to grow, it is beginning to seem improbably that out of all of the HIV proteins only HIV protease has no impact on host pathways.

Direct evidence for interactions between HIV protease and host pathways originated first from reports that recombinant HIV protease is able to cleave multiple cytoskeletal proteins *in vitro*, including vimentin, desmin and actin (Table 1.1) [41-43]. Later reports revealed cleavage of proteins involved in immune activation such as nuclear factor kappa B (NF-κB) and N-myc downstream regulated 1 and 2 (NDR1 and NDR2) as well as the eukaryotic translation proteins eIF4GI and PABP (Table 1.1). These reports were met with some skepticism due to the argument that they did not confirm that substrate cleavage occurs in the context of infection and even if they do, that the proteolysis events are meaningful. While some of these studies were able to demonstrate that endogenous host protein cleavage events could happen within cells exposed to HIV protease, they were not able to provide evidence that these interactions were important for viral replication (Table 1.1).

The Krogan laboratory (UCSF) began an initiative to establish a global framework of HIV host-protein interactions using affinity purification followed by mass spectrometry (AP-MS) of each HIV protein [44]. The results of this study are detailed in Chapter 2, but in summary for HIV protease, a strep tag was added to the C-terminus of a catalytically inactive version of the enzyme both as a mature protein and as a component of the gag-pol polyprotein. These constructs were expressed within two human cell lines and lysates were then made from the cells. HIV protease was isolated from the lysates by immunoprecipitation against the strep tag and liquid chromatography tandem mass spectrometry (LC-MS/MS) was performed to identify any host proteins associated with the enzyme [44].

A surprising finding from this study was that, despite it's relatively small size (22 kDa as a dimer), HIV protease had the greatest number of host protein interactions of any HIV protein.

Intriguingly, after grouping the identified host proteins into complexes, it was revealed that 12 out of 13 members of the eukaryotic translation initiation factor 3 complex (eIF3) interact with HIV protease. Given the precedence for viral manipulation of host translation initiation pathways, the nature of the interactions between this complex and HIV protease was immediately of interest. Experiments using purified, human eIF3 complex and recombinant active HIV-1 protease, led to the discovery that one subunit, (eIF3d) is cleaved at a single location. Subsequent experiments demonstrated that the full length eIF3d protein is a negative regulator of HIV replication and likely inhibits reverse transcription. These results led to the hypothesis that upon entry of the virus into a host cell, proteolysis of eIF3d by HIV protease may facilitate reverse transcription of the HIV RNA genome (see Chapters 2 and 3).

Recently, additional support for this hypothesis has come from the discovery that the catalytic activity of HIV-1 protease is increased by an order of magnitude in the presence of RNA [45]. This result suggests that association between HIV-1 protease and the viral genome promotes activation of the enzyme within viral particles, perhaps to facilitate viral maturation. However, it could also indicate that HIV protease acts as a guardian of the viral genome as it enters a newly infected host cell, preventing host RNA binding proteins from interfering with the process of reverse transcription by cleaving the offending proteins.

Finally, support for a role of HIV protease in manipulation of host pathways has been further substantiated by the recent discovery that the serine/threonine receptor interacting protein kinases 1 and 2 (RIPK1 and RIPK2) are cleaved during HIV infection of CD4+ T cells (Table 1.1) [46]. These kinases are involved in immunity and regulation of cell death and are known to play an important role in regulating the cellular response to viral and bacterial pathogens. Cleavage of these proteins by HIV protease was shown to disrupt their functions and likely represents yet another mechanism facilitating HIV immune evasion. An important advancement of this publication was the conclusive detection of cleavage of the native proteins in cells infected by wild type HIV. Previous reports of host-protein cleavage events have had to

rely on recombinant, over-expression systems or infection with modified HIV to be able to observe cleavage of host proteins.

Our understanding of the role of HIV protease is continuing to expand, with the discovery of additional host proteins cleaved by this enzyme, improvements in techniques for identification of substrates and new approaches for determining the relevance of these interactions (see Chapters 3 and 7). While many questions remain, it is clear that this enigmatic enzyme does more than process viral proteins. Discovering these additional functions may provide further opportunities for therapeutic intervention. Resistance to HIV protease inhibitors and other antiviral drugs is a serious problem due to the high mutation rate associated with HIV reverse transcription. Uncovering the host proteins and pathways that are hijacked by HIV could allow for the development of drugs targeting these cellular proteins. Because eukaryotic proteins are subject to much lower rates of mutation, therapeutics affecting them would be less prone to the development of resistance mutations.

1.5 *Cryptococcus neoformans*, identifying virulence mechanisms of an opportunistic fungal pathogen

Concurrent with the rise of the HIV epidemic, a dramatic increase in the number of life-threatening infections caused by a previously obscure fungus, *Cryptococcus neoformans* occurred. A ubiquitous environmental fungus found in association with rotting wood, plants and bird droppings, most individuals are exposed to desiccated yeast cells or spores through inhalation at some point in their lives. When an immunocompetent individual encounters *C. neoformans,* a minor respiratory infection may develop but in most cases the infection is quickly cleared by the adaptive immune system. In immunocompromised individuals however, an initial lung infection spreads to the central nervous system, (CNS), leading to menigoencephalitis and ultimately death. Currently, *C. neoformans* infections are the single leading cause of HIV/AIDS associated mortality, resulting in 40% of AIDS deaths or approximately 600,000 deaths per year

[47,48]. The virulence of this pathogen and the difficulty in treating it is highlighted by the fact that over 60% of the one million new infections per year are lethal. The greatest infection burden occurs in sub-Saharan Africa, a region with high HIV infection rates and limited access to anti-retroviral and anti-fungal therapies. However, even infections that arise in individuals with full access to medical services are extremely difficult to treat and mortality rates are relatively high, with a 10-week survival rate of approximately 70%. A further complicating factor for treating *C. neoformans* infections is the risk of immune reconstitution inflammatory syndrome (IRIS) [49,50]. IRIS is the over activation of the immune response to the presence of *C. neoformans* in the body, which can result in a worsening of clinical symptoms. It occurs when an individual with HIV/AIDS infected by *C. neoformans* is treated with antiretroviral therapy, meaning that in some cases administration of antiretroviral drugs must be delayed until the fungal infection has been treated, a process that can take months and can decrease the individual's chance of recovery.

While the majority of *C. neoformans* infections occur in HIV/AIDS patients, additional at-risk populations include pregnant women, organ transplant recipients and people receiving high doses of corticosteroids for treatment of autoimmune conditions [51,52]. Infections in immunocompetent individuals are rare but those that do occur have a 35% risk of mortality. An additional concern is a species closely related to *C. neoformans*, *Cryptococcus gattii*, which primarily causes disease in immunocompetent individuals [53,54]. Outbreaks of this fungal pathogen have thus far been fairly limited in scope, but there are concerns that its prevalence could increase. Therapeutic intervention for *C. gattii* infections is also challenging and mortality rates are high. Given the close evolutionary relationship between *C. neoformans* and *C. gattii*, advancements in the treatment of *C. neoformans* infections will likely be applicable to *C. gattii* infections as well.

A major challenge in the treatment of fungal infections is the extremely limited number of anti-fungal drugs available for clinical use. Only three classes of compounds are approved for

fungal diseases, with just one of these compounds, (the echinocandins), having been developed in the past two decades [55,56]. Treatment of cryptococcal infections typically relies on three drugs, amphotericin B, flucytosine and flucanazole [57]. Amphotericin B and flucytosine are well known for their numerous and serious side effects, including a high risk of kidney damage and bone marrow toxicity. Additionally, to affectively treat systemic mycoses amphotericin B and sometimes flucytosine must be administered intravenously, reducing accessibility in resource poor settings. Flucanazole has far fewer side effects and can be administered orally; however its low efficacy requires extended periods of dosing often for 6-12 months, which can be difficult to achieve in regions with limited access to healthcare infrastructure.

The majority of fungal species are innocuous, posing no disease risk. Thus the characteristics that have allowed a limited number of species such as *C. neoformans* to become pathogenic are the focus of great research interest. One of the most notable virulence-associated traits of *C. neoformans* is the production of a large polysaccharide capsule surrounding the cell wall, a feature unique to this and closely related species. During infection the capsule is extensively remodeled and expanded in size, in some cases exceeding the width of the cell itself. It posses many immune-modulatory properties, such as prevention of phagocytosis and absorption of complement, factors which make it essential for virulence [58]. Other virulence traits identified in *C. neoformans* include secretion of multiple types of enzymes, including laccase which produces melanin, phospholipases and enzymes involved in resistance to stress such as superoxide dismutase [59,60]. In the context of infection, these enzymes are thought to facilitate resistance to killing through oxidative and nitrositive stresses and allow the fungal cells to obtain nutrients.

In addition to the enzymes mentioned above, *C. neoformans* is known to secrete peptidases [61]. Initially, protein hydrolysis assays revealed that clinical strains of *C. neoformans* or those isolated from infected animals tended to secrete more abundant proteolytic activity than environmental isolates, implying an association between protease secretion and

infection [62-64]. Further support for the role of secreted proteases in *C. neoformans* virulence came from analysis of gene expression. Microarray studies suggested that peptidases may be under regulation concurrent with other fungal virulence factors. For example, exposure to host-like conditions induced expression of verified virulence factors but also some predicted secreted peptidases [65]. To investigate these secreted enzymes, researchers attempted to isolate individual proteolytic activities from supernatants of *C. neoformans* cultures and analyze their biochemical characteristics. However, in most cases results of this analysis were limited to identification of the type of peptidase and its molecular weight [66,67].   Information such as the substrate specificity, regulation of activity and the gene encoding the isolated proteolytic enzyme remained unknown. Only a single example of detailed characterization of a secreted *C. neoformans* peptidase can be found in the literature currently. This study identified an enzyme that is important for entry of the fungus into the CNS, confirming the role of secreted peptidases in virulence [68].

In data presented in Chapter 4, we describe a substrate level analysis of secreted proteolytic activity produced by *C. neoformans*. In these experiments we determine the global specificity of secreted peptidases in this organism and through incorporation of proteomic and genetic techniques we identify and characterize several individual enzymes that contribute to this activity. We find that one of these enzymes, a previously uncharacterized aspartyl peptidase, is required for low pH tolerance and virulence. We then screen a targeted library of aspartyl inhibitor compounds and isolate several molecules that are nanomolar inhibitors. Furthermore, in Chapter 5 we show that a novel quorum sensing pathway in this organism contributes to regulation of secreted peptidases in *C. neoformans*.

Investigating the regulation and substrate specificity of peptidases produced by pathogens is critical to understanding host-pathogen interactions and their impact on disease. In the following chapters I present a study of HIV protease host-protein substrate interactions

(Chapters 2-3), discovery and detailed characterization of secreted peptidases produced by *C. neoformans* (Chapter 4-5) and biochemical and functional characterization of a secreted peptidase from by the opportunistic fungal pathogen *C. albicans* (Chapter 6). These data represent important advancements in our understanding the biology of each of these pathogens. In addition, the methods and approaches described are applicable to the discovery and study of peptidases from other pathogenic organisms.

1.6 References

1. Chowdhury M, Enenkel C (2015) Intracellular Dynamics of the Ubiquitin-Proteasome-System. F1000Res 4: 367.
2. Rawlings ND, Waller M, Barrett AJ, Bateman A (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 42: D503-509.
3. Afonina IS, Muller C, Martin SJ, Beyaert R (2015) Proteolytic Processing of Interleukin-1 Family Cytokines: Variations on a Common Theme. Immunity 42: 991-1004.
4. Ashkenazi A, Salvesen G (2014) Regulated cell death: signaling and mechanisms. Annu Rev Cell Dev Biol 30: 337-356.
5. Mohammed FF, Smookler DS, Taylor SE, Fingleton B, Kassiri Z, et al. (2004) Abnormal TNF activity in Timp3-/- mice leads to chronic hepatic inflammation and failure of liver regeneration. Nat Genet 36: 969-977.
6. Coussens LM, Tinkle CL, Hanahan D, Werb Z (2000) MMP-9 supplied by bone marrow-derived cells contributes to skin carcinogenesis. Cell 103: 481-490.
7. Balbin M, Fueyo A, Tester AM, Pendas AM, Pitiot AS, et al. (2003) Loss of collagenase-2 confers increased skin tumor susceptibility to male mice. Nat Genet 35: 252-257.
8. Kryza T, Silva Munasinghage L, Loessner D, Heuze-Vourc'h N, Clements JA (2015) The kallikrein-related peptidase family: Dysregulation and functions during cancer progression. Biochimie.
9. Reinhard SM, Razak K, Ethell IM (2015) A delicate balance: role of MMP-9 in brain development and pathophysiology of neurodevelopmental disorders. Front Cell Neurosci 9: 280.
10. Twigg MS, Brockbank S, Lowry P, FitzGerald SP, Taggart C, et al. (2015) The Role of Serine Proteases and Antiproteases in the Cystic Fibrosis Lung. Mediators Inflamm 2015: 293053.
11. Voskoboinik I, Whisstock JC, Trapani JA (2015) Perforin and granzymes: function, dysfunction and human pathology. Nat Rev Immunol 15: 388-400.
12. Puente XS, Sanchez LM, Gutierrez-Fernandez A, Velasco G, Lopez-Otin C (2005) A genomic view of the complexity of mammalian proteolytic systems. Biochem Soc Trans 33: 331-334.
13. Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. Protein Expr Purif 80: 283-293.
14. Harris JL, Backes BJ, Leonetti F, Mahrus S, Ellman JA, et al. (2000) Rapid and general profiling of protease specificity by using combinatorial fluorogenic substrate libraries. Proc Natl Acad Sci U S A 97: 7754-7759.
15. Boulware KT, Daugherty PS (2006) Protease specificity determination by using cellular libraries of peptide substrates (CLiPS). Proc Natl Acad Sci U S A 103: 7583-7588.

16. auf dem Keller U, Schilling O (2010) Proteomic techniques and activity-based probes for the system-wide study of proteolysis. Biochimie 92: 1705-1714.
17. Mahrus S, Trinidad JC, Barkan DT, Sali A, Burlingame AL, et al. (2008) Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. Cell 134: 866-876.
18. O'Donoghue AJ, Eroy-Reveles AA, Knudsen GM, Ingram J, Zhou M, et al. (2012) Global identification of peptidase specificity by multiplex substrate profiling. Nat Methods 9: 1095-1100.
19. O'Donoghue AJ, Knudsen GM, Beekman C, Perry JA, Johnson AD, et al. (2015) Destructin-1 is a collagen-degrading endopeptidase secreted by Pseudogymnoascus destructans, the causative agent of white-nose syndrome. Proc Natl Acad Sci U S A 112: 7478-7483.
20. Yang Y, Wen Y, Cai YN, Vallee I, Boireau P, et al. (2015) Serine proteases of parasitic helminths. Korean J Parasitol 53: 1-11.
21. Vaitkevicius K, Rompikuntal PK, Lindmark B, Vaitkevicius R, Song T, et al. (2008) The metalloprotease PrtV from Vibrio cholerae. FEBS J 275: 3167-3177.
22. Naglik JR, Challacombe SJ, Hube B (2003) Candida albicans secreted aspartyl proteinases in virulence and pathogenesis. Microbiol Mol Biol Rev 67: 400-428, table of contents.
23. Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, et al. (2012) Viral immune modulators perturb the human molecular network by common and unique strategies. Nature 487: 486-490.
24. Royall E, Doyle N, Abdul-Wahab A, Emmott E, Morley SJ, et al. (2015) Murine norovirus 1 (MNV1) replication induces translational control of the host by regulating eIF4E activity during infection. J Biol Chem 290: 4748-4758.
25. Lozano G, Martinez-Salas E (2015) Structural insights into viral IRES-dependent translation mechanisms. Curr Opin Virol 12: 113-120.
26. Gradi A, Foeger N, Strong R, Svitkin YV, Sonenberg N, et al. (2004) Cleavage of eukaryotic translation initiation factor 4GII within foot-and-mouth disease virus-infected cells: identification of the L-protease cleavage site in vitro. J Virol 78: 3271-3278.
27. Castello A, Alvarez E, Carrasco L (2011) The multifaceted poliovirus 2A protease: regulation of gene expression by picornavirus proteases. J Biomed Biotechnol 2011: 369648.
28. Willcocks MM, Carter MJ, Roberts LO (2004) Cleavage of eukaryotic initiation factor eIF4G and inhibition of host-cell protein synthesis during feline calicivirus infection. J Gen Virol 85: 1125-1130.
29. Kuyumcu-Martinez M, Belliot G, Sosnovtsev SV, Chang KO, Green KY, et al. (2004) Calicivirus 3C-like proteinase inhibits cellular translation by cleavage of poly(A)-binding protein. J Virol 78: 8172-8182.
30. Frankel AD, Young JA (1998) HIV-1: fifteen proteins and an RNA. Annu Rev Biochem 67: 1-25.
31. Ganser-Pornillos BK, Yeager M, Pornillos O (2012) Assembly and architecture of HIV. Adv Exp Med Biol 726: 441-465.
32. Pettit SC, Everitt LE, Choudhury S, Dunn BM, Kaplan AH (2004) Initial cleavage of the human immunodeficiency virus type 1 GagPol precursor by its activated protease occurs by an intramolecular mechanism. J Virol 78: 8477-8485.
33. Wensing AM, van Maarseveen NM, Nijhuis M (2010) Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance. Antiviral Res 85: 59-74.
34. Arts EJ, Hazuda DJ (2012) HIV-1 antiretroviral drug therapy. Cold Spring Harb Perspect Med 2: a007161.
35. Guatelli JC (2009) Interactions of viral protein U (Vpu) with cellular factors. Curr Top Microbiol Immunol 339: 27-45.
36. Grandgenett DP, Pandey KK, Bera S, Aihara H (2015) Multifunctional facets of retrovirus integrase. World J Biol Chem 6: 83-94.

37. Simon V, Bloch N, Landau NR (2015) Intrinsic host restrictions to HIV-1 and mechanisms of viral escape. Nat Immunol 16: 546-553.
38. Pawlak EN, Dikeakos JD (2015) HIV-1 Nef: a master manipulator of the membrane trafficking machinery mediating immune evasion. Biochim Biophys Acta 1850: 733-741.
39. Meng B, Lever AM (2013) Wrapping up the bad news: HIV assembly and release. Retrovirology 10: 5.
40. Malim MH, Bieniasz PD (2012) HIV Restriction Factors and Mechanisms of Evasion. Cold Spring Harb Perspect Med 2: a006940.
41. Shoeman RL, Honer B, Stoller TJ, Kesselmeier C, Miedel MC, et al. (1990) Human immunodeficiency virus type 1 protease cleaves the intermediate filament proteins vimentin, desmin, and glial fibrillary acidic protein. Proc Natl Acad Sci U S A 87: 6336-6340.
42. Shoeman RL, Sachse C, Honer B, Mothes E, Kaufmann M, et al. (1993) Cleavage of human and mouse cytoskeletal and sarcomeric proteins by human immunodeficiency virus type 1 protease. Actin, desmin, myosin, and tropomyosin. Am J Pathol 142: 221-230.
43. Shoeman RL, Kesselmier C, Mothes E, Honer B, Traub P (1991) Non-viral cellular substrates for human immunodeficiency virus type 1 protease. FEBS Lett 278: 199-203.
44. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, et al. (2012) Global landscape of HIV-human protein complexes. Nature 481: 365-370.
45. Potempa M, Nalivaika E, Ragland D, Lee SK, Schiffer CA, et al. (2015) A Direct Interaction with RNA Dramatically Enhances the Catalytic Activity of the HIV-1 Protease In Vitro. J Mol Biol 427: 2360-2378.
46. Wagner RN, Reed JC, Chanda SK (2015) HIV-1 protease cleaves the serine-threonine kinases RIPK1 and RIPK2. Retrovirology 12: 74.
47. Armstrong-James D, Meintjes G, Brown GD (2014) A neglected epidemic: fungal infections in HIV/AIDS. Trends Microbiol 22: 120-127.
48. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, et al. (2009) Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. AIDS 23: 525-530.
49. Antinori S (2013) New Insights into HIV/AIDS-Associated Cryptococcosis. ISRN AIDS 2013: 471363.
50. Jenny-Avital ER, Abadi M (2002) Immune reconstitution cryptococcosis after initiation of successful highly active antiretroviral therapy. Clin Infect Dis 35: e128-133.
51. La Hoz RM, Pappas PG (2013) Cryptococcal infections: changing epidemiology and implications for therapy. Drugs 73: 495-504.
52. Voelz K, May RC (2010) Cryptococcal interactions with the host immune system. Eukaryot Cell 9: 835-846.
53. Kwon-Chung KJ, Fraser JA, Doering TL, Wang Z, Janbon G, et al. (2014) Cryptococcus neoformans and Cryptococcus gattii, the etiologic agents of cryptococcosis. Cold Spring Harb Perspect Med 4: a019760.
54. McMullan BJ, Sorrell TC, Chen SC (2013) Cryptococcus gattii infections: contemporary aspects of epidemiology, clinical manifestations and management of infection. Future Microbiol 8: 1613-1631.
55. Krysan DJ (2015) Toward improved anti-cryptococcal drugs: Novel molecules and repurposed drugs. Fungal Genet Biol 78: 93-98.
56. Roemer T, Krysan DJ (2014) Antifungal drug development: challenges, unmet clinical needs, and new approaches. Cold Spring Harb Perspect Med 4.
57. Perfect JR, Dismukes WE, Dromer F, Goldman DL, Graybill JR, et al. (2010) Clinical practice guidelines for the management of cryptococcal disease: 2010 update by the infectious diseases society of america. Clin Infect Dis 50: 291-322.

58. O'Meara TR, Alspaugh JA (2012) The Cryptococcus neoformans capsule: a sword and a shield. Clin Microbiol Rev 25: 387-408.

59. Kronstad JW, Attarian R, Cadieux B, Choi J, D'Souza CA, et al. (2011) Expanding fungal pathogenesis: Cryptococcus breaks out of the opportunistic box. Nat Rev Microbiol 9: 193-203.

60. Coelho C, Bocca AL, Casadevall A (2014) The tools for virulence of Cryptococcus neoformans. Adv Appl Microbiol 87: 1-41.

61. Eigenheer RA, Jin Lee Y, Blumwald E, Phinney BS, Gelli A (2007) Extracellular glycosylphosphatidylinositol-anchored mannoproteins and proteases of Cryptococcus neoformans. FEMS Yeast Res 7: 499-510.

62. Ruma-Haynes P, Brownlee AG, Sorrell TC (2000) A rapid method for detecting extracellular proteinase activity in Cryptococcus neoformans and a survey of 63 isolates. J Med Microbiol 49: 733-737.

63. Chan MY, Tay ST (2010) Enzymatic characterisation of clinical isolates of Cryptococcus neoformans, Cryptococcus gattii and other environmental Cryptococcus spp. Mycoses 53: 26-31.

64. Vidotto V, Koga-Ito CY, Canella D, Sinicco A, Di Perri G, et al. (2000) Extracellular activity in Cryptococcus neoformans strains isolated from AIDS patients and from environmental sources. Rev Iberoam Micol 17: 14-19.

65. Haynes BC, Skowyra ML, Spencer SJ, Gish SR, Williams M, et al. (2011) Toward an integrated model of capsule regulation in Cryptococcus neoformans. PLoS Pathog 7: e1002411.

66. Yoo Ji J, Lee YS, Song CY, Kim BS (2004) Purification and characterization of a 43-kilodalton extracellular serine proteinase from Cryptococcus neoformans. J Clin Microbiol 42: 722-726.

67. Pinti M, Orsi CF, Gibellini L, Esposito R, Cossarizza A, et al. (2007) Identification and characterization of an aspartyl protease from Cryptococcus neoformans. FEBS Lett 581: 3882-3886.

68. Vu K, Tham R, Uhrig JP, Thompson GR, 3rd, Na Pombejra S, et al. (2014) Invasion of the central nervous system by Cryptococcus neoformans requires a secreted fungal metalloprotease. MBio 5: e01101-01114.

Table 1.1	
Protein name	**Citation**
Actin	[1,2]
Alzheimer amyloid precursor protein (AAP)	[2]
Bcl-2	[3]
Calmodulin	[4]
Caspase-8	[5]
Complement factor C3	[6]
DDX5	[7]
Desmin	[8]
eIF3d	[9]
eIF4G	[10]
Fibronectin	[11]
Filamin	[12]
Fimbrin	[12]
Focal adhesion plaque kinase	[12]
GCN2	[13]
Glia fibrillary acidic protein	[8]
Integrin α3	[12]
Integrin β4	[12]
Lyric	[14]
Microtubule-associated protein MAP1	[15]
Microtubule-associated protein MAP2	[15]
Myosin	[1]
NDR1	[16]
NDR2	[16]
NF-kB	[17]
PABP (poly(A)-binding protein)	[10]
Phosphorylase kinase	[4]
Pro-interleukin 1β	[2]
RIPK1	[18]
RIPK2	[18]
Spectrin	[12]
Talin	[12]
Tropomyosin	[19]
Troponin C	[2]
Vimentin	[8]
α-actinin	[20]

Table 1.1 References

1. Shoeman RL, Sachse C, Honer B, Mothes E, Kaufmann M, et al. (1993) Cleavage of human and mouse cytoskeletal and sarcomeric proteins by human immunodeficiency virus type 1 protease. Actin, desmin, myosin, and tropomyosin. Am J Pathol 142: 221-230.

2. Tomasselli AG, Hui JO, Adams L, Chosay J, Lowery D, et al. (1991) Actin, troponin C, Alzheimer amyloid precursor protein and pro-interleukin 1 beta as substrates of the protease from human immunodeficiency virus. J Biol Chem 266: 14548-14553.

3. Strack PR, Frey MW, Rizzo CJ, Cordova B, George HJ, et al. (1996) Apoptosis mediated by HIV protease is preceded by cleavage of Bcl-2. Proc Natl Acad Sci U S A 93: 9571-9576.

4. Daube H, Billich A, Mann K, Schramm HJ (1991) Cleavage of phosphorylase kinase and calcium-free calmodulin by HIV-1 protease. Biochem Biophys Res Commun 178: 892-898.

5. Nie Z, Phenix BN, Lum JJ, Alam A, Lynch DH, et al. (2002) HIV-1 protease processes procaspase 8 to cause mitochondrial release of cytochrome c, caspase cleavage and nuclear fragmentation. Cell Death Differ 9: 1172-1184.

6. Kisselev AF, Mentele R, von der Helm K (1997) Cleavage of the complement system C3 component by HIV-1 proteinase. Biol Chem 378: 439-442.

7. Impens F, Timmerman E, Staes A, Moens K, Arien KK, et al. (2012) A catalogue of putative HIV-1 protease host cell substrates. Biol Chem 393: 915-931.

8. Shoeman RL, Honer B, Stoller TJ, Kesselmeier C, Miedel MC, et al. (1990) Human immunodeficiency virus type 1 protease cleaves the intermediate filament proteins vimentin, desmin, and glial fibrillary acidic protein. Proc Natl Acad Sci U S A 87: 6336-6340.

9. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, et al. (2012) Global landscape of HIV-human protein complexes. Nature 481: 365-370.

10. Castello A, Franco D, Moral-Lopez P, Berlanga JJ, Alvarez E, et al. (2009) HIV- 1 protease inhibits Cap- and poly(A)-dependent translation upon eIF4GI and PABP cleavage. PLoS One 4: e7997.

11. Oswald M, von der Helm K (1991) Fibronectin is a non-viral substrate for the HIV proteinase. FEBS Lett 292: 298-300.

12. Shoeman RL, Hartig R, Hauses C, Traub P (2002) Organization of focal adhesion plaques is disrupted by action of the HIV-1 protease. Cell Biol Int 26: 529-539.

13. del Pino J, Jimenez JL, Ventoso I, Castello A, Munoz-Fernandez MA, et al. (2012) GCN2 has inhibitory effect on human immunodeficiency virus-1 protein synthesis and is cleaved upon viral infection. PLoS One 7: e47272.

14. Engeland CE, Oberwinkler H, Schumann M, Krause E, Muller GA, et al. (2011) The cellular protein lyric interacts with HIV-1 Gag. J Virol 85: 13322-13332.

15. Wallin M, Deinum J, Goobar L, Danielson UH (1990) Proteolytic cleavage of microtubule-associated proteins by retroviral proteinases. J Gen Virol 71 (Pt 9): 1985-1991.

16. Devroe E, Silver PA, Engelman A (2005) HIV-1 incorporates and proteolytically processes human NDR1 and NDR2 serine-threonine kinases. Virology 331: 181-189.

17. Riviere Y, Blank V, Kourilsky P, Israel A (1991) Processing of the precursor of NF-kappa B by the HIV-1 protease during acute infection. Nature 350: 625-626.

18. Wagner RN, Reed JC, Chanda SK (2015) HIV-1 protease cleaves the serine-threonine kinases RIPK1 and RIPK2. Retrovirology 12: 74.

19. Honer B, Shoeman RL, Traub P (1992) Degradation of cytoskeletal proteins by the human immunodeficiency virus type 1 protease. Cell Biol Int Rep 16: 603-612.

20. Shoeman RL, Kesselmier C, Mothes E, Honer B, Traub P (1991) Non-viral cellular substrates for human immunodeficiency virus type 1 protease. FEBS Lett 278: 199-203.

Chapter 2: Global landscape of HIV-human protein interactions

2.1 Forward

At the time of this publication, the concept that HIV protease could serve a dual role in viral protein processing and host-pathway modification was not accepted by most of the virology field. Previously, researchers had demonstrated that HIV protease was capable of cleaving a range of host proteins but most of these experiments were missing critical information such as the identification of the cleavage site within the protein and plausible data showing that these cleavage events impacted HIV virulence. In addition, large screens of HIV host-protein interactions were viewed with some skepticism, as there was very little overlap between screens published by different researchers. It was unclear how many interactions discovered in this way were impactful and how many were artifacts of the experimental system.

My contribution to this paper was the discovery that HIV protease cleaves a single subunit of the eukaryotic translation initiation factor 3 complex, (subunit d), which had been identified in this study as a HIV protease host-protein interaction. I identified the cleavage site through N-terminal sequencing and conclusively demonstrated location of the scissile bond by mutating the P1 methionine to a valine. I then determined that HIV protease cleaves eIF3d within its RNA binding domain and that upon cleavage, the N-terminal fragment of this protein remains associated with the eIF3 complex while the C-terminal fragment dissociates. Using the purified, recombinant HIV protease that I had expressed, I determined through *in vitro* digestion assays and Western blot analysis that eIF3d was proteolyzed at a rate that is on par with the native HIV protease substrate, gag. Furthermore, I found that eIF3d is cut more efficiently than some previously published host protein substrates. In addition to discovering and characterizing an important host-protein interaction of HIV protease, these results contributed to validating the

overall findings of this study and the hundreds of other HIV-host interactions reported within it.

2.2 Abstract

Human immunodeficiency virus (HIV) has a small genome and therefore relies heavily on the host cellular machinery to replicate. Identifying which host proteins and complexes come into physical contact with the viral proteins is crucial for a comprehensive understanding of how HIV rewires the host's cellular machinery during the course of infection. Here we report the use of affinity tagging and purification mass spectrometry[1-3] to determine systematically the physical interactions of all 18 HIV-1 proteins and polyproteins with host proteins in two different human cell lines (HEK293 and Jurkat). Using a quantitative scoring system that we call MiST, we identified with high confidence 497 HIV–human protein–protein interactions involving 435 individual human proteins, with ~40% of the interactions being identified in both cell types. We found that the host proteins hijacked by HIV, especially those found interacting in both cell types, are highly conserved across primates. We uncovered a number of host complexes targeted by viral proteins, including the finding that HIV protease cleaves eIF3d, a subunit of eukaryotic translation initiation factor 3. This host protein is one of eleven identified in this analysis that act to inhibit HIV replication. This data set facilitates a more comprehensive and detailed understanding of how the host machinery is manipulated during the course of HIV infection.

2.3 Introduction

A map of the physical interactions between proteins within a particular system is necessary for studying the molecular mechanisms that underlie the system. The analysis of protein–protein interactions (PPIs) has been successfully accomplished in different organisms using a variety of technologies, including mass spectrometry approaches[1,3,4] and those designed to detect pairwise physical interactions, including the two-hybrid yeast system[5,6] and

protein-fragment complementation assays[7]. Although two-hybrid methodologies have been used to systematically study host–pathogen interactions[8,9], so far no systematic affinity tagging/purification mass spectrometry (AP–MS) study has been carried out on any host–pathogen system. Here we have targeted HIV-1 for such an analysis, uncovering a wide variety of host proteins, complexes and pathways that are hijacked by the virus during the course of infection.

2.4 Methods

Briefly, affinity tagging and purification was carried out as previously described[2] and the protein samples were analysed on a Thermo Scientific LTQ Orbitrap XL mass spectrometer. For the evolutionary analysis, genome-wide alignments to rhesus macaque were downloaded from the University of California, Santa Cruz genome browser (http://genome.ucsc.edu/) and evolutionary rates for each group of genes considered were measured using the synonymous and non-synonymous rates of evolution. For the *in vitro* protease assay, maltose binding protein (MBP)-tagged PR was expressed in BL21 (Gold) DE3 cells in the presence of 100 μM Saquinavir and purified on an MBP trap column. Purified eIF3 was obtained from J. Cate (University of California, Berkeley). For the infection assays, HeLa P4.R5 cells were transfected with short interfering RNAs and after 48 h infected with pNL4-3 or a pNL4-3-derived VSV-G-pseudotyped reporter virus. Infection levels were determined by luminescence read-out.

2.5 Results

We aimed to identify host proteins associated with HIV-1 proteins systematically and quantitatively using an AP–MS approach[2,3]. To this end, we cloned the genes corresponding to all 18 HIV-1 proteins and polyproteins, including the accessory factors (Vif, Vpu, Vpr and Nef), Tat, Rev, the polyproteins (Gag, Pol and Gp160) and the corresponding processed products

(MA, CA, NC and p6; PR, RT and IN; and Gp120 and Gp41, respectively) (Supplementary Fig. 1 and Supplementary Table 1). Each clone was fused to a purification tag (consisting of 2×Strep and 3×Flag) and transiently transfected into HEK293 cells; each also was used to generate stably expressed, tetracycline-inducible, affinity-tagged versions of the proteins in Jurkat cells (Fig. 2.1a and Supplementary Fig. 2). Following multiple purifications of each factor from both cell lines, the material on the anti-FLAG or Strep-Tactin beads, as well as the eluted material, was analysed by mass spectrometry (Fig. 2.1a and Supplementary Table 2). Finally, an aliquot of each purified factor was subjected to SDS–polyacrylamide gel electrophoresis, stained (Supplementary Fig. 2.3) and subjected to analysis by mass spectrometry. For each HIV factor, we identified co-purifying host proteins that were reproducible regardless of the protocol used (Supplementary Figs 4, 5 and 7 and Supplementary Data 1). Several scoring systems can quantify PPIs from AP–MS proteomic data sets, including PE[10], CompPASS[4] and SAINT[11]. For this data set, we devised a scoring system particularly suited for identifying AP–MS-derived host–pathogen PPIs, which we call MiST (mass spectrometry interaction statistics). The MiST score is a weighted sum of three measures: protein abundance measured by peak intensities from the mass spectrum (abundance); invariability of abundance over replicated experiments (reproducibility); and uniqueness of an observed host–pathogen interaction across all viral purifications (specificity) (Fig. 2.1b and Supplementary Methods). These three metrics are summed by principal component analysis into a composite score (Fig. 2.1c and Supplementary Data 2). By comparing our dataset with a benchmark of well-characterized HIV–human PPIs (Supplementary Table 3), analysis of the MiST scoring system revealed superior performance on our data set when compared to CompPASS or SAINT (Supplementary Fig. 6) (and comparable performance using other data sets (Supplementary Fig. 8)) and allowed us to define a MiST cut-off of 0.75, corresponding to ~4% of all detected interactions. To estimate how many interactions would exceed this threshold by chance, we randomly shuffled our data set 1,000 times. A random MiST score of 0.75 or greater was assigned to an interaction ten times less

frequently than we saw among the MiST scores for the real data, and the probability of an interaction assignment with a random MiST score greater than 0.75 was 2.5×10^{-4} (Fig. 2.1d).

At the MiST threshold of 0.75, the number of host proteins we found associated with each HIV protein ranged from 0 (CA and p6) to 63 (Gp160) (Fig. 2.1e). In total, we observed 497 different HIV–human PPIs (347 and 348 identified from HEK293 cells and Jurkat cells, respectively) (Supplementary Data 3). We detected 196 interactions (~40%) in both cell types; 150 and 151 were specific to the HEK293 cells and the Jurkat cells, respectively (Fig. 2.1e). Only some of these specificities could be explained by differential gene expression in the two cell lines (Supplementary Fig. 9). Using antibodies against 26 of the human proteins, and affinity-tagged versions of an additional 101, we could confirm 97 of the 127 AP–MS derived HIV–human PPIs using co-immunoprecipitation/western blot analysis (76% success rate) (Supplementary Figs 10 and 11), suggesting that we derived a high-quality physical interaction data set.

We next analysed the functional categories of host proteins associated with each HIV protein, and in doing so uncovered many expected connections. These included an enrichment of host factors involved in transcription physically linked to the HIV transcription factor Tat and an enrichment of host machinery implicated in the regulation of ubiquitination associating with Vpu, Vpr and Vif, HIV accessory factors that hijack ubiquitin ligases[12] (Fig. 2.1f and Supplementary Data 4). When we considered domain types instead of whole proteins (Fig. 2.1g and Supplementary Table 4), we found that host proteins interacting with IN are enriched in 14-3-3 domains, which generally bind phosphorylated regions of proteins[13], and that proteins containing β-propellers have a higher propensity for binding to Vpr (for additional domain enrichment analysis, see Supplementary Fig. 12). These domain analyses could facilitate future structural modelling of HIV–human PPIs.

Next we compared our data to other HIV-related data sets, including previously published HIV–human PPIs and host factors implicated in HIV function from genome-wide RNA interference (RNAi) screens. For example, the VirusMint database[14] contains 587 HIV–human literature-curated PPIs (Supplementary Data 5), which are mostly derived from small-scale, targeted studies. Although the overlap between the 497 interactions identified in this work and those in VirusMint is statistically significant ($P = 8 \times 10^{-8}$), it corresponds to only 19 PPIs (Fig. 2.2a and Supplementary Table 5). However, a greater overlap exists, one that remains statistically significant, when interactions below the MiST threshold of 0.75 are considered using a sliding cut-off (for example, at a MiST score of 0.2 there exists an overlap of 67 PPIs ($P = 1 \times 10^{-3}$); Fig. 2.2c, red lines, and Supplementary Data 6). This overlap indicates that we have indeed identified many interactions that have been previously reported. However, it is likely that the higher scoring interactions identified here have a greater chance of being biologically relevant with respect to HIV function than do many of those in VirusMint.

Recently, four RNAi screens identified host factors that have an adverse effect on HIV-1 replication when knocked down[15-18]. In total, 1,071 human genes were identified in these four studies (Supplementary Data 7), 55 of which overlap with the 435 proteins ($P = 2.7 \times 10^{-10}$; Fig. 2.2b, Supplementary Fig. 12 and Supplementary Table 6). Again, this overlap increases (as does its statistical significance) if we consider proteins participating in HIV–human PPIs with MiST scores below 0.75 (Fig. 2.2c, blue lines, and Supplementary Data 8).

To identify the evolutionary forces operating on host proteins interacting with HIV-1, we performed a comparative genomics analysis of divergence patterns between human and rhesus macaque. The proteins identified in both HEK293 and Jurkat cell lines had stronger signatures of evolutionary constraint than those identified exclusively in one cell line or in VirusMint (Fig. 2.2d). Points in the lower-right quadrant of Fig. 2.2d show signatures of strong purifying selection, whereas the upper-right quadrant shows signatures more consistent with neutral

evolution. This observation suggests that the PPIs identified in our study, especially the ones identified in both cell types, are more physiologically relevant to mammalian evolution than those reported in VirusMint.

We next plotted the 497 HIV–human interactions identified in this study in a network representation (Fig. 2.3) containing nodes corresponding to 16 HIV (yellow) and 435 human factors that were derived from the HEK293 cells (blue), Jurkat cells (red) or both. We also introduced 289 interactions between human proteins (black edges) derived from several databases[19,20] (Supplementary Data 9). These human–human interactions helped to identify many host complexes, including several that have been previously characterized (see Supplementary Information for a detailed discussion of the HIV–human interaction data sets). Ultimately, all data will be accessible for searching and comparison to other HIV-related data sets using the web-based software GPS-PROT[21] (http://www.gpsprot.org/).

Notably, we found that Pol and PR, which we needed to make catalytically inactive (Supplementary Fig. 1), bound the translational initiation complex eIF3, a 13-subunit complex (eIF3a to eIF3m). We detected 12 of the subunits bound to Pol and/or PR, except eIF3j, which is only loosely associated with the complex[22] (Fig. 2.4a). Even though PR is the smallest of the *pol*-encoded proteins, we find it associated with the greatest number of host factors (Fig. 2.4a). To determine whether components of the translation complex are substrates for PR, FLAG-tagged versions of ten eIF3 subunits were individually co-transfected, each with a small amount of active HIV-1 PR, into HEK293 cells. The cell lysates were analysed by western blotting and only eIF3d was found to be cleaved (Fig. 2.4b). Purification of tagged versions of the amino and carboxy termini of cleaved eIF3d revealed that only the N terminus of 114 amino-acid residues associates with the eIF3 complex (Supplementary Table 7). The cleavage occurred with an efficiency similar to that of the processing of the natural PR substrate Gag (Fig. 2.4c), whereas two cellular proteins previously described to be cleaved by HIV PR,

PAPBC1[23] and BCL2[24], were cleaved only at higher PR concentrations or not at all, respectively. To confirm this result *in vitro*, we incubated purified human eIF3 with active PR, resulting in the removal of a 70-kDa band and the appearance of a ~60-kDa protein product (Fig. 2.4d). Analysis of the cleaved product by N-terminal sequencing revealed a cleavage of eIF3d between Met 114 and Leu 115, which corresponds to the consensus sequence for HIV-1 protease[25] and falls within the RNA-binding domain (RRM) of eIF3d (ref. [26]; Fig. 2.4d).

Next we used four to six short interfering RNAs against different eIF3 subunits in HIV infectivity assays (Fig. 2.4e, f, Supplementary Fig. 14 and Supplementary Table 8). Using a fusion of HIV with vesicular stomatitis virus glycoprotein (VSV-G), which only allows for a single round of replication, knockdown of eIF3d, but not other eIF3 subunits, resulted in an increase in infectivity (Fig. 2.4e), suggesting that this factor acts in early stages of infection. In assays requiring multiple rounds of HIV infection, knockdown of eIF3d, eIF3e and eIF3f enhanced HIV NL4.3 infectivity by a factor of three to five, whereas inhibition of eIF3c, eIF3g and eIF3i had no promoting effect (Fig. 2.4f). Consistent with these results, a previous overexpression screen for factors that restrict HIV-1 replication identified eIF3f as the most potent inhibitory clone[27]. Furthermore, using assays monitoring both early and late products we found that knockdown of eIF3d results in an increase in accumulation of reverse transcription product (Fig. 2.4g and Supplementary Fig. 15). This suggests that eIF3 does in fact have a role in the early stages of infection, perhaps by binding to the viral RNA through the RNA-binding domain in eIF3d, and thus inhibiting RT, an effect that is overcome by PR cleavage of eIF3d (Supplementary Fig. 16). These results suggest that our data set will be enriched not only for host proteins the virus requires for efficient replication (Fig. 2.2b, c), but also those that have an inhibitory role during infection. Indeed, we have found that an additional ten factors from our list of inter-actors, when knocked down by RNAi, produce an increase in HIV infection (Supplementary Figs 17–19, Supplementary Tables 12 and 13 and Supplementary Methods). Knockdown of two of these,

DESP and HEAT1, also resulted in an increase in HIV integration (Supplementary Fig. 20 and Supplementary Table 14), consistent with their physical association with IN.

2.6 Discussion

As well as performing the systematic AP–MS study reported here, we explored in further detail the biological significance of two newly identified HIV–human interactions: HIV protease targeting a component of eIF3 that is inhibitory to HIV replication; and CBF-β, a new component of the Vif–CUL5 ubiquitin ligase complex required for APOBEC3G stability and HIV infectivity[28]. Further work will be required to determine whether, how and at what stage of infection the remaining host factors impinge on HIV function. Ultimately, our analysis of the host factors co-opted by different viruses using the same proteomic pipeline will allow for the identification of protein complexes routinely targeted by different pathogens, which may represent better therapeutic targets for future studies.

2.7 References

1. Gavin AC, et al. Proteome survey reveals modularity of the yeast cell machinery. Nature. 2006; 440:631–636. [PubMed: 16429126]

2. Jäger S, et al. Purification and characterization of HIV-human protein complexes. Methods. 2011; 53:13–19. [PubMed: 20708689]

3. Krogan NJ, et al. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. Nature. 2006; 440:637–643. [PubMed: 16554755]

4. Sowa ME, Bennett EJ, Gygi SP, Harper JW. Defining the human deubiquitinating enzyme interaction landscape. Cell. 2009; 138:389–403. [PubMed: 19615732]

5. Yu H, et al. High-quality binary protein interaction map of the yeast interactome network. Science. 2008; 322:104–110. [PubMed: 18719252]

6. Stelzl U, et al. A human protein-protein interaction network: a resource for annotating the proteome. Cell. 2005; 122:957–968. [PubMed: 16169070]

7. Tarassov K, et al. An *in vivo* map of the yeast protein interactome. Science. 2008; 320:1465–1470. [PubMed: 18467557]

8. Calderwood MA, et al. Epstein-Barr virus and virus human protein interaction maps. Proc Natl Acad Sci USA. 2007; 104:7606–7611. [PubMed: 17446270]

9. Shapira SD, et al. A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. Cell. 2009; 139:1255–1267. [PubMed: 20064372]

10. Collins SR, et al. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. Mol Cell Proteomics. 2007; 6:439–450. [PubMed: 17200106]

11. Choi H, et al. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. Nature Methods. 2011; 8:70–73. [PubMed: 21131968]

12. Malim MH, Emerman M. HIV-1 accessory proteins–ensuring viral survival in a hostile environment. Cell Host Microbe. 2008; 3:388–398. [PubMed: 18541215]

13. Yaffe MB, et al. The structural basis for 14-3-3:phosphopeptide binding specificity. Cell. 1997; 91:961–971. [PubMed: 9428519]

14. Chatrar-yamontri A, et al. VirusMINT: a viral protein interaction database. Nucleic Acids Res. 2009; 37:D669–D673. [PubMed: 18974184]

15. Brass AL, et al. Identification of host proteins required for HIV infection through a functional genomic screen. Science. 2008; 319:921–926. [PubMed: 18187620]

16. König R, et al. Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. Cell. 2008; 135:49–60. [PubMed: 18854154]

17. Yeung ML, Houzet L, Yedavalli VS, Jeang KT. A genome-wide short hairpin RNA screening of Jurkat T-cells for human proteins contributing to productive HIV-1 replication. J Biol Chem. 2009; 284:19463–19473. [PubMed: 19460752]

18. Zhou H, et al. Genome-scale RNAi screen for host factors required for HIV replication. Cell Host Microbe. 2008; 4:495–504. [PubMed: 18976975]

19. Ruepp A, et al. CORUM: the comprehensive resource of mammalian protein complexes–2009. Nucleic Acids Res. 2010; 38:D497–D501. [PubMed: 19884131]

20. Stark C, et al. BioGRID: a general repository for interaction datasets. Nucleic Acids Res. 2006; 34:D535–D539. [PubMed: 16381927]

21. Fahey ME, et al. GPS-Prot: a web-based visualization platform for integrating host-pathogen interaction data. BMC Bioinformatics. 2011; 12:298. [PubMed: 21777475]

22. Hinnebusch AG. eIF3: a versatile scaffold for translation initiation complexes. Trends Biochem Sci. 2006; 31:553–562. [PubMed: 16920360]

23. Alvarez E, Castello A, Menendez-Arias L, Carrasco L. HIV protease cleaves poly(A)-binding protein. Biochem J. 2006; 396:219–226. [PubMed: 16594896]

24. Strack PR, et al. Apoptosis mediated by HIV protease is preceded by cleavage of Bcl-2. Proc Natl Acad Sci USA. 1996; 93:9571–9576. [PubMed: 8790371]

25. Schilling O, Overall CM. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. Nature Biotechnol. 2008; 26:685–694. [PubMed: 18500335]

26. Asano K, et al. Structure of cDNAs encoding human eukaryotic initiation factor 3 subunits. Possible roles in RNA binding and macromolecular assembly. J Biol Chem. 1997; 272:27042–27052. [PubMed: 9341143]

27. Valente ST, Gilmartin GM, Mott C, Falkard B, Goff SP. Inhibition of HIV-1 replication by eIF3f. Proc Natl Acad Sci USA. 2009; 106:4071–4078. [PubMed: 19237569]

28. Jäger, S., et al. Vif hijacks CBF-β to degrade APOBEC3G and promote HIV-1 infection. Nature. http://dx.doi.org/10.1038/nature10693(this issue)

29. Zhou M, et al. Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. Proc Natl Acad Sci USA. 2008; 105:18139–18144. [PubMed: 18599441]

Figure 2.1. Affinity purification of HIV-1 proteins, analysis and scoring of mass spectrometry data a, Flowchart of the proteomic AP–MS used to define the HIV–host interactome. PAGE, polyacrylamide gel electrophoresis. SF, 2×Strep–3×Flag affinity tag. b, Data from AP–MS experiments are organized in an interaction table with cells representing amount of prey protein purified (for example spectral counts or peptide intensities). Three features are used to describe bait–prey relationships: abundance (blue), reproducibility (the invariability of bait–prey pair quantities; red) and specificity (green). c, All bait–prey pairs are mapped into the three-feature space (abundance, reproducibility and specificity). The MiST score is defined as a projection on the first principal component (red line). All interactions, represented as nodes, above the defined threshold (0.75) are shown in red. This procedure separates the interactions more likely to be biologically relevant (for example Vif–ELOC (ELOC also known as TCEB1), Vpr–VPRBP and Tat–CCNT1) from the interactions that are likely to be less relevant owing to low reproducibility (Vpu–ATP4A) or specificity (RT–HSP71 (HSP71 also known as HSPA1A) and NC–RL23A (RL23A also known as RPL23A)). d, The histogram of MiST scores (real data) is compared with a randomized set of scores obtained from randomly shuffling the bait–prey table (simulated data). The MiST score threshold (0.75) was defined using a benchmark (Supplementary Table 3) whereby the predictions are enriched for these interactions by a factor of at least ten relative to random predictions (as well as through ROC (receiver operating characteristic) and recall plots (Supplementary Fig. 6)). e, Bar graph of the number of host proteins we found interacting with each HIV factor (MiST score, >0.75). The cell type in which the interaction was found is represented in blue (HEK293 only), yellow (Jurkat only) or red (both). f, g, Heat maps representing enriched biological functions (f) and domains (g) from the human proteins identified as interacting with HIV proteins (Supplementary Methods). ER, endoplasmic reticulum; mRNA, messenger RNA; tRNA, transfer RNA. TPR, tetratricopeptide repeat; HTH, helix-turn-helix; SPFH, stomatin–prohibitin–flotillin–HflK/C.
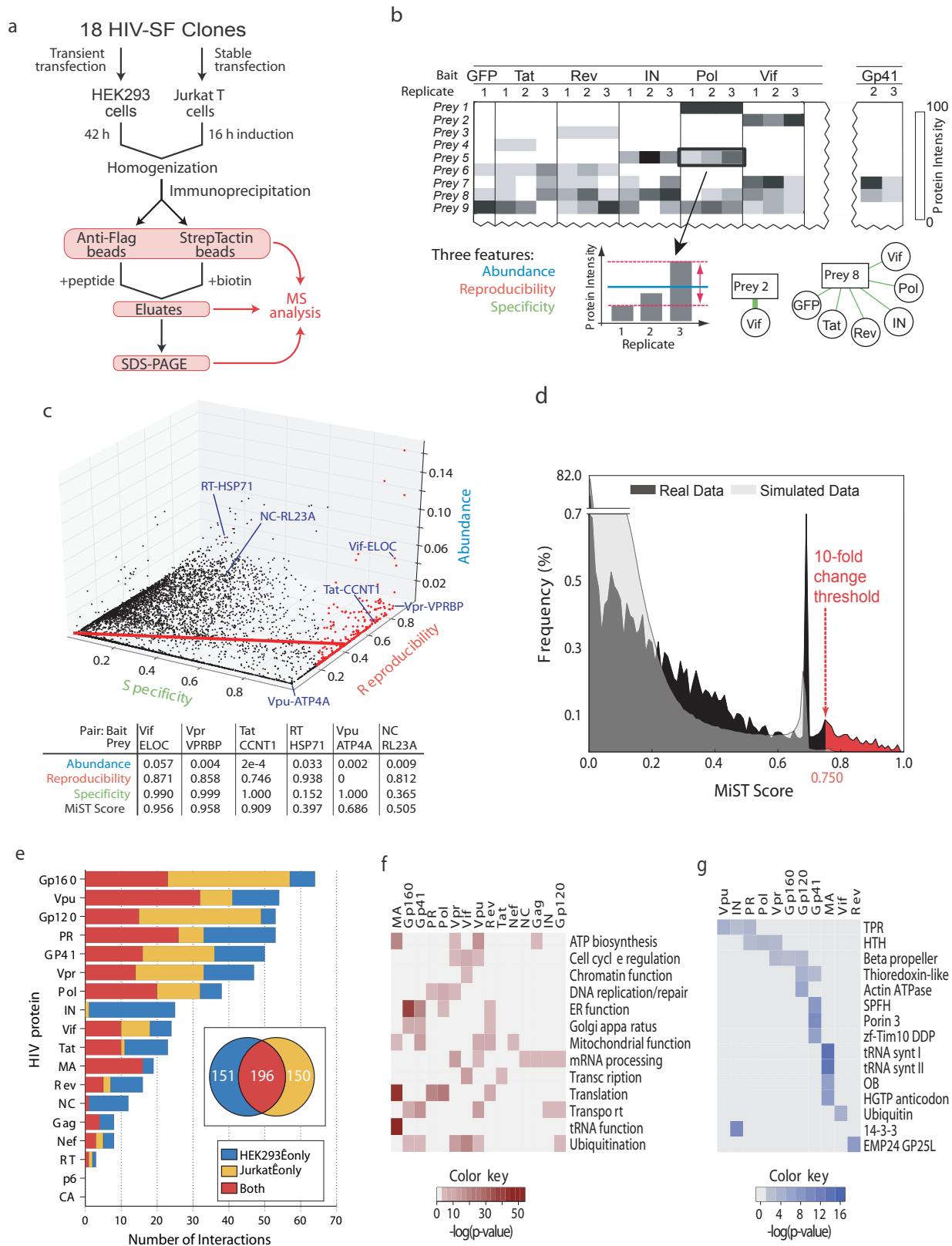
Figure 2.1

Figure 2.2. Comparison of PPI data with other HIV data sets a, Overlap of the 497 HIV–human PPIs with the 587 PPIs reported in VirusMint (Supplementary Table 5). b, Overlap of the 435 human proteins with the genes identified in four HIV-dependency RNAi screens15-18 (Supplementary Table 6). c, Number of interactions overlapping with VirusMint (solid red line) and proteins with RNAi screens (solid blue line) as functions of the MiST cut-off. The P values of the overlap are represented as dashed lines using the same colours (Supplementary Data 6 and 8). d, Comparative genomics analysis of divergence patterns between human and rhesus macaque reveals strong evolutionary constraints on human proteins binding to HIV proteins. The x and y axes represent P values for the synonymous (dS) and non-synonymous (dN) rates of evolution (Supplementary Methods). Horizontal and vertical dotted lines are drawn at 0.5% to indicate the Bonferroni significance threshold for each axis. For the VirusMint data, the significance of ω (dN/dS) is primarily driven by higher rates of synonymous evolution. U, union; ∩, intersection; Pω, bootstrap-based P value for ω.

a Protein - Protein Interactions

this study
n = 497

478 | 19* | 568

VirusMint
n = 587

Jurkat   19
HEK293   13

* p = 8e −08

b Protein Factors Identified

this study
n = 435

380 | 55* | 1016

RNAi Screens
n = 1071

Jurkat   41
HEK293   44

reference	n	p-value
Brass et al.	17	5e −05
Zhou et al.	11	3e −02
Konig et al.	17	2e −04
Yeung et al.	12	6e −03

* p = 2.7e −010

c

(LH scale)
RNAi overlap
VirusMint overlap

(RH scale)
RNAi p-values
VirusMint p-values

d

293 ∪ Jurkat
293 ∩ Jurkat
293 specific
all 293
Jurkat specific
all Jurkat
VirusMint

$p_\omega < 1$
$p_\omega < 0.01$
$p_\omega < 0.001$
$p_\omega < 1e-04$

Figure 2.3. Network representation of the HIV–human PPIs In total, 497 HIV–human interactions (blue) are represented between 16 HIV proteins and 435 human factors. Each node representing a human protein is split into two colours and the intensity of each colour corresponds to the MiST score from interactions derived from HEK293 (blue) or Jurkat (red) cells. Black edges correspond to interactions between host factors (289) that were obtained from publicly available databases; dashed edges correspond to interactions also found in VirusMint.
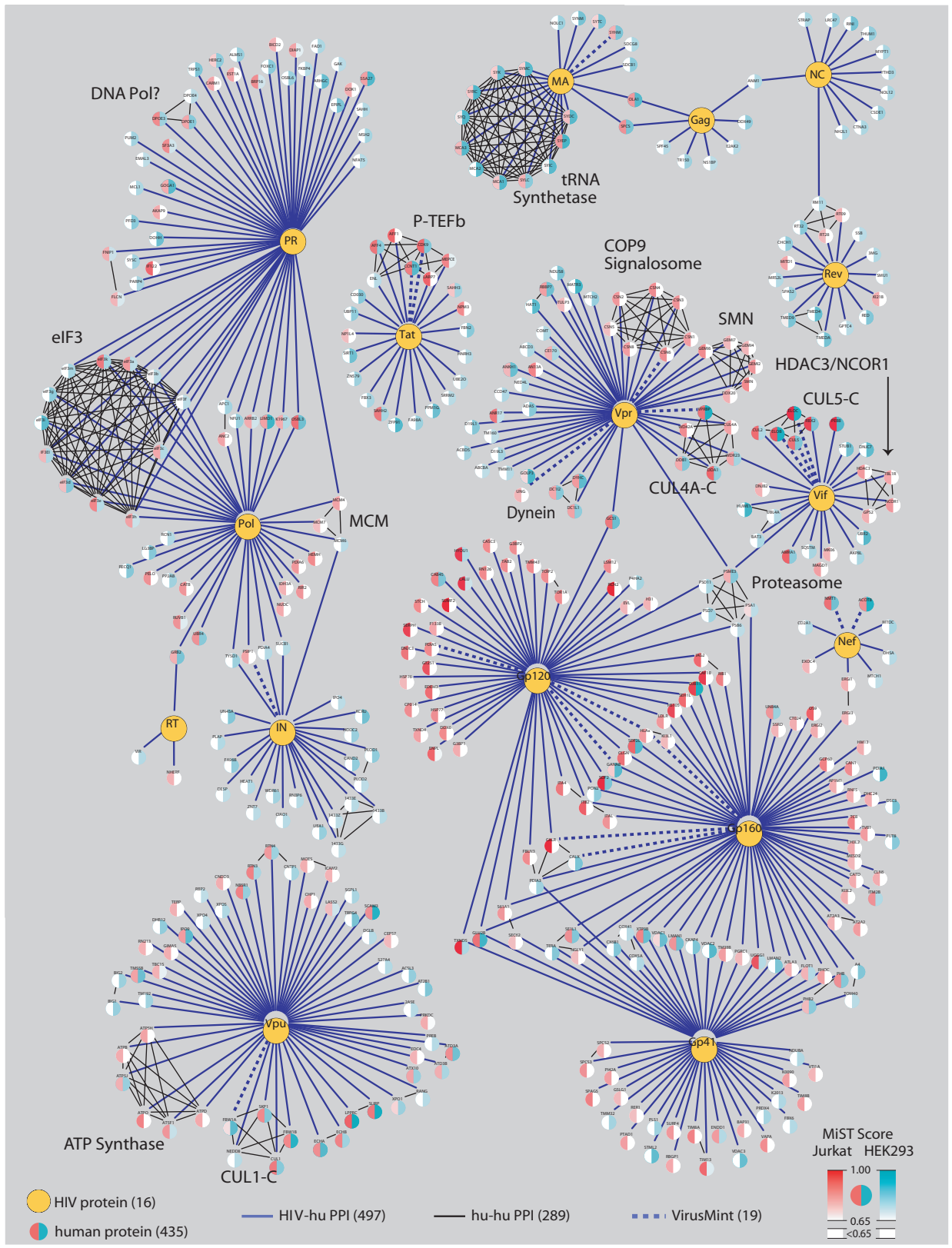
Figure 2.4. eIF3d is cleaved by HIV-1 PR and inhibits infection a, MiST scores for eIF3 subunits associated with PR (right) and Pol (left) in HEK293 and Jurkat cells. Sizes of the proteins and numbers of significant interactions (MiST score, >0.75) detected for Pol and its subunits are shown below, as is a modular representation of the eIF3 complex29. The cleaved subunit, eIF3d, is in red. b, Western blot of HEK293 cell lysate expressing FLAG-tagged eIF3 subunits in the absence (−) or presence (+) of active PR probed with an anti-FLAG antibody. c, HEK293 cells were co-transfected with Gag, or FLAG-tagged eIF3d, PABPC1, BCL2 and increasing amounts of PR. Cell lysates were probed against Gag (upper panel), FLAG-tagged eIF3d (middle panel) or tubulin as control (lower panel). d, Silver stain of purified eIF3 complex incubated with recombinant HIV-1 PR. The residues corresponding to the eIF3d cleavage site (red) is located within the RNAbinding domain26. e, f, HeLa-derived P4/R5 MAGI cells were transfected with two different short interfering RNAs (siRNAs) targeting individual subunits of the eIF3 complex (Supplementary Tables 7 and 9) and subsequently infected with either a pNL4-3-derived, VSV-G-pseudotyped, single-cycle virus (HIV–VSV-G) (e) or wild-type pNL4-3 (f). NC, negative control. g, Early (left) and late (right) HIV-1 DNA levels measured by quantitative PCR amplification in cells transfected with two independent eIF3d siRNAs or with control siRNAs. Samples were normalized by input DNA amount or by cellular gene (HMBS) copy number. The RT and replication assays were done three to five times and the standard deviations are shown (Supplementary Tables 7, 10 and 11). *P < 0.05 (Kruskal–Wallis test with Dunn's correction for multiple comparisons).

a

	Pol Jurkat	Pol HEK293	PR Jurkat	PR HEK293
eIF3d	0.803	0.931	0.791	0.842
eIF3k	0.841	0.816	0.868	0.784
eIF3c	0.848	0.876	0.786	0.764
eIF3a	0.833	0.881	0.809	0.748
eIF3l	0.825	0.875	0.758	0.771
eIF3e	0.837	0.786	0.847	0.756
eIF3h	0.860	0.794	0.797	0.726
eIF3f	0.864	0.886	0.592	0.751
eIF3b	0.729	0.815	0.664	0.751
eIF3i	0.686	0.896	n/a	0.739
eIF3g	0.686	0.823	n/a	0.731
eIF3m	0.686	0.686	0.443	0.788
eIF3j	n/a	n/a	n/a	n/a

MiST Score
Jurkat   HEK293
1.00
0.65
<0.65
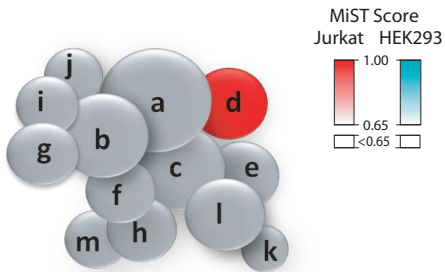
	Pol	PR	RT	IN
Size (kDa)	107.4	10.8	64.4	32.2
Interactions	38	53	3	25
Interactions/kDa	0.35	4.9	0.047	0.78

b

kDa   -PR   +PR   PR
a
b, c
72 Đ   d, l
55 Đ   e, f
g, h, m
i, j
12 Đ

P3 P2 P1 P1' P2' P3'

eIF3d   RNM-LQF   (114/115)
1   86   118   548
RNA binding domain

c

c   d   e   f   g   h   i   k   l   m
kDa   Đ +   Đ +   Đ +   Đ +   Đ +   Đ +   Đ +   Đ +   Đ +   Đ +
150 Đ
100 Đ
75 Đ
50 Đ
37 Đ
25 Đ

d

0 0.5 1 2 5 10 20 50 100 300   ng PR

Gag
eIF3d
PABPC1
BCL2
tubulin

e   pNL4.3? nef (VSVg)
Normalized luciferase expression
siRNA
NC
eIF3c
eIF3c
eIF3d
eIF3d
eIF3e
eIF3e
eIF3f
eIF3f
eIF3g
eIF3g
eIF3i
eIF3i
0 1 2 3 4 5 6 7 8

f   pNL4.3
Normalized §-gal expression
siRNA
NC
eIF3c
eIF3c
eIF3d
eIF3d
eIF3e
eIF3e
eIF3f
eIF3f
eIF3g
eIF3g
eIF3i
eIF3i
0 1 2 3 4 5 6

g   Early RT
Fold Increase
siRNA
NC
eIF3d
eIF3d
0 1 2 3 4 5

Late RT
Fold Increase
siRNA
NC
eIF3d
eIF3d
0 1 2

37

Chapter 3: Unpublished analysis of HIV protease host-protein interactions

3.1 Introduction

This chapter includes further analysis of HIV protease host-protein interactions which were not included in the publication "Global landscape of HIV-human protein interactions" [1]. These findings were the result of studies of the candidate HIV protease interacting proteins identified in the above publication, as well as additional affinity purification-mass spectrometry (AP-MS) experiments using catalytically inactive HIV-1 protease, simian immunodeficiency virus (SIV) protease, and two variations of HIV-1 protease containing mutations associated with resistance to HIV protease inhibitors.

These experiments provide further elucidation of the published interaction between HIV protease and eIF3d, as well as revealing a number of other host proteins that are cleaved by this enzyme and have not been published as substr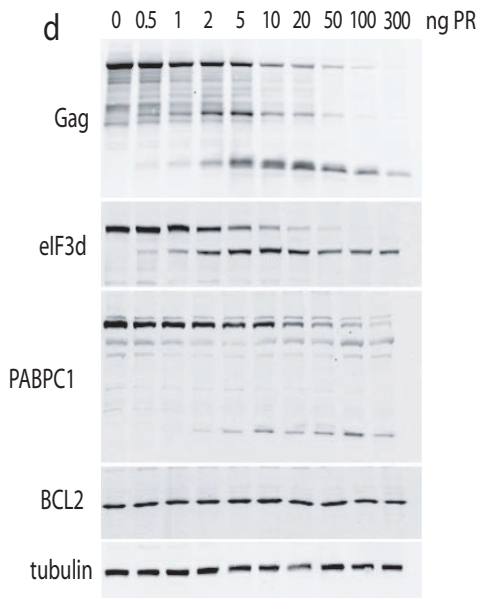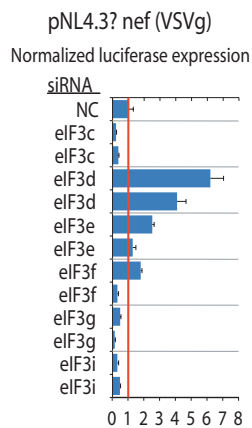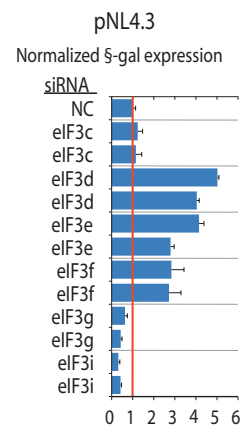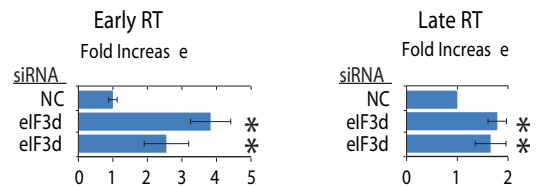ates. Furthermore, analysis of the overlap in host-protein interactions between the four above mentioned enzymes is described. Also included in these results is data demonstrating that cleavage of endogenous or over expressed host protein substrates including eIF3d and splicing factor 3A3 (SF3A3) within virally infected cells is challenging to detect using the techniques currently available. Finally, we describe the use of a nanosensor for sensitive observation of HIV protease activity.

As revealed by our studies of HIV protease cleavage of eIF3d and other host protein substrates, determining the impact of HIV protease activity within a virally infected cell will require the ability to accurately measure small changes in viral replication as well as the local cleavage of a subset of proteins within an infected cell. These factors imply that approaches such as Western blot may not be sensitive enough to detect these subtle alterations. Alternative techniques with greater sensitivity must be explored.

3.2 Methods

<u>Figure 3.1</u>: Anti-flag immunoprecipitation was conducted using anti-FLAG M2 affinity gel

agarose (Sigma-Aldrich) per manufacturer's directions. FLAG peptide (Sigma-Aldrich) was

used for the elution step following the manufacturer's provided protocol. Purified human eIF3

complex and eIF4G was supplied by Duane Smith of the Cate laboratory (UC Berkeley). The

eIF3 complex was digested using purified, recombinant HIV-1 protease (NL4.3) under

conditions previously described (Jäger *et. al*) [1]. The anti-eIF3b Western blot was performed

with antibody targeting endogenous human eIF3b (Santa Cruz Biotechnology) 1:1,000 dilution

in 2% BSA followed by goat anti-rabbit HRP conjugated secondary antibody, 1:2,000 dilution in

2% BSA.

<u>Figure 3.2</u>: C-terminally FLAG tagged human proteins were transfected into HEK 293T cells

using calcium phosphate. Lysates were made from the transfected cells after approximately 48

hours using NP-40 as the detergent as detailed [1]. Lysates were incubated *in vitro* with

recombinant, purified HIV protease followed by an anti-FLAG Western blot performed using

anti-FLAG M2 antibody (Sigma-Aldrich) at a 1:1,000 dilution in 2% BSA followed by an HRP-

conjugated goat anti-mouse secondary antibody at a 1:2,000 dilution in 2% BSA.   .

<u>Figure 3.3</u>: Panel A, HEK 293 cells were infected by HIV-1 virus pseudotyped with vesicular

stomatitis virus envelope glycoprotein (VSV-g) resulting in a single round of infection. A range

of multiplicity of infection (MOI) was used. 48 hours after infections cells were lysed and

endogenous eIF3d and tubulin was analyzed using antibodies against the native proteins.

ImageJ was used to calculate the ratio between these two proteins based on intensity of the

Western blot signal. Panel B, HEK 293 cells were infected with the maximum MOI using HIV-1

VSV-g pseudotyped virus. Cells were collected at time points indicated following infection from

0-10 hours. Lysates were prepared and endogenous eIF3d and tubulin was analyzed as above.

Panel C, Jurkat T-REx cells (Life Technologies), were stabling transfected with C-terminally

FLAG-tagged SF3A3 or eIF3d. The cells were then infected with wild type NL4.3 HIV-1 virus

using polybrene and spinoculation to increase the MOI to approximately 10-20 virions/cell. Cells were collected at the designated time points and lysates were analyzed by an anti-FLAG Western blot as above.

Figure 3.4: Multiple drug resistant HIV-1 protease mutations were identified through the Standford HIV drug resistance database [2]. Models of HIV-1 protease and SIVmac protease were constructed using PyMol (open access, Schrödinger).

Figure 3.5: All protease constructs were strep tagged and contained the inactivating D25N mutation of the catalytic aspartic acid. They were expressed within HEK 293 cells and AP-MS and protein identification was then performed as published [1]. The identified host proteins were compared to the Krogan lab database of proteins that have been observed to co-immunoprecipitate with viral protein constructs. Host proteins were excluded if they have been found to associate with more than one other viral protein, as this indicates proteins that are non-specific binders. Purified *Saccharomyces cerevisiae* anaphase promoting complex (APC) was provided by Vanessa Van Voorhies of the Morgan laboratory (UCSF) and was digested *in vitro* with purified, recombinant HIV-1 protease using the conditions described [1] and protease at either 40nM or 160nM. The digestion was then analyzed by a silver stain kit (Thermo Fisher).

Figure 3.6: Gold nanosensors were generated as in *Tajon et al*. The peptide sequence used to link the gold nanoparticles was SQNYPIVQ derived from the matrix-capsid cleavage site within the HIV-1 gag protein. 500nM HIV-2 protease was used to cleave the nanoparticles using a flow chamber, as detailed in *Tajon et al* [3]. The scattering intensity of individual nanoparticles were recorded over the length of one hour and plotted. Stable decreases in scattering intensity indicate cleave events.

3.3 Results

Cleavage of eIF3d does not disrupt associated between eIF3 and eIF4G

Many viruses disrupt their host's translation machinery. In particular, eIF4G is often targeted as

it is required for cap-dependent host protein translation, whereas many viruses have cap-independent methods of translating their own proteins [4-6]. Therefore, we investigated whether cleavage of eIF3d by HIV protease could be a mechanism to interfere with the association of the eIF3 complex and eIF4G, an essential interaction for eukaryotic translation initiation.

We obtained purified, human eIF4G bearing a C-terminal FLAG tag as well as purified, human eIF3 complex. A portion of the eIF3 complex was then incubated with recombinant, purified, wild type HIV protease in order to cleave the eIF3d subunit. After fully cleaving eIF3d and inactivating the HIV protease using an inhibitor, the tagged eIF4G was combined with either the digested eIF3 complex or an undigested eIF3 control sample and an anti-FLAG immunoprecipitation was carried out (Fig 3.1A). The presence of a subunit of the eIF3 complex, (eIF3b), was analyzed by Western blot to probe for co-immunoprecipitation between eIF3 and eIF4G. The quantity of eIF3b recovered in the anti-FLAG pull-down was approximately equal regardless of whether digested or undigested eIF3 complex was used, indicating that cleavage of eIF3d by HIV protease likely does not significantly disrupt binding between eIF3 and eIF4G (Fig 3.1B).

Multiple host proteins identified by AP-MS are HIV protease substrates

While eIF3d was the first host protein substrate identified through AP-MS of HIV-1 protease, over 40 additional eukaryotic proteins were also detected by AP-MS with this enzyme [1]. To determine whether any of these other eukaryotic proteins are also HIV protease substrates, we individually cloned the candidate proteins into a vector containing a C-terminal FLAG tag and a CMV promoter. These proteins were then transiently expressed in HEK 293T cells, a cell line optimized for high levels of protein expression. Lysates were made from the cells as described in Jäger *et. al* [1] and these samples were then digested *in vitro* using purified, recombinant HIV protease under the conditions previously detailed [1]. An anti-FLAG Western blot was performed to determine which of the candidate proteins were cleaved by HIV protease as

evidenced by the appearance of one or more distinct product bands after incubation with the

enzyme. 41 candidate host proteins (including all subunits of the eIF3 complex), were cloned

from the original 53 proteins identified in the AP-MS dataset. Analysis of these proteins

revealed that 8 are cleaved by HIV-1 protease, as illustrated in figure 3.2A. This result indicates

that approximately one fifth of proteins identified using this AP-MS approach are HIV protease

substrates. The cleaved host proteins are involved in many different cellular pathways,

however, we noted that a majority of proteins are involved in nucleic acid binding (Fig 3.2B).

This is an interesting result given the proposed function of HIV protease-mediated cleavage of

eIF3d and the recent finding that RNA enhances HIV protease activity [1,7].

Efforts to detect cleavage of host proteins within HIV infected cells

Although it is clear that HIV protease cleaves host proteins *in vitro*, demonstration of cleavage

within virally infected cells is important because it contributes to verifying that these events are

physiologically relevant. We attempted to detect cleavage products of eIF3d or a decrease in

full length eIF3d using Western blot analysis of virally infected cells (Fig 3.3A-B). To determine

if the quantity of full length protein was decreased, we performed a Western blot against the

endogenous eIF3d protein and endogenous tubulin. The ratio between eIF3d and tubulin was

calculated using ImageJ and normalized to the ratio present in a mock infection. In Fig 3.3A, we

calculated the ratio between eIF3d and tubulin 48 hours after infection of HEK 293 cells with

VSV-G pseudotyped HIV-1 over a range of multiplicity of infection (MOI). No decrease in this

ratio was observed at any MOI tested. To determine if we would see a decrease in full length

eIF3d at earlier time points after infection, we repeated this experiment using the highest MOI

feasible and collected samples from 0-10 hours following HIV infection (Fig 3.3B). In this

experiment we observed that during the first five hours after infection there was a small

reduction in the ratio of eIF3d/tubulin, however, no cleavage products were detected.

Because the observation of cleavage products is the most conclusive evidence to support

proteolysis of these substrates by HIV protease during a viral infection, we attempted to detect them using a T cell line (Jurkat T-REx) stably transfected with FLAG-tagged versions of HIV protease host-protein substrates (Fig 3C). This system had the advantages of a higher expression level of the substrate proteins, coupled with the ability to use an anti-FLAG antibody to detect cleavage products rather than antibodies against the endogenous protein, which were more difficult to use. The cells were infected with wild type HIV-1 (NL4.3) at an MOI of 10-20, the maximum possible MOI under these conditions. Samples of infected cells were collected at the designated time points, lysates made and analyzed by anti-FLAG Western blots. No cleavage products were observed for either substrate protein tested (SF3A3 or eIF3d) at any time point after infection.

Comparison of host-protein interactions between related retroviral proteases

If interactions between HIV-1 protease and host proteins are important for viral replication, it is anticipated that these interactions should be conserved between related retroviral proteases. Similarly, these interactions would be expected to be preserved in virulent drug resistant mutants of the protease, (or compensatory host protein interactions formed). To test this hypothesis, we conducted AP-MS using tagged, catalytically inactive constructs of a commonly occurring multiple drug resistant (MDR) variant of HIV-1 protease and an enzyme related to HIV-1 protease, SIV protease (SIVmac239) (Fig 3.4A). The MDR protease contained the mutations: M46I, I54V, I84V, L90M which are located in or near the active site and are strongly associated with drug resistance as well as the mutations I13M, L19I, K20I, M36I, N37S, I62V, I64V, I66V, C67F, H69Q, I72T, I93L which are mostly located at sites distant from the active site and are not associated with drug resistance (Fig 3.4A). To investigate the impact of these non-drug resistance associated mutations, an additional HIV-1 protease mutant (referred to as PR minor) was made bearing only these mutations.

AP-MS analysis was conducted using the methods published previously and the overlap

in host protein interactions was graphed (Fig 3.5A) [1]. The number of total host proteins co-immunoprecipitating with each construct is listed and in parenthesis, the number of these proteins that were unique to that construct, meaning that they have never been observed with another bait protein in any proteomic studies of viral proteins conducted by the Krogan laboratory. PR minor was found to have the greatest overlap in host-protein interactions with wild type HIV protease, suggesting that a majority of host proteins interact with these protease constructs through their active sites, (identical in these two enzymes). This is perhaps not a surprising finding because this is the region of the enzyme that is designed to bind proteins. The SIV protease construct had the least overlap with wild type HIV-1 protease, an expected result given that the SIV construct has many more sequence differences from wild type HIV-1 than either of the mutant HIV-1 constructs. Of particular interest are the six proteins that were observed to bind to all four constructs as these indicate potential evolutionarily conserved interactions. These proteins were adenosylhomocysteinase (AHCY), histone deacetylase 2 (HDAC2), protein KIAA1967 (referred to as K1967 in Fig 3.2), and three members of the anaphase promoting complex (APC), subunits 5, 7 and 16. Because three proteins belonging to the APC complex were identified, we investigated the interactions between HIV protease and the APC through *in vitro* digestion assays, modeled after our investigation of the eIF3 complex. Since purified human APC was not accessible, our initial experiments were conducted with purified APC from *S. cerevisiae*. *In vitro* digestion assays revealed that at least one subunit appears to be cleaved (APC2), while at least three smaller product bands were detected.

Nanosensors can be used to detect single proteolytic events
Due to the challenges we faced in attempting to detect cleavage of host protein substrates in virally infected cells through Western blot, we investigated the possibility of designing a gold nanosensor specific for HIV protease activity. Gold nansensors are two or more nanoparticles composed of a zinc, iron and silicon dioxide dielectric core covered in a gold shell and

connected by a protease sensitive linker sequence [3,8]. The linker is designed to position the

nanoparticles at the optimal distance for resonance energy transfer to occur, causing a high

scattering intensity when viewed under a light microscope. When the linker sequence is

cleaved, the nanoparticles diffuse away from each other, causing a sharp decrease in the

observed scattering intensity.

We designed gold nanosensors linked by the peptide sequence cleaved by HIV-1

protease to separate the gag proteins matrix and capsid into their individual components

(SQNYPIVQ) (Fig 3.6A). We then incubated these sensors with 500nM of purified, recombinant

HIV protease in a flow chamber and recorded the scattering intensity of particles over one hour.

By plotting the scattering intensity of individual sensors, the occurrence of cleavage events is

evident by sudden, sustained decreases in scattering intensity. As multiple nanoparticles can

be attached to each other, multiple cleavage events can be observed by monitoring a single

nanosensor. In Fig 3.6B, we show examples of a nanosensor that was cleaved three times (top

panel) or one time (bottom panel). While these experiments were conducted in an *in vitro*

system, these nanoparticles have been used successfully to analyze protease activity in cells

and cell lysates [3,8,9]. Analysis of HIV infected whole cells or lysates could be conducted

using these sensors to interrogate HIV protease activity in the context of a viral infection.

Replacement of the gag linker sequence with a sequence from a host protein substrate could

then serve as a reporter for cleavage of host proteins during viral infection.

3.4 Discussion

In the experiments presented here, we report evidence that HIV protease-mediated

cleavage of eIF3d does not disrupt association between the eIF3 complex and eIF4G. While

many viruses are known to affect their host's translation pathway by cleaving or inactivating

eIF4G or other proteins involved in cap dependant translation, this does not appear to be the

purpose of HIV protease mediated cleavage of eIF3d. Further experiments can be done to

investigate this, such as observing the impact of HIV protease treatment on cap-dependent and cap-independent protein translation in an *in vitro* transcription/translation system, and by studying interactions between eIF3d and the viral RNA.

We also present evidence that HIV protease is capable of cleaving multiple eukaryotic proteins, many of which are involved in nucleic acid binding. This result is consistent with previous reports demonstrating that HIV protease cleaves a range of host proteins [10-12]. It suggests that this enzyme may exert influence on a number of host pathways, each contributing to small improvements in viral fitness. All together it is possibly that these cleavage events make a significant contribution to viral replication but individually it may be challenging to detect their impact. This result highlights the importance of developing more sensitive methods for detecting proteolysis, such as the development of gold nanosensors that are capable of reporting a single proteolytic event.

Finally, we report analysis of the overlap in host protein interactions between three protease constructs related to wild type HIV-1 protease. These enzymes include a commonly occurring drug-resistant variant of HIV-1 protease containing mutations associated with drug resistance and additional mutations of unknown relevance, a variant of HIV-1 protease synthesized to include only the mutations not associated with drug resistance found in the MDR construct, and the protease from SIV, a close relative of HIV. From AP-MS analysis of these constructs, we found that it is likely that most HIV protease host-protein interactions are mediated through binding to the active site of the enzyme, since the differences in host protein binding between the enzymes correlated with differences in amino acids surrounding the active site. We also identified six proteins which may be evolutionarily conserved host protein interactions since they co-immunoprecipitate with all four protease constructs but not with other viral proteins. Curiously, half of these proteins were members of the APC, a large complex regulating the eukaryotic cell cycle. HIV accessory proteins are known to influence cell cycle progression, however, no link between HIV protease and the cell cycle has yet been established

[13]. Using purified APC from *S. cerevisiae*, we revealed that at least one subunit is cleaved *in vitro* by HIV protease. This result is important to investigate further using human APC, as it may indicate an additional host pathway regulated by HIV protease activity.

The extended results presented in this chapter provide further support of the initial conclusions published in Jäger *et. al.* regarding a dual role for HIV protease in viral and host protein processing. While many questions remain, it is becoming clear that the role of this enzyme is more extensive than cleavage of viral proteins. Uncovering these additional functions and determining their significance will likely require the development of innovative and sensitive techniques for measuring proteolytic activity.
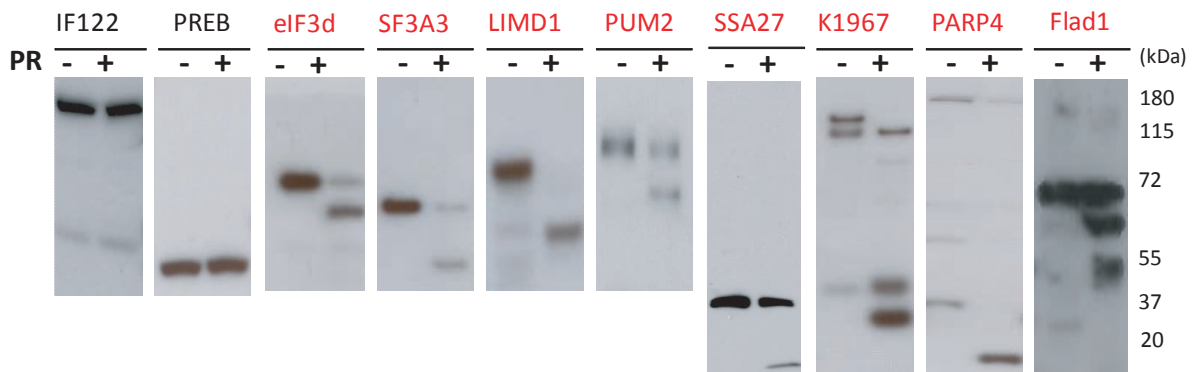
3.5 References

1. Jager S, Cimermancic P, Gulbahce N, Johnson JR, McGovern KE, et al. (2012) Global landscape of HIV-human protein complexes. Nature 481: 365-370.
2. Shafer RW (2006) Rationale and uses of a public HIV drug-resistance database. J Infect Dis 194 Suppl 1: S51-58.
3. Tajon C, Jun YW, Craik CS (2014) Single-molecule sensing of caspase activation in live cells via plasmon coupling nanotechnology. Methods Enzymol 544: 271-297.
4. Gradi A, Foeger N, Strong R, Svitkin YV, Sonenberg N, et al. (2004) Cleavage of eukaryotic translation initiation factor 4GII within foot-and-mouth disease virus-infected cells: identification of the L-protease cleavage site in vitro. J Virol 78: 3271-3278.
5. Song QQ, Lu MZ, Song J, Chi MM, Sheng LJ, et al. (2015) Coxsackievirus B3 2A protease promotes encephalomyocarditis virus replication. Virus Res 208: 22-29.
6. Glaser W, Skern T (2000) Extremely efficient cleavage of eIF4G by picornaviral proteinases L and 2A in vitro. FEBS Lett 480: 151-155.
7. Potempa M, Nalivaika E, Ragland D, Lee SK, Schiffer CA, et al. (2015) A Direct Interaction with RNA Dramatically Enhances the Catalytic Activity of the HIV-1 Protease In Vitro. J Mol Biol 427: 2360-2378.
8. Tajon CA, Seo D, Asmussen J, Shah N, Jun YW, et al. (2014) Sensitive and selective plasmon ruler nanosensors for monitoring the apoptotic drug response in leukemia. ACS Nano 8: 9199-9208.
9. Jun YW, Sheikholeslami S, Hostetter DR, Tajon C, Craik CS, et al. (2009) Continuous imaging of plasmon rulers in live cells reveals early-stage caspase-3 activation at the single-molecule level. Proc Natl Acad Sci U S A 106: 17735-17740.
10. Shoeman RL, Honer B, Stoller TJ, Kesselmeier C, Miedel MC, et al. (1990) Human immunodeficiency virus type 1 protease cleaves the intermediate filament proteins vimentin, desmin, and glial fibrillary acidic protein. Proc Natl Acad Sci U S A 87: 6336-6340.

11. Shoeman RL, Kesselmier C, Mothes E, Honer B, Traub P (1991) Non-viral cellular substrates for human immunodeficiency virus type 1 protease. FEBS Lett 278: 199-203.
12. Wagner RN, Reed JC, Chanda SK (2015) HIV-1 protease cleaves the serine-threonine kinases RIPK1 and RIPK2. Retrovirology 12: 74.
13. Rice AP, Kimata JT (2015) Subversion of Cell Cycle Regulatory Mechanisms by HIV. Cell Host Microbe 17: 736-740.
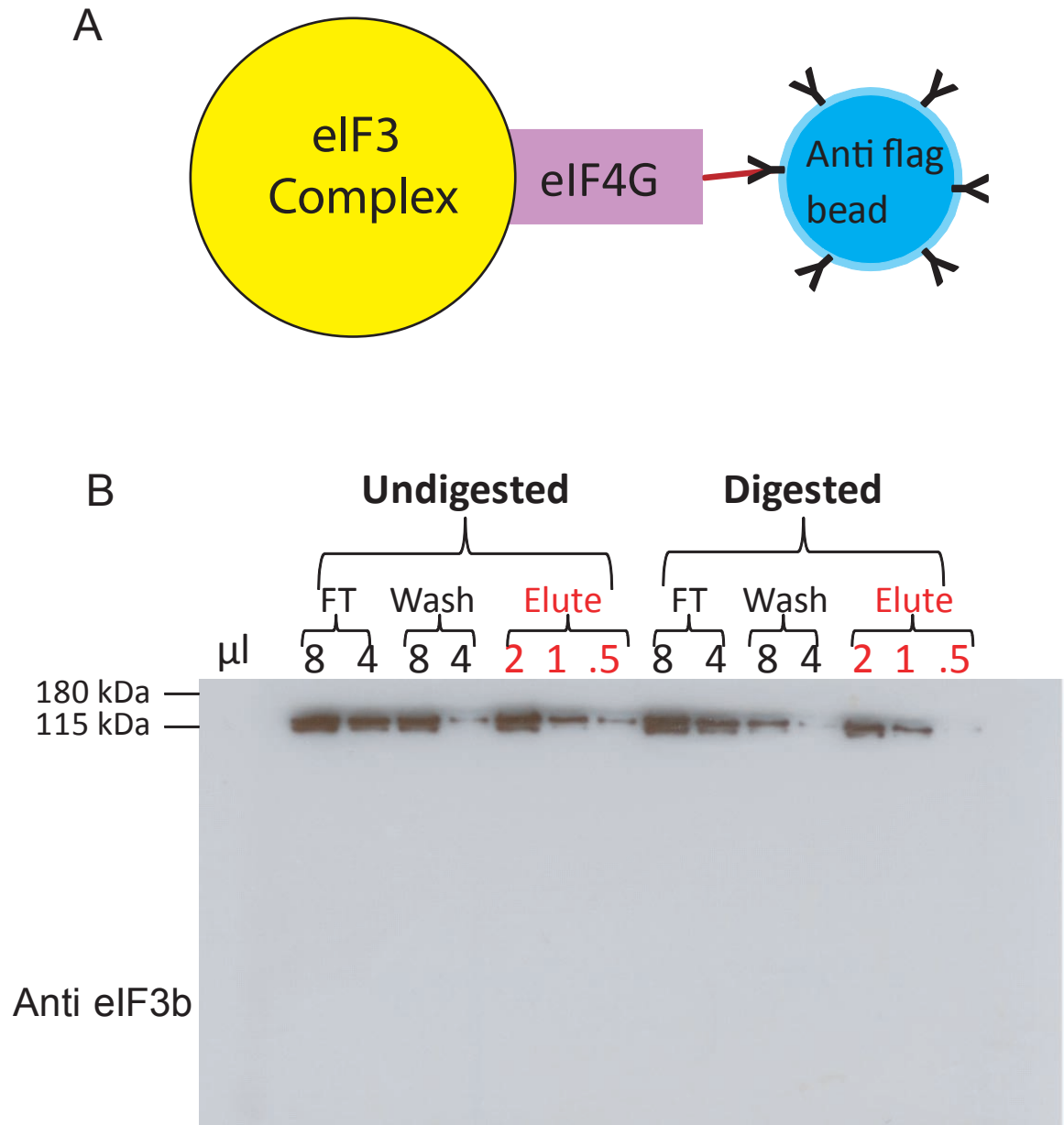
Addendum Figure 3.1 The impact of eIF3d cleavage on the association between the eIF3 complex and eIF4G. (A) Experimental schematic: FLAG tagged eIF4g was incubated with eIF3 complex that had been either treated with active HIV protease or not. An anti-FLAG pull-down was then conducted and the subunits of the eIF3 complex analyzed by Western blot. (B) Western blot against the eIF3b subunit reveals that proteolysis of eIF3d by HIV protease does not effect co-immunoprecipitation of eIF3b with eIF4G. FT stands for flow through, FT, Wash and Elute refer to the steps of the eIF4G anti-FLAG pull-down.

A

| | IF122 | PREB | eIF3d | SF3A3 | LIMD1 | PUM2 | SSA27 | K1967 | PARP4 | Flad1 | (kDa) |

PR − + − + − + − + − + − + − + − + − + − +

180
115
72
55
37
20

B

Name	Function
SSA27	Cell growth
eIF3d	Translation
PUM2	Translation
SF3A3	Splicing
K1967	Apoptosis
PARP4	Vault protein
LIMD1	Gene silencing
Flad1	FAD biosynthesis

Addendum Figure 3.2 Cleavage of host proteins identified as HIV protease interaction partners. (A) Numerous host proteins were found to be cleaved into distinct products by recombinant, purified HIV-1 protease. The proteins were tagged with a C-terminal FLAG tag, expressed in human cells and then the cell lysates were digested *in vitro* by HIV protease. PR - or + indicates treatment with HIV protease or a control. (B) These host protein substrates are involved in numerous cellular processes, although many are involved in nucleic acid binding.
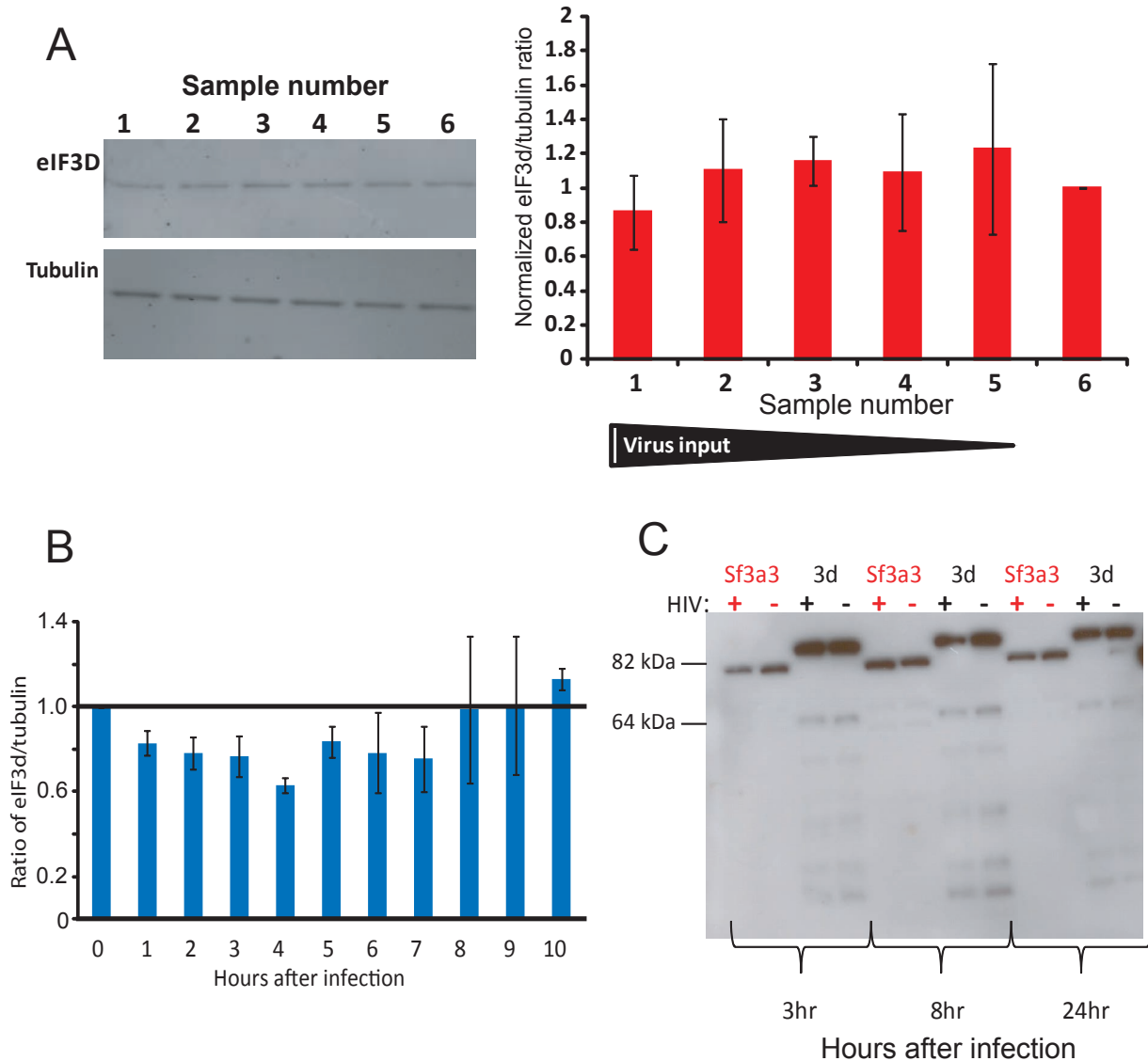
Addendum Figure 3.3 Detecting cleavage of host proteins within virally infected cells.
(A) HEK 293 cells were infected by a VSV-g pseudotyped HIV-1 virus (NL4.3) over a range of MOI. 48 hours after infection lysates were made and ratio of endogenous eIF3d to tubulin was measured by Western blot. (B) HEK 293 cells were infected at maximum MOI with VSV-g pseudotyped HIV-1. Samples were colleected 0-10 hours after infection and the ratio of endogenous eIF3d to tubulin was determined by Western blot. (C) Jurkat T-REx cells lines stably expressing C-terminally FLAG tagged SF3A3 or eIF3d were infected by HIV-1 virus (NL4.3) using polybrene and spinoculation to increase the MOI to 10-20. Samples were collected at the indicated timepoints, lysates made and anti-FLAG Western blots used to analyze cleavage.

A
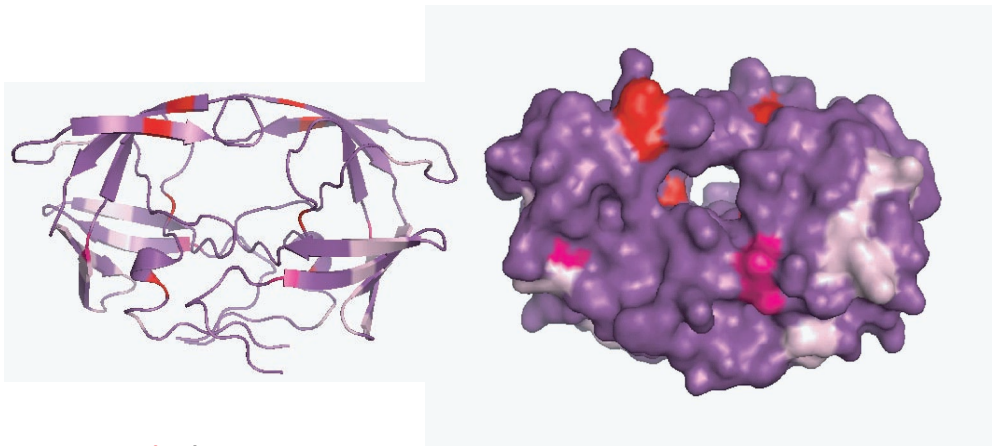
Red = drug resistance mutations
Pink = mutations contributing to drug resistance
Light Pink = non drug resistance associated mutations

B
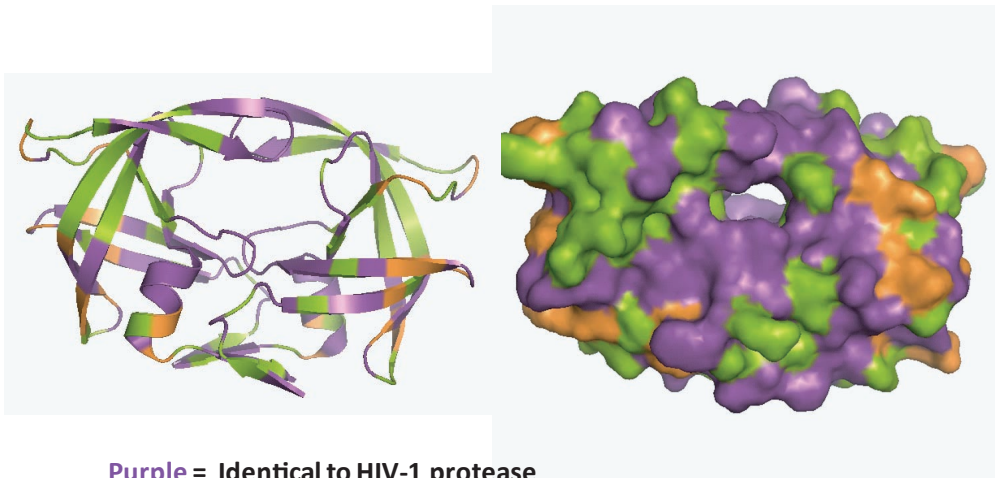
Purple = Identical to HIV-1 protease
Green = Conservative mutations from HIV-1 protease
Orange = Non-conservative mutations from HIV-1 protease

Addendum Figure 3.4. Analysis of sequence similarity between a multiple drug-resistent variant of HIV-1 protease, SIV protease and wild type HIV-1 protease. (A) The 8th most commonly observed multiple-drug resistant protease sequence was identified through the Standford HIV drug resistence database. The major mutations associated with resistence are highlighted in red and are: M46I, I54V, I84V, L90M. Additional mutations observed in this variant that do not have a significant association with drug resistance are colored in pink and include: I13M, L19I, K20I, M36I, N37S, I62V, I64V, I66V, C67F, H69Q, I72T, I93L. A mutant containing all of these mutations or only the non-drug resistence associated mutations was generated. (B) Analysis of the similarity between SIVmac protease and wild type HIV-1 protease (NL4.3).
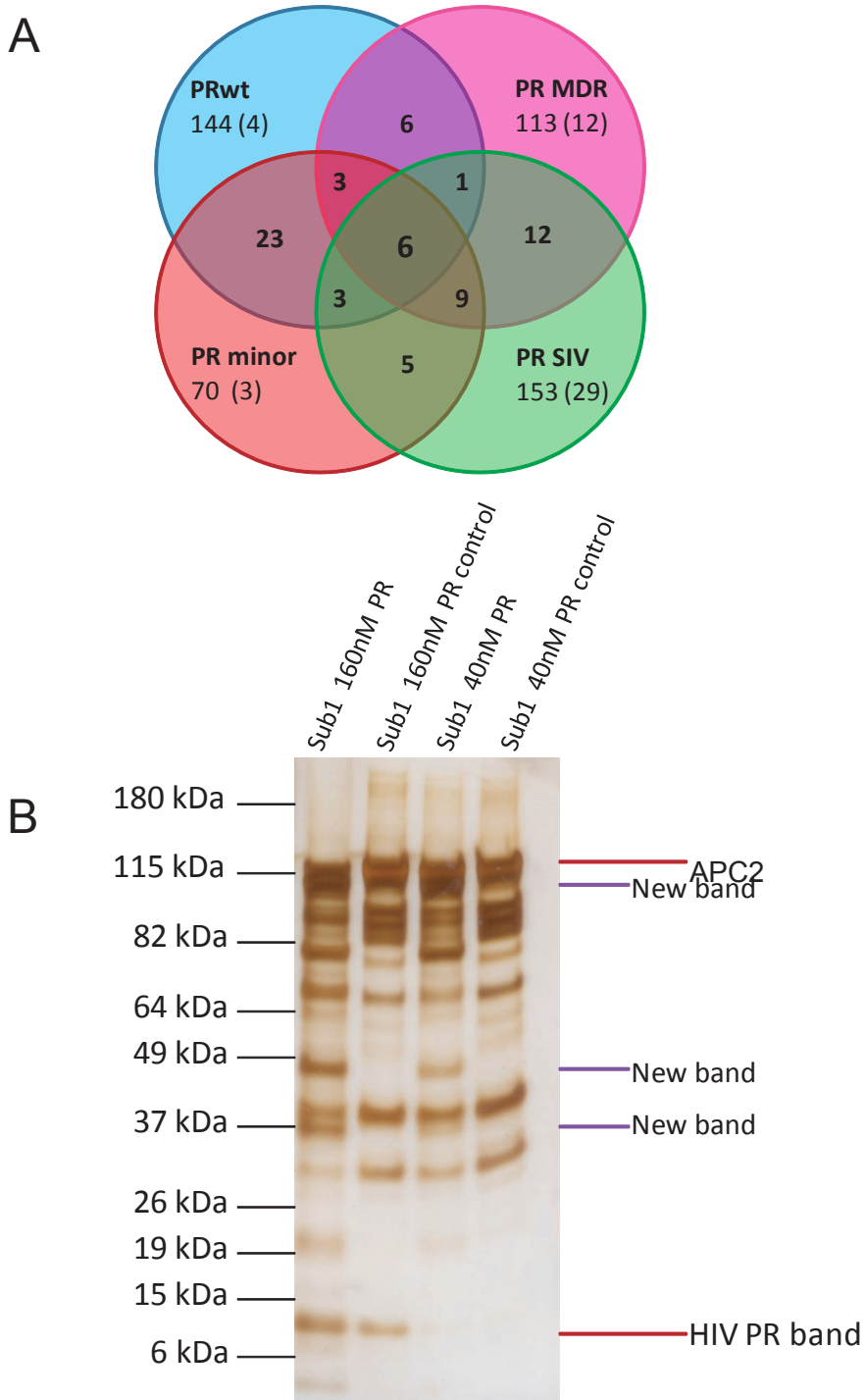
Addendum Figure 3.5 (A) Overlap in AP-MS interactions between wild type HIV-1 protease (PRwt), A multiple drug resistent HIV protease (PR MDR), an HIV protease with mutations not associated with drug resistence (PR minor) and SIVmac protease (PR SIV). The total number of identified host proteins is listed and the number of proteins unique to that construct (ie have never been observed in any AP-MS study conducted in the Krogan laboratory) are indicated in parenthesis. (B) Wild type HIV protease is capable of cleaving some subunits of the *Saccharomyces cerevisiae* anaphase promoting complex (APC). Members of the human APC were found to bind all four protase constructs.
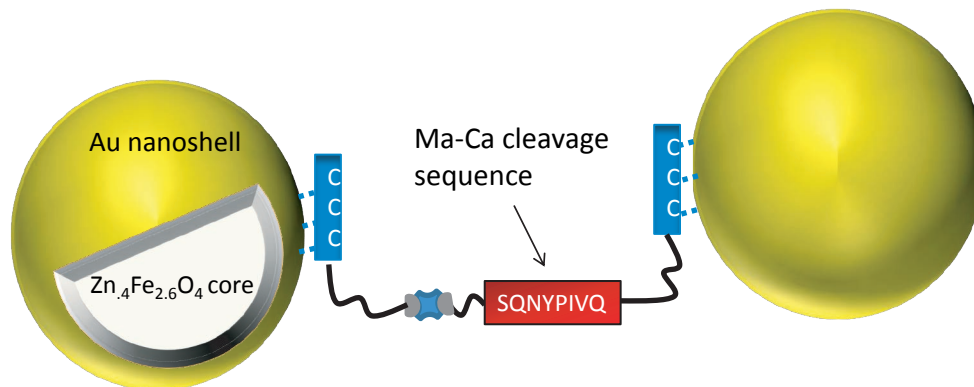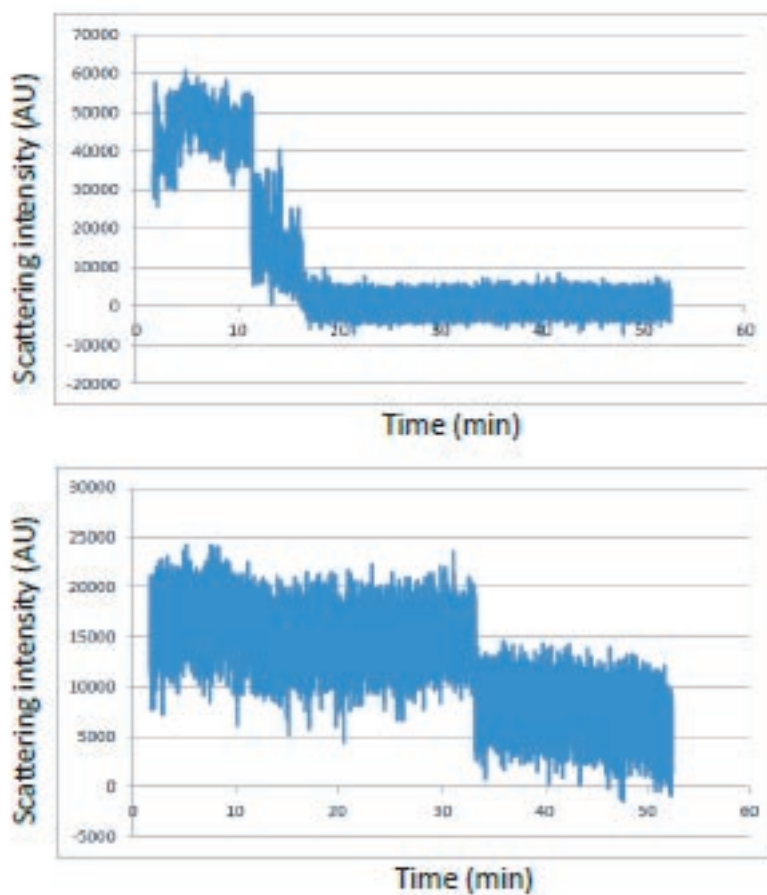
Addendum Figure 3.6 Nanosensors can be used to detect single proteolytic events.
(A) Schematic of a gold nanosensor containing an HIV protease native cleavage site, the sequence cleaved between matrix and capsid in the HIV gag protein. (B) Observation of single proteolytic events *in vitro*. As each peptide sequence is cleaved by HIV protease, the scattering intensity decreases. In the top graph three distinct proteolysis events occur, while in the bottom graph a single proteolytic event is recorded.

Chapter 4. Integrated Activity and Genetic Profiling of Secreted Peptidases in *Cryptococcus neoformans* Reveals an Aspartyl Peptidase Required for Low pH Survival and Virulence.

4.1 Abstract

The opportunistic fungal pathogen *Cryptococcus neoformans* is a major cause of mortality in immunocompromised individuals, resulting in more than 600,000 deaths per year. Many human fungal pathogens secrete peptidases that influence virulence, but in most cases the substrate specificity and regulation of these enzymes remains poorly understood. The paucity of such information is a roadblock to our understanding of the biological functions of peptidases and whether or not these enzymes are viable therapeutic targets. We report here an unbiased analysis of secreted peptidase activity and specificity in *C. neoformans* using a mass spectrometry-based substrate profiling strategy and subsequent functional investigations. Our initial studies revealed that global peptidase activity and specificity are dramatically altered by environmental conditions. To uncover the substrate preferences of individual enzymes and interrogate their biological functions, we constructed and profiled a ten-member gene deletion collection of candidate secreted peptidases. Through this deletion approach, we characterized the substrate specificity of three peptidases within the context of the *C. neoformans* secretome, including an enzyme known to be important for fungal entry into the brain. We selected a previously uncharacterized peptidase, which we term **M**ajor **a**spart**y**l peptidase 1 (May1), for detailed study due to its substantial contribution to extracellular proteolytic activity. Based on the preference of May1 for proteolysis between hydrophobic amino acids, we screened a focused library of aspartyl peptidase inhibitors and identified four high-affinity antagonists. Finally, we tested *may1Δ* strains in a mouse model of *C. neoformans* infection and found that strains lacking this enzyme are significantly attenuated for virulence. Our study reveals the secreted peptidase activity and specificity of an important human fungal pathogen, identifies

responsible enzymes through genetic tests of their function, and demonstrates how this information can guide the development of high affinity small molecule inhibitors.

Author Summary

Many pathogenic organisms secrete peptidases. The activity of these enzymes often contributes to virulence, making their study crucial for understanding host-pathogen biology and developing therapeutics. In this report, we employed an unbiased, activity-based profiling assay to examine the secreted peptidases of a fungal pathogen, *Cryptococcus neoformans* that is responsible for 40% of AIDS-related deaths. We discovered which peptidases are secreted, identified their substrate specificity, and interrogated their biological functions. Through this analysis, we identified a principal enzyme responsible for the extracellular peptidase activity of *C. neoformans,* May1, and demonstrated its importance for growth in acidic environments. Characterization of its substrate preferences allowed us to identify compounds which are potent substrate-based inhibitors of May1 activity. Finally, we found that the presence of this enzyme promotes virulence in a mouse model of infection. Our comprehensive study reveals the expression, regulation and function of *C. neoformans* secreted peptidases, including evidence for the role of a novel aspartyl peptidase in virulence.

4.2 Introduction

Cryptococcus neoformans is an opportunistic fungal pathogen responsible for 40% of all AIDS-related deaths [1,2]. Of the one million new infections occurring worldwide annually, greater than 60% result in death due to the limited efficacy and availability of therapeutics [3]. Only three classes of drugs are currently approved for treatment of fungal infections, thus there is a significant need for development of new antifungal compounds [3-5].

Peptidases are secreted by many types of pathogens including bacteria, fungi and parasites and often serve critical roles related to survival and virulence [6-11]. Direct targeting

of peptidases expressed by pathogenic organisms has proven to be a successful therapeutic strategy, notably in the development of Hepatitis C Virus (HCV) and Human Immunodeficiency Virus (HIV) protease inhibitors [12,13]. Additionally, the identification and characterization of peptidases secreted by pathogens have contributed to the formulation of new diagnostic approaches based on detection of these proteolytic activities [14-16].

Pathogenic fungi express extracellular peptidases for wide-ranging functions including host tissue invasion, nutrient acquisition and regulation of mating [17-19]. A single organism may simultaneously secrete multiple peptidases with divergent substrate specificities and requirements for activity that are tailored to their biological functions. In addition, peptidase secretion and activation are often stimulated by extracellular conditions, as distinct proteolytic functions can be important for different environments. *Candida albicans* and *Aspergillus fumigatus,* two prominent fungal pathogens, each secrete several peptidases with defined roles in virulence, while dermatophytes and the causative agent of white-nose syndrome *Pseudogymnoascus destructans* use extracellular peptidases to degrade host tissues [20-26]. Multiple peptidases have been identified in the secreted proteome of *C. neoformans*, including a metallopeptidase that is required for dissemination to the central nervous system (CNS) in a mouse infection model [27-34]. Interestingly, the level of peptidase secretion has been shown to vary between isolates in *Cryptococcus* species and in many cases higher secretion has been correlated with increased virulence [35-38]. Although these findings suggest that extracellular peptidases are involved in *C. neoformans* pathogenicity, the delineation of their functions and their validation as therapeutic targets is limited by poor understanding of their activity, specificity and regulation.

In this work, we used a comprehensive activity-based approach to characterize secreted peptidases in *C. neoformans* culture supernatants. This strategy, termed Multiplex Substrate Profiling by Mass Spectrometry (MSP-MS), relies on mass spectrometry to identify cleavage events within a defined 228-member library comprised of physiochemically diverse

tetradecapeptides [39]. The scope and design of the library allows detection of cleavage events from multiple peptidases simultaneously, and the resulting data are informative for understanding activity on both a global and individual enzyme level. Activity-based profiling stands in contrast to traditional proteomics methods that catalog which peptidases are present but does not provide information on how each enzyme contributes to the overall proteolytic activity [11,27]. Likewise, candidate-based approaches focusing on single proteolytic activities isolated from cultures may not accurately represent how these enzymes function within the secreted peptidase milieu [31,32].

To investigate the secreted peptidases of *C. neoformans* and test the influence of environment on global proteolytic activity, we cultured fungal cells under two different conditions and then isolated the cell-free supernatants for substrate specificity profiling. These experiments revealed that overall peptidase specificity differs greatly in response to extracellular conditions. To uncover the contribution of individual enzymes to the total proteolytic activity, ten candidate peptidases were individually deleted and conditioned media generated from each mutant strain was compared to the parental strain. Through this approach, we identified and defined the substrate preferences of three peptidases, including a previously uncharacterized secreted aspartyl peptidase. We found that this enzyme is the dominant contributor to extracellular endopeptidase activity at acidic pH and determined that this activity is required for tolerance to low pH environments. Analysis of its substrate specificity enabled us to screen an appropriately focused library of aspartyl peptidase inhibitors, which led to the identification of potent *in vitro* antagonists. Finally, we found that deletion strains of this enzyme are attenuated for virulence in a mouse inhalation model of *C. neoformans* infection.

Our in-depth characterization of extracellular peptidases in *C. neoformans* establishes a framework for uncovering the biological functions of these enzymes. As demonstrated by our identification of a peptidase required for virulence, examining the roles of these enzymes is critical to understanding the pathogenicity of *C. neoformans* and offers an avenue for the

development of alternative therapeutic approaches. Furthermore, the methods described here are applicable to the discovery and characterization of secreted peptidases from other pathogenic organisms.

4.3 Materials and Methods

Ethics statement: Studies in mice were carried out according to the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. All protocols were reviewed and approved by the Institutional Animal Care and Use Committee, University of California, San Francisco, approval number AN091509-02C. During infections, mice were anesthetized by an intraperitoneal injection of ketamine (75 mg/kg) and dexmedetomidine (0.5 mg/kg), which was reversed by an intraperitoneal injection of atipamezole (1.5 mg/kg). Mice were sacrificed in accordance with protocol guidelines by CO_2 inhalation and cervical spine dislocation.

Peptide-based detection of peptidase activity and characterization of substrate specificity

Fluorogenic peptide assays: Assays were conducted at room temperature on a Biotek Synergy H4 plate reader in a 50µL volume using 96-well round bottom, polystyrene plates (Corning) with λ_{ex} 328nm λ_{em} 393 nm unless otherwise stated. Substrates were each 7 or 8 amino acids with 7-methoxycoumarin-4-acetic acid or 7-methoxycoumarin-4-yl-acetyl-L-lysine on the amino terminus, and at the carboxyl terminus 4-dinitrophenyl-L-lysine or 4-dinitrophenyl bound directly to the carboxyl terminus as indicated (For sequences see S 4.1 Table). Substrates were assayed at a 10µM final concentration from DMSO stocks. Biotek Gen5 software was used to calculate initial velocities in relative fluorescent units per second (RFU/sec) from 20 points over the linear portion of each assay. To assay activity, YNB conditioned media was adjusted to pH 4.5 using 100mM MES, 100mM NaCl buffer, pH 6.5 at a final concentration of 26mM MES, 26mM NaCl, unless otherwise stated. Conditioned DMEM

media was first buffer exchanged into PBS using a centrifugal filter unit with a 3 kDa cutoff (Millipore) before use. The peptidase inhibitors pepstatin A, 1-10-phenanthroline, AEBSF and E64 were dissolved in DMSO and were obtained from Sigma-Aldrich.

Multiplex substrate profiling by mass spectrometry: Full methods are available elsewhere [39]. Minor modifications to the published method are as follows: The library contained 104 additional tetradecapeptides designed using the same algorithm as published for a total of 228 synthetic peptides. The library was split into two pools of 114 peptides to optimize detection by LC-MS/MS and 500nM of each peptide was present in the assay. YNB supernatants (32 hour cultures) were adjusted to pH 5 and diluted 1:2 in fresh YNB prior to assaying by MSP-MS, whereas DMEM supernatants (32 hour cultures) were buffer exchanged into PBS and used undiluted in the assay. The assay was conducted at room temperature and samples were removed at the time points defined [39].

Mass spectrometry was conducted on either the LTQ Orbitrap XL or an LTQ FT machine as described [39]. The full length sequences of all substrates were then deduced by comparison to the intact peptides found in the library using the Protein Prospector program v5.10.15, (UCSF), and an excel format of the results was generated using the Extractor program (UCSF) [39,40]. The frequency with which each amino acid was detected in the P4 to P4' positions was illustrated using iceLogo software [41]. All possible P4 to P4' sequences in the 228-member library were used as the reference dataset .

Proteomics Conditioned media was prepared from wild type *C. neoformans* cultured in YNB (32-hours) or DMEM (48-hours) as described below and concentrated using Millipore centrifugation filters, (3kDa molecular weight cutoff). Trypsin digestion was conducted as previously described using a 1/40 mass ratio of trypsin/protein [39]. Peptides were recovered and desalted using C18 tips (Rainin). Peptide identification was conducted as previously described using the LTQ-Orbitrap XL mass spectrometer (Thermo) [25]. To identify proteins, searches were carried out against the Uniprot database (downloaded March 21, 2012), with

Cryptococcus species entered as the taxonomy. This database was concatenated with a fully randomized set of proteins for determination of false-identification rate. Peptides were matched with up to 2 missed trypsin cleavages, carbamidomethylated cysteines as a fixed modification and oxidation of methionine, N-terminal methionine loss with or without acetylation, N-terminal acetylation or oxidation and pyroglutamate from glutamine at the N-terminus as variable modifications. Tolerance for mass accuracy was 20 ppm for parent and 0.8 Da for fragment errors.

For protein identification from the database search, the Protein Prospector settings were: 15 for the minimum protein score and 10 for the minimum peptide score. The maximum expectation value for proteins was set at 0.009 and for peptides it was 0.05. At the time of this study, the Uniprot database did not contain annotated *C. neoformans* var *grubbii* genes, thus protein matches were identified within other *C. neoformans* serotypes and the var *grubbii* orthologs were identified by searching the H99 genome either manually or through BLASTp searches using the NCBInr database, (http://blast.ncbi.nlm.nih.gov/blast/Blast.cgi).

SignalP version 4.0 was used to predict secretion signals, while SecretomeP version 2.0 was used to predict non-classical secretion pathways [42,43]. Data is reported in S 4.2 Table.

Identification of May1 orthologs was conducted by searching for *CNAG_05872* in FungiDB (www.fungidb.org) [44]. The functional domains of May1 were annotated using BLASTp. Isoelectric point and molecular weights were determined using ExPASy http://www.expasy.org/ [45]. Sites of putative N-glycosylation or O-glycosylation were predicted by NetNGlyc version 1.0 (http://www.cbs.dtu.dk/services/NetNGlyc/) [46] or NetOGlyc version 4.0 (http://www.cbs.dtu.dk/services/NetOGlyc/), respectively [47].

Yeast genetics

Yeast strains: *C. neoformans* genes were defined by Broad Institute (Cambridge, MA) annotations of the var. *grubii* H99 genome (http://www.broadinstitute.org/annotation/genome/cryptococcus_neoformans/MultiHome.html),

where each gene is named numerically as "CNAG_#" [48] All *C. neoformans* strains used in this study were derived from strain H99 using standard procedures [49] (S 4.3 Table). If unpublished, names for *C. neoformans* peptidases were assigned following the guidelines established in Inglis *et. al.* [50].

Preparation of conditioned media: Yeast cultures were grown in either YNB (1.5 g/L yeast nitrogen base, 5 g/L ammonium sulfate, 2% glucose) or in Dulbecco's Modified Eagle Medium (DMEM) without phenol red (4.5 g/L glucose, 0.584 g/L L-glutamine, 3.7 g/L $NaHCO_3$). YNB media is unbuffered and has a starting pH of 5.0, acidifying to a final pH between 1.5-2.0 in saturated cultures, while DMEM is buffered to pH 7.4. For YNB conditioned media, a single yeast colony was inoculated into 100 mL YNB and grown with shaking for a defined duration at 30°C (32 hours unless otherwise noted). The cultures were then centrifuged; the supernatant was filtered (0.45µm), flash frozen and stored at -20°C. For DMEM conditioned media, 90 ODs of log-phase *C. neoformans* cells, (the equivalent of 90mLs of a culture at an optical density at 600nm (OD_{600}) of 1), grown in YNB were spun down and inoculated into mammalian cell culture dishes containing 25 mL DMEM and maintained in a tissue culture incubator at 37°C with 5% CO_2 for 32 hours (unless otherwise noted), at which point media was harvested as described above. Because *C. neoformans* responds to light, strains were grown in darkness [51].

Mutant characterization assays: Overnight YNB cultures were adjusted to an OD_{600} of 5 and 3µL spots of 10-fold dilutions were spotted onto 2% YNB agar plates. Growth at 37°C was measured as well as tolerance to low pH, high pH, NaCl, hydrogen peroxide, sorbital, caffeine and SDS were measured through inclusion of 25mM succinic acid, 25mM MES pH 6.5, 0.75M NaCl, 0.5mM peroxide, 1M sorbital, 26mM caffeine and 0.02% SDS respectively. Melanization using L-DOPA containing plates was assayed as previously described [52]. Melanization and all plate assays apart from growth at 37°C were conducted at 30°C. Capsule production was induced through overnight growth in 100% Sabourad's liquid media followed by dilution to 10% Sabourad's liquid media adjusted to pH 8 and 48 hours growth at 37°C. Capsule was visualized

through contrast staining by India ink. Doubling times were calculated using http://doubling-time.com/compute.php [53].

May1 characterization and enzymatic assays

Immunoblot: Samples were collected at the designated time points from liquid YNB cultures and OD_{600} determined by diluting cultures into a cuvette and measuring density with a Nanodrop. 42 ODs, (the equivalent of 42mLs of a culture at an OD_{600} of 1), were then spun down and the supernatant removed and frozen. The samples were lyophilized and then dissolved in 0.17mM Tris base pH 8 and 1X SDS loading dye containing TCEP. After boiling for 15 minutes the samples were loaded onto a 4-12% Bis-Tris gel and run using MES buffer (Life Technologies). Gels were transferred to nitrocellulose membranes using the iBlot dry transfer system (Life Technologies) and blocked in 2% BSA. The monoclonal mouse anti-flag primary antibody (Sigma-Aldrich) diluted 1:2,000 in 2% BSA was used followed by HRP conjugated goat anti-mouse secondary antibody (Thermo Scientific) diluted 1:10,000 in 2% BSA. The Luminata Forte Western HRP substrate was used (EMD Millipore) and the blot imaged using a BioRad ChemiDoc imager.

Determination of May1 cleavage site within IQ-2: Matrix assisted laser desorption ionization-time of flight (MALDI-TOF) analysis was conducted to identify the site of May1 cleavage within IQ-2. 100µM IQ-2 was digested in a 50µL assay with 14.6nM May1 in 100mM MES, 100mM NaCl pH 4.5. 10µL samples were collected at the start of the reaction and after 24 hours. Peptides were recovered and desalted using Rainin C18 tips, lyphophilized, and redissolved in 5µL 0.1% formic acid. 0.5µL of sample was combined with 0.5µL matrix and analyzed by MALDI-TOF (Shimadzu Biotech Axima Performance).

May1 enrichment: To concentrate secreted May1, YNB conditioned media was prepared as described above from 32-hour cultures of wild-type *C. neoformans*. The media was then diluted 2.7-fold into buffer A (50mM Tris base pH 8), chosen to increase the pH in order to limit May1 autoproteolysis, dilute salts in the media (final conductivity ~6%), and confer a negative charge

to the peptidase domain. The media was then loaded onto a 1mL HiTrap DEAE fast flow

column (GE Healthcare) using a fast protein liquid chromatography system with a flow rate of

1.5mL/min. May1 was eluted using a 30 minute linear gradient of 0-100% buffer B (50mM Tris

base, 1M NaCl, pH 8). Active fractions were determined by measuring activity using the

substrate IQ-2. They were then combined and approximate May1 concentration determined by

active-site titration.

Active site titration, Km and IC$_{50}$ calculations: For the following experiments the plate reader

conditions were as described for fluorogenic assays and the substrate used was IQ-2 at 10µM

final concentration since the Km of this substrate was found to be 19.64µM. Published methods

were followed with minor modifications [54]. In brief: May1 active sites were titrated using the

potent inhibitor peptstatin A and GraphPad Prism 6 software was used to determine May1

concentration from a plot of V$_i$/V$_o$ versus the log of inhibitor concentration.

Km was determined using 73nM May1 (100mM MES 100mM NaCl pH 4.5) and 0.5µM to

140µM IQ-2. A correction factor was calculated to adjust for sensitivity of the plate reader by

plotting the RFU value of complete cleavage versus the product concentration of IQ-2 from

0.5µM -10µM and dividing the V$_{max}$ values from the Km calculation by 1/slope of this line, (units:

RFU/[P]). Km was calculated by GraphPad Prism 6 software using the Michaelis-Menten

equation.

IC$_{50}$ calculations were conducted using 14.6nM May1 (100mM MES 100mM NaCl pH

4.5). Inhibitor stocks were dissolved in DMSO and incubated with May1 for 10 minutes before

addition of substrate. GraphPad Prism 6 was used to calculate IC$_{50}$ values from a plot of the log

of inhibitor concentration versus normalized response.

pH titration of May1 activity: Concentrated May1 was diluted to 14.6nM in 100mM MES,

100mM NaCl buffers from pH 1.5-7. Fluorogenic activity assays were conducted using IQ-2 and

the conditions described above.

Mouse virulence studies

C. neoformans strains were grown in liquid YNB cultures overnight, and then centrifuged and washed twice in PBS. For competitive co-infection experiments, mixtures of a wild-type strain and a deletion strain of interest were prepared by determining cell concentration using a hemocytometer and then mixing strains in a 1:1 ratio to a final concentration of $1x10^7$ cells per ml PBS, which was confirmed by plating serial dilutions. A/J female mice (Jackson Laboratory) aged 5-6 weeks were anesthetized by intraperitoneal injection of ketamine (75 mg/kg) and dexmedetomidine (0.5 mg/kg), then suspended from a silk thread by their front incisors, as described previously [49]. Intranasal infections of 50 μl were delivered by pipette, resulting in a dose of $5x10^5$ cells. After an additional 10 minutes of suspension, the mice were lowered and anesthesia reversed by intraperitoneal injection of atipamezole (1.5 mg/kg). Three mice were infected with each *C. neoformans* genotype, and were monitored until a defined terminal time point of ten days after infection. At this time, mice were sacrificed by CO_2 inhalation and cervical spine dislocation. Next, lungs were harvested and homogenized in PBS using a PRO200 homogenizer (Grainger). The ratios of *C. neoformans* strains in the input and organ samples were determined by plating in serial dilutions on Sabouraud agar plates containing 40 mg/ml gentamicin and 50 mg/ml carbenicillin, and then testing the nourseothricin resistance status of ~200 colonies. As a negative control, mice were infected with a 1:1 ratio of wild-type cells and a *sxi1Δ* strain, which is known to have a wild-type phenotype [55].

For monotypic infections, female A/J mice were intranasally infected with 50 μl PBS containing *C. neoformans* cells of a single genotype at a concentration of $1.0x10^7$ cells per ml, as described above. Ten mice were infected per genotype, and were monitored until severe morbidity (as indicated by a loss of 15% of initial body weight or other symptoms), at which point they were sacrificed. Survival data was analyzed using the Online Application for the Survival Analysis of Lifespan Assays Performed in Aging Research [56].

4.4 Results

Global secreted peptidase profiling in *C. neoformans* reveals abundant activity and environment-dependent specificity *C. neoformans* was cultured in either the microbial minimal media, yeast nitrogen base (YNB) pH 5.0, or mammalian tissue-culture media (DMEM) pH 7.4, and supernatants from each condition were assayed using a panel of internally quenched (IQ) fluorogenic peptides (Fig 4.1A, S 4.1 Table for sequences). These substrates were previously developed to detect a broad range of microbial peptidases from diverse peptidase families [57-59]. The speed and flexibility of this assay allowed us to optimize the conditions for peptidase activity and to determine which class-specific inhibitors affect it.

Although peptidase activity was evident under both culture conditions, differential substrate cleavage reflected differences in specificity. Notably, IQ-2 and IQ-6 were cleaved more efficiently by peptidases in YNB media, while proportionally higher activity was observed against IQ-3 and IQ-4 in DMEM media (Fig 4.1A). These differences suggested that alternate peptidases were active in each culture condition, which was further confirmed by assaying the substrates in the presence of class-specific peptidase inhibitors. This analysis revealed that aspartyl peptidase activity was present in YNB conditioned media while metallopeptidase activity could be detected in DMEM media (S 4.1 Fig). Adjustment of YNB supernatants from pH 5.0 to 7.4 (the pH of DMEM media) yielded only very low levels of any peptidase activity, while lowering the pH of DMEM supernatants from 7.4 to pH 5.0 produced a peptidase activity pattern similar to YNB media (S 4.1 Fig). This result suggests that growth in DMEM media using mammalian cell culture conditions induces peptidases optimized for neutral pH in addition to the acidic pH activities detected after growth in YNB media.

To investigate global peptidase substrate specificity, MSP-MS was conducted on YNB and DMEM supernatants at the optimal pH for activity observed for each condition, pH 5.0 and 7.4,

respectively. In the MSP-MS assay, peptide sequencing by LC-MS/MS is used to identify all peptide cleavage products within the 228-member library, revealing peptidase substrate specificity preferences (Fig 4.1B). Importantly, since there are no modifications to either the N or C-termini in the peptide library, both exo- and endo-peptidases can be identified in an unbiased manner.

Using MSP-MS we observed that peptidases in YNB media cleaved at 423 total sites while peptidases in DMEM media cleaved at 283 total sites (Fig 4.1C). Only 107 of these sites were cleaved by enzymes in both samples. This difference in cleavage site preference indicated that peptidase activity and specificity differs between the two culture conditions. Positional analysis of all bonds cleaved within the tetradecapeptides of the MSP-MS library illustrates the proportion of endo- and exo-peptidase activity in each sample (Fig 4.1D). In YNB supernatants, the most frequently cleaved bond was the carboxyl terminal bond between amino acids thirteen and fourteen, representing 32% of total proteolysis events. In fact, 137 of the 228 peptides had their carboxyl-terminal amino acid cleaved. Moreover, we observed that single amino acids were often sequentially hydrolyzed from the carboxyl termini of substrates until an unfavored residue was reached, consistent with the presence of abundant carboxypeptidase activity. In contrast, peptidases in DMEM media cleaved all peptide bonds within the tetradecapeptides at approximately the same frequency. These studies indicate that carboxypeptidases represent the majority of secreted proteolytic activity in *C. neoformans* cultured in YNB media and that DMEM cultures contain an equivalent representation of endo- and exopeptidase-like activity. To further illustrate the differences in proteolytic activity between the two conditions, representative examples of peptides cleaved in both samples are shown (Fig 4.1E).

Identification of secreted peptidases by proteomic and genetic approaches To identify which *C. neoformans* peptidases may be contributing to the global substrate specificity profile, we conducted a proteomic analysis of secreted proteins. We observed 199 and 131 proteins in

YNB and DMEM conditioned media respectively, with 52 proteins common to both conditions (S 4.2 Table). In total, 23 of the proteins detected had predicted secretion signals (SignalP 4.0) [42], 127 were predicted to be non-classically secreted (SecretomeP 2.0) [43], and 17 have been associated with extracellular microvesicles [60]. The remaining proteins identified had no known mechanism of secretion (S 4.2 Fig, S 4.2 Table). Seven of the proteins with predicted signal sequences were peptidases and included members of the aspartyl, metallo and serine peptidase families. Both endopeptidases and carboxypeptidases were identified, consistent with our predictions from analysis of *C. neoformans* extracellular proteolytic activity (Fig 4.1). Five of these enzymes have been detected in studies of the *C. neoformans* secretome by other groups; however Prc1 and CNAG_05872 have not been observed previously.

To determine which enzymes are responsible for the proteolytic activity present in *C. neoformans* conditioned media, we performed targeted gene deletions on ten candidate secreted peptidases (Table 4.1, S 4.3 Table). Eight of these peptidases were chosen based on their positive identification in our proteomics study. One of the trypsin-derived peptides identified in our study as originating from carboxypeptidase D could not be mapped unambiguously to a single protein, as three paralogs of this enzyme exist in the *C. neoformans* var *grubii* genome [44]. Therefore, all three genes were individually targeted for deletion (*CNAG_00919*, *CNAG_01040* and *CNAG_02966*). Because these genes are unnamed and lack orthologs in *Saccharomyces cerevisiae*, we propose naming them ***C**arbo**x**ypeptidase **D 1, 2** and **3** (*CXD1-3*), respectively. One of the aspartyl peptidases observed in our study was predicted to be GPI-anchored (CNAG_04380) [27,61], and was therefore excluded from further analysis, as our study was focused on non-cell wall anchored enzymes. Finally, two secreted peptidases that were not identified here but have been reported in previous proteomics studies were also included in the mutagenesis [27]. Two independent isolates of each of the ten strains were generated and are indicated in the text and figures by the gene name or CNAG number followed by "-1" or "-2" (S 4.3 Table).

Based on our characterization of secreted peptidase activity present in wild type

C. neoformans, we selected deletion strains for in-depth substrate profiling analysis by MSP-MS

under either DMEM or YNB culture conditions. Subsequently, by comparing the secreted

peptidase activity in conditioned media from the wild type and mutant strains, we were able to

correlate extracellular proteolytic activities to specific candidate enzymes.

DMEM conditioned media contains a metallopeptidase Mpr1 and trypsin-like

endopeptidase activity To analyze the peptidase substrate specificity of DMEM media

conditioned by wild-type cells, we generated a frequency plot from the 283 cleavage events

detected by MSP-MS (Fig 4.2A) [41]. The amino acid preferences are shown for four positions

on either side of the cleaved bond, (P4-P4'), as the majority of substrate specificity is

determined by residues closest to the scissile bond. This analysis revealed that peptidases in

DMEM supernatants prefer positively charged residues on either side of the cleaved bond, as

well as hydrophobic residues in the P1' position. Negatively charged amino acids are

disfavored at the majority of positions, and proline and glycine are both highly disfavored in

most positions from P2-P2' (Fig 4.2A).

To identify the enzymes responsible for this activity, we examined proteolytic activity in

peptidase deletion strains. Because DMEM conditioned media contained metallopeptidase

activity (S 4.1 Fig) and a single metallopeptidase, Mpr1 was identified by proteomics (Table 4.1,

S 4.2 Table), we began by investigating the contribution of this enzyme to the global specificity

profile. Mpr1 had previously been characterized as a secreted factor that is important for *C.*

neoformans invasion of the CNS [28]. Matched comparison of the substrate specificity profiles

obtained from DMEM media conditioned by wild type or *mpr1Δ* cells revealed that Mpr1

deficiency caused a loss of the P1' preference for hydrophobic amino acids seen in wild type,

(Fig 4.2A-B). However, the selection for positively charged residues on either side of the

cleaved bond remained unaltered and the same amino acids were disfavored in most positions.

To further analyze the impact of *MPR1* deletion, a Venn diagram was used to compare the overlap of cleavage events between wild type and *mpr1Δ* (Fig 4.2C). A majority of cleaved peptides were detected in both samples; however 107 cleavage events were detected in wild type but not media conditioned by *mpr1Δ*. These cleavages, presumed to be absent due to the loss of this enzyme, were used to generate a frequency plot representing Mpr1 specificity (Fig 4.2C). A prominent feature of this substrate specificity profile is enrichment for phenylalanine, leucine and norleucine (a replacement for methionine in the MSP-MS library), at the P1' position, a result that is consistent with the specificity of other peptidases predicted to be related to this enzyme (members of the M36 peptidase family) [62]. It is also notable that the P1' site exhibits the greatest degree of selectivity of any position from P4-P4'. To further illustrate the changes in substrate specificity observed in the *mpr1Δ* deletion strain, a representative example of a peptide cleaved by enzymes in both wild type and *mpr1Δ* supernatants is shown (Fig 4.2D).

An additional activity in DMEM media conditioned by wild type displays a trypsin-like preference for proteolysis between two positively charged residues, indicating the presence of serine peptidase activity [62]. This specificity is particularly evident in the *mpr1Δ* culture media (Fig 4.2B). Two serine endopeptidases were present in the deletion collection and DMEM conditioned media was analyzed from both strains (*prb1Δ* and *CNAG_00150Δ*). Deletion of either gene did not substantially impact the extracellular peptidase activity profile, suggesting functional redundancy or the existence of additional, unidentified peptidases (S 4.3 Fig). One predicted serine peptidase with a secretion signal, *KEX2,* was identified in a genome search. However, our attempts to delete this gene were unsuccessful, indicating it may be essential for *C. neoformans* survival.

An aspartyl endopeptidase May1 and the carboxypeptidase Cxd1 are the major activities in conditioned YNB media. A substrate specificity profile constructed from the 423 cleavages

observed in YNB media conditioned by wild-type cells indicated a preference for hydrolysis between hydrophobic residues, while positively charged residues, proline and glycine are disfavored (Fig 4.3A). From positional analysis of these cleavage sites, we identified carboxypeptidase activity as the dominant proteolytic activity in this media (Fig 4.1D). Since carboxypeptidases cleave the carboxyl-terminal bond, no enrichment of amino acids is evident in the P2' to P4' positions in these substrates (Fig 4.3A).

To determine whether any of the three carboxypeptidase D paralogs we identified in our proteomics analysis were responsible for the observed carboxypeptidase activity, the gene for each enzyme was deleted and conditioned media from the resulting mutant strains *(cxd1Δ, cxd2Δ* and *cxd3Δ)* was profiled by MSP-MS. After comparison of the specificity profiles, it was clear that Cxd1 is the dominant enzyme with carboxypeptidase specificity. In media conditioned by this deletion strain only 17 primary carboxyl-termini were cleaved compared to 137 in wild type (Fig 4.3B, D-E). These results indicate that Cxd1 cleaves over half of all carboxy-termini present in the library, which includes both hydrophobic and charged amino acids. By comparison, 134 and 123 primary carboxyl-termini were cleaved in media conditioned by *cxd2Δ* and *cxd3Δ,* respectively, suggesting that these enzymes do not contribute substantially to the extracellular carboxypeptidase activity (S 4.4 Fig). As anticipated, endopeptidase cleavages were not affected in any of the three carboxypeptidase deletion strains (Fig 4.3B, 3D, S 4.4 Fig).

Fluorogenic assays demonstrated aspartyl endopeptidase activity in wild type YNB supernatants (S 4.1 Fig). To assign activity to the candidate aspartyl peptidases, conditioned YNB media was profiled from the two aspartyl peptidase deletion strains listed in table 4.1 (*CNAG_05872Δ* and *pep4Δ)*. Proteolytic activity remained unchanged relative to wild type in the *pep4Δ* strain (S 4.4 Fig). In contrast *CNAG_05872Δ* conditioned media exhibited a near-total loss of endopeptidase cleavage events as well as substantially decreased carboxypeptidase activity as evidenced by proteolysis of only 55 primary carboxyl termini (Fig

4.3C-E). This result indicates that CNAG_05872 is the dominant endopeptidase under these

culture conditions. This finding is consistent with fluorogenic assays, where deletion of

CNAG_05872 led to a loss of endopeptidase activity in conditioned YNB media, while all other

strains exhibited activity levels similar to wild type (Fig 4.3F, S 4.5 Fig). As the dominant

endopeptidase, we propose renaming *CNAG_05872* to **M**ajor **A**spart**y**l peptidase 1 (*MAY1*).

May1 is a pepsin-like aspartyl peptidase, with optimal expression and activity at acidic

pH. Due to its substantial contribution to peptidase activity in YNB supernatants, we performed

an in-depth biochemical characterization of May1. This enzyme consists of a 16 residue

secretory signal (SignalP 4.0) [42] followed by an 82 residue prodomain (Fig 4.4A). The

prodomain is positively charged (pI = 9.97), which likely facilitates interaction with the negatively

charged catalytic domain (pI = 4.03) at neutral or slightly acidic pH [45]. The pepsin-like

aspartyl peptidase domain includes residues 100-434 and is expected to auto-activate in acidic

environments, causing release of the pro-domain. The N-terminal region, (position 101-223), is

also an N-terminal xylanase inhibitor domain, TAXi_N [63]. Homology between xylanase

inhibitors and fungal aspartyl peptidases has been noted previously and likely indicates an

evolutionary relationship [64].

May1 readily cleaves IQ-2 between phenylalanine and leucine (S 4.1 Table, S 4.5 Fig),

which allowed us to use fluorogenic assays to monitor enrichment of this enzyme from YNB

supernatants and investigate the impact of pH on its activity. Ion exchange chromatography

was used to enrich May1 from conditioned YNB media, resulting in a 292nM peptidase stock

solution. May1 was diluted from this stock into buffers ranging from pH 1.5 to 7.0 and activity

against IQ-2 was tested. Optimal activity was observed between pH 3.5-4.5, a range that is

consistent with other members of the aspartyl peptidase family of enzymes (Fig 4.4B) [65]. The

aspartyl peptidase antagonist pepstatin A fully inhibited proteolysis of IQ-2 in this assay,

providing further verification that May1 is the predominant endopeptidase activity under these conditions.

To investigate the time and growth dependent secretion of May1, we added a CBP-2xFLAG tag to the carboxyl-terminus through homologous recombination. By monitoring activity using IQ-2, we confirmed that the addition of this tag did not diminish May1 activity in wild type YNB conditioned media. Interestingly, although recombinant May1 activity could be detected in the culture supernatant after overnight growth, it could not be detected by immunoblot even after three days of growth. We hypothesized that the enzyme rapidly hydrolyses the C-terminal tag; therefore pepstatin A was added to the culture to inhibit this processing. This inhibitor also prevented activation of pro-May1 to mature May1, resulting in the observation of only the pro form of recombinant May1. Under these conditions the protein was detected by immunoblot after 48 hours of growth (Fig 4.4C). When *C. neoformans* was cultured in YNB media buffered to pH 6.5, the May1 activity detectable in supernatants using IQ-2 was approximately 10-fold lower than in unbuffered YNB and no signal could be seen by immunoblot after 48 hours of growth (Fig 4.4D). This result suggests that low extracellular pH could stimulate May1 secretion and activation.

The apparent molecular weight of pro-May1 detected by immunoblot was approximately 13 kDa greater than the predicted 56.8 kDa for the tagged enzyme, suggesting that this protein could contain post-translational modifications. Indeed, five asparagines within May1 are predicted to be N-glycosylated (NetNGlyc 1.0) [46] and multiple O-glycosylation sites are also predicted throughout the protein (NetOGlyc 4.0) [47].

May1 activity is required for wild-type saturation density in YNB cultures. We observed that *may1Δ* strains grown in YNB had a lower cell density at saturation than wild type or any of the other nine peptidase deletion strains (Fig 4.5A, S 4.6 Fig, S 4.4 Table). Higher cell density could not be achieved by the *may1Δ* strains even after culturing for 96 hours. These data

suggested that *may1Δ* strains were not merely slow to replicate but were incapable of growth at high density. In fact, during the exponential growth phase, *may1Δ* strains had an average doubling time of 2.36 hours which was on par with the other nine peptidase deletion strains, however all of these strains exhibited wild-type saturation densities (S 4.4 Table).

It was unclear from studies of the deletion strains, whether May1 activity or simply the presence of the May1 protein was required to reach a saturation density equivalent to wild type. Therefore, we assessed growth of wild type *C. neoformans* in the presence and absence of pepstatin A. Treatment of wild type cultures with this aspartyl peptidase inhibitor resulted in a saturation density defect equivalent to that observed for the *MAY1* deletion strains (Fig 4.5B). Importantly, no additional defects were seen in *may1Δ* strains in the presence of the inhibitor, suggesting that the growth defect observed in wild-type cells treated with pepstatin A was mediated through inhibition of May1.

Mutant characterization assays reveal a *may1Δ* growth defect at low pH. Plating assays on various stress conditions were conducted with two independent isolates of each of the ten peptidase deletion strains (S 4.7-S 4.8 Fig). After 48 hours both *may1Δ* strains exhibited a growth defect at pH 3.5 but not at pH 5.0 or 6.5 (Fig 4.5C). Longer growth periods at pH 3.5 did not allow *may1Δ* colonies to overcome this sensitivity and after three days of growth, it became apparent that *may1Δ* colonies also had a slight defect at pH 5.0 but not pH 6.5 (S 4.7 Fig). None of the other peptidase deletion strains displayed sensitivity to acidic pH, although *pep4Δ* was sensitive to hydrogen peroxide and SDS, and *prb1Δ* had a slight sensitivity to hydrogen peroxide (S 4.7-8 Fig). Based on this result, we hypothesized that the inability of *may1Δ* strains to reach wild-type saturation densities in YNB was a result of acidification of the media and would be rescued by buffering the media to pH 6.5. These culture conditions fully rescued the saturation density of *may1Δ* (Fig 4.5D). Surprisingly, it also allowed the final saturation densities of both wild type and *may1Δ* cultures to approximately double, revealing that low pH is

a condition limiting growth even for wild-type *C. neoformans*.

We also assessed the established virulence factors of capsule production and melanization for each of the ten peptidase deletion strains. *C. neoformans* secretes a polysaccharide capsule surrounding its cell wall which modulates the host immune response, while the synthesis and secretion of melanin is believed to help fungal cells resist oxidative and nitrositive stresses [66,67]. Because both capsule and melanin production occur extracellularly, we hypothesized that either process could be influenced by secreted peptidase activity. None of the deletion strains displayed abnormal capsule induction under the conditions tested; however, the serine endopeptidase deletion strain *prb1Δ* exhibited a hypomelanization phenotype (S 4.8-S 4.9 Fig).

A screen of an aspartyl peptidase inhibitor library yields compounds antagonistic to May1 While pepstatin A inhibits May1 with an IC_{50} of 1.4nM, it is an antagonist of many members of the aspartyl class of peptidases, thereby limiting its utility. To determine whether additional inhibitors targeting May1 could be obtained, we conducted an *in vitro* screen using knowledge of May1 substrate specificity derived from MSP-MS analysis. We screened a panel of 21 peptidomimetic molecules with similarities to May1 substrate preferences but with a non-cleavable bond between the P1 and P1' position (S 4.5 Table). Compounds 1 through 11 are linear inhibitors, while compounds 12 through 21 are macrocycles (S 4.5 Table) [68-71]. We also screened ten HIV protease inhibitors because some of these molecules have been reported to inhibit *C. neoformans* peptidase activity [72,73].

May1 was incubated with 100µM, 10µM and 1µM of each molecule and activity was detected using IQ-2 (Fig 4.6A, S 4.10 Fig). IC_{50} values were then calculated for the most potent compounds. The best inhibition by an HIV protease inhibitor was observed with Brecanavir, which reduced activity by 80% at 1µM and had an IC_{50} of approximately 352 nM (S 4.10 Fig). Among the peptidomimetic molecules, the macrocycles were the most potent, with the best

compounds (15 to 21) containing P2 – P3' tethered side chains, statines in P1 and an α–amino acid in P2' (Fig 4.6, S 4.5 Table). Compounds 16, 21 and 18 all exhibited nanomolar IC_{50} values of 1.6nM, 9.4nM and 41nM, respectively (S 4.10 Fig). Among the linear peptidomimetic inhibitors, those with a phenylstatine or hydroxyethylamine scissile bond isoster, (compounds 4, 7, 8 and 11), were superior to compounds with a reduced bond (1, 2, 5, 6 and 9) or a homo-amide (2). Compound 4 was the most potent May1 antagonist out of this group of inhibitors, with an IC_{50} of 3.1nM (S 4.10 Fig).

From analysis of the four most effective inhibitors identified in our screen, (compounds 4, 16, 18 and 21), it is clear that a phenylalanine side chain, either unsubstituted (4 and 16) or with a small substituent (18 and 21), is preferred in P1 while a bulkier P1 side chain leads to decreased potency, for example compounds 17, 19 and 20. These results match the P1 substrate preference for phenylalanine that we had predicted for May1 and fit our expectations that bulkier residues such as tryptophan are not well tolerated in this position (Fig 4.3, Fig 4.6B).

Next, we selected the two best *in vitro* antagonists to test their potency in culture relative to pepstatin A by measuring inhibition of May1 and restriction of culture growth using fluorogenic assays and OD_{600} respectively. Wild-type *C. neoformans* was grown in YNB treated with 5µM, 1µM or 0.1µM of compound 4, 16 or pepstatin A and the culture density and May1 activity were measured at saturation. While compound 16 exhibited an *in vitro* IC_{50} comparable to pepstatin A, it was not as effective at inhibiting May1 activity or restricting culture growth (Fig 4.6D, S 4.11 Fig). Curiously, despite having an *in vitro* IC_{50} approximately twice that of compound 16, compound 4 was better at inhibiting culture growth. These results demonstrate that May1 can be targeted by small molecule inhibitors and provide a discovery framework for further inhibitor development.   However, additional medicinal chemistry efforts are necessary for *in vivo* applications. Therefore, subsequent studies investigating the role of May1 in virulence were carried out using deletion strains of this enzyme.

***May1* is required for virulence.** Because *may1Δ* strains exhibit phenotypes in both peptidase

activity assays and growth at low pH, we examined the role of this protein in virulence using an

established mouse inhalation model of Cryptococcal infection [49]. Wild-type cells were mixed

with an equivalent amount of *may1Δ* cells and used to infect mice intranasally (Fig 4.7A).

These experiments were conducted with two independent isolates of *may1Δ* as well as a

negative control known not to affect fungal replication in this assay (*sxi1Δ*) [55]. Ten days after

infection, mouse lungs were harvested and plated for colony forming units (CFUs), at which

point *may1Δ* strains contributed only 22 ± 3% of the colonies recovered from the lungs, a

substantial decrease from the approximately 50% present upon infection. This result reveals

that *may1Δ* cells have a growth defect within a mammalian host because the ratio between

deletion strain and wild type cells was reduced after host infection.

The defect in accumulation of *may1Δ* cells during infection suggested that these strains

would be attenuated for virulence. We directly investigated the virulence of *may1Δ* strains by

performing monotypic infections [49]. Ten mice per group were infected intranasally with wild

type, *may1Δ*-1 or *may1Δ*-2 (Fig 4.7B). Loss of May1 caused significant attenuation of virulence,

with mice infected by *may1Δ*-1 or *may1Δ*-2 exhibiting a mean survival time of 60.1 and 60.7

days respectively, whereas those infected by wild type had a mean survival time of 25 days.

The results from these *in vivo* experiments indicate an important role for May1 during

mammalian infection.

Discussion

In this work, we investigated secreted proteolytic activity in *C. neoformans* var. *grubii* culture

media using an unbiased approach that can detect both endo- and exo-peptidase activity. In

combination with proteomic methods and single gene deletion approaches, this strategy allowed

us to characterize peptidase activity from a global perspective as well as interrogate the roles of

individual enzymes in the *C. neoformans* secretome. By comparing the overlap in peptidase

activity between wild type and these deletion strains, we were able to identify and define the

substrate specificities of a carboxy, aspartyl and metallopeptidase which contribute substantially

to the total peptidase activity profile. Additionally, we delineated the substrate specificity of an

unidentified trypsin-like peptidase activity, an intriguing result given previous reports implicating

secreted serine peptidases in *C. neoformans* pathogenicity [31,34].

Deletion of some peptidase genes, such as the predicted carboxypeptidase D genes

CXD2 and *CXD3*, caused no significant change in secreted proteolytic activity or cellular

phenotype. Instead, it appears that a third carboxypeptidase D paralog *CXD1,* is responsible for

the majority of exopeptidase activity under these conditions. The broad specificity of Cxd1

suggests that one possible role for this enzyme could be in nutrient acquisition by providing *C.*

neoformans with free amino acids from extracellular protein sources [17,18,74].

The serine peptidase deletion strain *prb1Δ* also had a minimal effect on total secreted

peptidase activity; however, a phenotype of reduced melanin production was evident, indicating

function under these conditions (S 4.8 Fig). One possibility is that this gene encodes an

enzyme with very strict substrate specificity, thus its deletion did not have a substantial impact

on total extracellular peptidase activity as measured by the MSP-MS assay (S 4.3 Fig).

Through the application of our global profiling approach to different culture conditions, we

were able to demonstrate that the landscape of secreted peptidase activity shifts in response to

alterations in environment. This result raises the possibility that changes in extracellular

proteolytic activity could be relevant for adaptation. For example, we detected the activity of the

metallopeptidase, Mpr1 only after growth under neutral pH conditions, whereas we find that

May1 is optimally active under acidic conditions. Thus, these enzymes may function in different

settings within the host or within other environments encountered by *C. neoformans*.

Through proteolytic profiling and mutant characterization assays, we identified the

aspartyl peptidase May1 as the dominant endopeptidase at low pH and found that its activity is

required for tolerance to acidic environments. The strongest determinant of May1 substrate specificity was shown to be a preference for cleavage between hydrophobic residues, in particular phenylalanine, leucine and norleucine (Fig 4.3A). Based on these results, we screened a focused panel of aspartyl peptidase inhibitors with similarity to the P1-P1' substrate specificity of May1. Several of these compounds had IC_{50} values in the nanomolar range, whereas the HIV protease inhibitors had relatively poor affinity for May1. Previous reports have shown that some HIV protease inhibitors reduce secreted aspartyl peptidase activity produced by *C. neoformans* [72,73]. The concentrations of inhibitors required to achieve statistically significant inhibition in previous studies were much higher than those used in the experiments reported here although the trends for inhibitor potency match our results [72]. Therefore, it is possible that the inhibition of *C. neoformans* aspartyl peptidase activity seen in previous publications could be explained by the inhibition of May1.

We found that strains lacking May1 are attenuated in both a competition infection assay and in a monotypic infection assay. In microbial culture it is likely that May1 cleaves one or more secreted or cell wall-bound fungal proteins to facilitate low pH tolerance. However, it is possible that during an infection May1 cleaves host proteins and either or both of these proteolytic events impacts virulence. An additional important consideration for defining the role of May1 in *C. neoformans* pathogenicity is the cleavage context within the host. Our pH titration determined that May1 has very low levels of activity above pH 6.5; however few environments of lower pH than this exist within the mammalian host. Therefore, it is possible that residual May1 activity at neutral pH is important, or alternatively that May1 could be relevant for survival in acidic host environments such as the macrophage or dendritic cell phagolysosomes. A third possibility is that a combination of these factors contributes to the attenuated virulence of *may1Δ* strains.

We have identified orthologs of *MAY1* in many other basidiomycetes including the opportunistic pathogens *Trichosporan asahii* and *Cryptococcus gattii,* the latter of which is

capable of infecting immunocompetent individuals [75] [44]. Many pathogenic ascomycetes

also contain *MAY1* orthologs, including *Histoplasma capsulatum*, *Coccidioides immitis* and

Aspergillus species*, although the sequence identity is low [44]. None of the *MAY1* orthologs in

basidiomycetes have been well studied and only one ortholog in an ascomycete has been

examined. This enzyme, from *A.fumigatus,* encodes a protein secreted during infection of the

virulence model *Galleria mellonella* [76]. The hypovirulent phenotype observed in *C.*

neoformans may1Δ strains and the identification of May1 orthologs in other fungal pathogens

raises the possibility that this peptidase family displays a conserved virulence function and

suggests that the roles of these orthologs are important to investigate.

Due to the role of secreted peptidases in virulence, inhibition of these enzymes is a

potential alternative approach to current anti-fungal therapy. However, small molecule drug

development requires a thorough understanding of the target enzyme as well as the

surrounding peptidase milieu [77-81]. The results described in this report lay the groundwork for

investigating the functions of *C. neoformans* secreted peptidases and the use of inhibitors to

modulate their activity*. The substrates and inhibitors presented here may also be of value for

interrogating related fungal peptidases. Furthermore, our approach to investigating secreted

peptidases through integration of activity profiling, proteomics, and genomics strategies is

broadly applicable to other genetically tractable pathogens.

Acknowledgements

References

1. Armstrong-James D, Meintjes G, Brown GD (2014) A neglected epidemic: fungal infections in HIV/AIDS. Trends Microbiol 22: 120-127.
2. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, et al. (2009) Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. AIDS 23: 525-530.
3. Krysan DJ (2015) Toward improved anti-cryptococcal drugs: Novel molecules and repurposed drugs. Fungal Genet Biol 78: 93-98.
4. Brown GD, Denning DW, Gow NA, Levitz SM, Netea MG, et al. (2012) Hidden killers: human fungal infections. Sci Transl Med 4: 165rv113.
5. Roemer T, Krysan DJ (2014) Antifungal drug development: challenges, unmet clinical needs, and new approaches. Cold Spring Harb Perspect Med 4.
6. Ruiz-Perez F, Nataro JP (2014) Bacterial serine proteases secreted by the autotransporter pathway: classification, specificity, and role in virulence. Cell Mol Life Sci 71: 745-770.
7. Kolar SL, Ibarra JA, Rivera FE, Mootz JM, Davenport JE, et al. (2013) Extracellular proteases are key mediators of Staphylococcus aureus virulence via the global modulation of virulence-determinant stability. Microbiologyopen 2: 18-34.
8. McKerrow JH, Sun E, Rosenthal PJ, Bouvier J (1993) The proteases and pathogenicity of parasitic protozoa. Annu Rev Microbiol 47: 821-853.
9. Klemba M, Goldberg DE (2002) Biological roles of proteases in parasitic protozoa. Annu Rev Biochem 71: 275-305.
10. Zhang YZ, Ran LY, Li CY, Chen XL (2015) Diversity, Structures, and Collagen-Degrading Mechanisms of Bacterial Collagenolytic Proteases. Appl Environ Microbiol 81: 6098-6107.
11. Yang Y, Wen Y, Cai YN, Vallee I, Boireau P, et al. (2015) Serine proteases of parasitic helminths. Korean J Parasitol 53: 1-11.
12. Elbaz T, El-Kassas M, Esmat G (2015) New era for management of chronic hepatitis C virus using direct antiviral agents: A review. J Adv Res 6: 301-310.
13. Konvalinka J, Krausslich HG, Muller B (2015) Retroviral proteases and their roles in virion maturation. Virology 479-480: 403-417.
14. Kaman WE, Hays JP, Endtz HP, Bikker FJ (2014) Bacterial proteases: targets for diagnostics and therapy. Eur J Clin Microbiol Infect Dis 33: 1081-1087.
15. Dixit AK, Dixit P, Sharma RL (2008) Immunodiagnostic/protective role of cathepsin L cysteine proteinases secreted by Fasciola species. Vet Parasitol 154: 177-184.
16. Aoki W, Kitahara N, Fujita A, Shibasaki S, Morisaka H, et al. (2013) Detection of Candida albicans by using a designed fluorescence-quenched peptide. J Biosci Bioeng 116: 573-575.
17. Yike I (2011) Fungal proteases and their pathophysiological effects. Mycopathologia 171: 299-323.
18. Girard V, Dieryckx C, Job C, Job D (2013) Secretomes: the fungal strike force. Proteomics 13: 597-608.
19. Alby K, Schaefer D, Bennett RJ (2009) Homothallic and heterothallic mating in the opportunistic pathogen Candida albicans. Nature 460: 890-893.
20. Naglik J, Albrecht A, Bader O, Hube B (2004) Candida albicans proteinases and host/pathogen interactions. Cell Microbiol 6: 915-926.
21. Rambach G, Dum D, Mohsenipour I, Hagleitner M, Wurzner R, et al. (2010) Secretion of a fungal protease represents a complement evasion mechanism in cerebral aspergillosis. Mol Immunol 47: 1438-1449.
22. Monod M, Capoccia S, Lechenne B, Zaugg C, Holdom M, et al. (2002) Secreted proteases from pathogenic fungi. Int J Med Microbiol 292: 405-419.

23. Monod M, Borg-von ZM (2002) Secreted aspartic proteases as virulence factors of Candida species. Biol Chem 383: 1087-1093.
24. Baldo A, Monod M, Mathy A, Cambier L, Bagut ET, et al. (2012) Mechanisms of skin adherence and invasion by dermatophytes. Mycoses 55: 218-223.
25. O'Donoghue AJ, Knudsen GM, Beekman C, Perry JA, Johnson AD, et al. (2015) Destructin-1 is a collagen-degrading endopeptidase secreted by Pseudogymnoascus destructans, the causative agent of white-nose syndrome. Proc Natl Acad Sci U S A 112: 7478-7483.
26. Sriranganadane D, Waridel P, Salamin K, Feuermann M, Mignon B, et al. (2011) Identification of novel secreted proteases during extracellular proteolysis by dermatophytes at acidic pH. Proteomics 11: 4422-4433.
27. Eigenheer RA, Jin Lee Y, Blumwald E, Phinney BS, Gelli A (2007) Extracellular glycosylphosphatidylinositol-anchored mannoproteins and proteases of Cryptococcus neoformans. FEMS Yeast Res 7: 499-510.
28. Vu K, Tham R, Uhrig JP, Thompson GR, 3rd, Na Pombejra S, et al. (2014) Invasion of the central nervous system by Cryptococcus neoformans requires a secreted fungal metalloprotease. MBio 5: e01101-01114.
29. Campbell LT, Simonin AR, Chen C, Ferdous J, Padula MP, et al. (2015) Cryptococcus strains with different pathogenic potentials have diverse protein secretomes. Eukaryot Cell 14: 554-563.
30. Casadevall A, Steenbergen JN, Nosanchuk JD (2003) 'Ready made' virulence and 'dual use' virulence factors in pathogenic environmental fungi--the Cryptococcus neoformans paradigm. Curr Opin Microbiol 6: 332-337.
31. Yoo Ji J, Lee YS, Song CY, Kim BS (2004) Purification and characterization of a 43-kilodalton extracellular serine proteinase from Cryptococcus neoformans. J Clin Microbiol 42: 722-726.
32. Pinti M, Orsi CF, Gibellini L, Esposito R, Cossarizza A, et al. (2007) Identification and characterization of an aspartyl protease from Cryptococcus neoformans. FEBS Lett 581: 3882-3886.
33. Aoki S, Ito-Kuwa S, Nakamura K, Kato J, Ninomiya K, et al. (1994) Extracellular proteolytic activity of Cryptococcus neoformans. Mycopathologia 128: 143-150.
34. Rodrigues ML, dos Reis FC, Puccia R, Travassos LR, Alviano CS (2003) Cleavage of human fibronectin and other basement membrane-associated proteins by a Cryptococcus neoformans serine proteinase. Microb Pathog 34: 65-71.
35. Ruma-Haynes P, Brownlee AG, Sorrell TC (2000) A rapid method for detecting extracellular proteinase activity in Cryptococcus neoformans and a survey of 63 isolates. J Med Microbiol 49: 733-737.
36. Chan MY, Tay ST (2010) Enzymatic characterisation of clinical isolates of Cryptococcus neoformans, Cryptococcus gattii and other environmental Cryptococcus spp. Mycoses 53: 26-31.
37. Vidotto V, Koga-Ito CY, Canella D, Sinicco A, Di Perri G, et al. (2000) Extracellular activity in Cryptococcus neoformans strains isolated from AIDS patients and from environmental sources. Rev Iberoam Micol 17: 14-19.
38. Vidotto V, Melhem M, Pukinskas S, Aoki S, Carrara C, et al. (2005) Extracellular enzymatic activity and serotype of Cryptococcus neoformans strains isolated from AIDS patients in Brazil. Rev Iberoam Micol 22: 29-33.
39. O'Donoghue AJ, Eroy-Reveles AA, Knudsen GM, Ingram J, Zhou M, et al. (2012) Global identification of peptidase specificity by multiplex substrate profiling. Nat Methods 9: 1095-1100.
40. Chalkley RJ, Baker PR, Medzihradszky KF, Lynn AJ, Burlingame AL (2008) In-depth analysis of tandem mass spectrometry data from disparate instrument types. Mol Cell Proteomics 7: 2386-2398.

41. Colaert N, Helsens K, Martens L, Vandekerckhove J, Gevaert K (2009) Improved visualization of protein consensus sequences by iceLogo. Nat Methods 6: 786-787.
42. Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8: 785-786.
43. Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S (2004) Feature-based prediction of non-classical and leaderless protein secretion. Protein Eng Des Sel 17: 349-356.
44. Stajich JE, Harris T, Brunk BP, Brestelli J, Fischer S, et al. (2012) FungiDB: an integrated functional genomics database for fungi. Nucleic Acids Res 40: D675-681.
45. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, et al. (2012) ExPASy: SIB bioinformatics resource portal. Nucleic Acids Res 40: W597-603.
46. R. Gupta EJaSB (2004) Prediction of N-glycosylation sites in human proteins. Available: http://www.cbs.dtu.dk/services/NetNGlyc/.
47. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, et al. (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBO J 32: 1478-1488.
48. Institute B (2008) Cryptococcus Sequencing initiative, Broad Institute (broadinstitute.org).
49. Chun CD, Madhani HD (2010) Applying genetics and molecular biology to the study of the human pathogen Cryptococcus neoformans. Methods Enzymol 470: 797-831.
50. Inglis DO, Skrzypek MS, Liaw E, Moktali V, Sherlock G, et al. (2014) Literature-based gene curation and proposed genetic nomenclature for cryptococcus. Eukaryot Cell 13: 878-883.
51. Idnurm A, Heitman J (2005) Photosensing fungi: phytochrome in the spotlight. Curr Biol 15: R829-832.
52. Liu OW, Chun CD, Chow ED, Chen C, Madhani HD, et al. (2008) Systematic genetic analysis of virulence in the human fungal pathogen Cryptococcus neoformans. Cell 135: 174-188.
53. Roth V (2006) Doubling Time. Available: http://doubling-time.com/compute.php.
54. Furfine ES (2001) HIV protease assays. Curr Protoc Pharmacol Chapter 3: Unit3 2.
55. Hull CM, Cox GM, Heitman J (2004) The alpha-specific cell identity factor Sxi1alpha is not required for virulence of Cryptococcus neoformans. Infect Immun 72: 3643-3645.
56. Yang JS, Nam HJ, Seo M, Han SK, Choi Y, et al. (2011) OASIS: online application for the survival analysis of lifespan assays performed in aging research. PLoS One 6: e23525.
57. O'Brien TC, Mackey ZB, Fetter RD, Choe Y, O'Donoghue AJ, et al. (2008) A parasite cysteine protease is key to host protein degradation and iron acquisition. J Biol Chem 283: 28934-28943.
58. Small JL, O'Donoghue AJ, Boritsch EC, Tsodikov OV, Knudsen GM, et al. (2013) Substrate specificity of MarP, a periplasmic protease required for resistance to acid and oxidative stress in Mycobacterium tuberculosis. J Biol Chem 288: 12489-12499.
59. O'Donoghue AJ, Mahon CS, Goetz DH, O'Malley JM, Gallagher DM, et al. (2008) Inhibition of a secreted glutamic peptidase prevents growth of the fungus Talaromyces emersonii. J Biol Chem 283: 29186-29195.
60. Rodrigues ML, Nakayasu ES, Oliveira DL, Nimrichter L, Nosanchuk JD, et al. (2008) Extracellular vesicles produced by Cryptococcus neoformans contain protein components associated with virulence. Eukaryot Cell 7: 58-67.
61. Levitz SM, Specht CA (2006) The molecular basis for the immunogenicity of Cryptococcus neoformans mannoproteins. FEMS Yeast Res 6: 513-524.
62. Rawlings ND, Waller M, Barrett AJ, Bateman A (2014) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res 42: D503-509.
63. Letunic I, Doerks T, Bork P (2015) SMART: recent updates, new developments and status in 2015. Nucleic Acids Res 43: D257-260.

64. Sansen S, De Ranter CJ, Gebruers K, Brijs K, Courtin CM, et al. (2004) Structural basis for inhibition of Aspergillus niger xylanase by triticum aestivum xylanase inhibitor-I. J Biol Chem 279: 36022-36028.
65. Dunn BM (2001) Overview of pepsin-like aspartic peptidases. Curr Protoc Protein Sci Chapter 21: Unit 21 23.
66. Nosanchuk JD, Casadevall A (2003) Budding of melanized Cryptococcus neoformans in the presence or absence of L-dopa. Microbiology 149: 1945-1951.
67. Bose I, Reese AJ, Ory JJ, Janbon G, Doering TL (2003) A yeast under cover: the capsule of Cryptococcus neoformans. Eukaryot Cell 2: 655-663.
68. Rinnova M, Hradilek M, Barinka C, Weber J, Soucek M, et al. (2000) A picomolar inhibitor of resistant strains of human immunodeficiency virus protease identified by a combinatorial approach. Arch Biochem Biophys 382: 22-30.
69. Skalova T, Hasek J, Dohnalek J, Petrokova H, Buchtelova E, et al. (2003) An ethylenamine inhibitor binds tightly to both wild type and mutant HIV-1 proteases. Structure and energy study. J Med Chem 46: 1636-1644.
70. Skalova T, Dohnalek J, Duskova J, Petrokova H, Hradilek M, et al. (2006) HIV-1 protease mutations and inhibitor modifications monitored on a series of complexes. Structural basis for the effect of the A71V mutation on the active site. J Med Chem 49: 5777-5784.
71. Petrokova H, Duskova J, Dohnalek J, Skalova T, Vondrackova-Buchtelova E, et al. (2004) Role of hydroxyl group and R/S configuration of isostere in binding properties of HIV-1 protease inhibitors. Eur J Biochem 271: 4451-4461.
72. Sidrim JJ, Perdigao-Neto LV, Cordeiro RA, Brilhante RS, Leite JJ, et al. (2012) Viral protease inhibitors affect the production of virulence factors in Cryptococcus neoformans. Can J Microbiol 58: 932-936.
73. Monari C, Pericolini E, Bistoni G, Cenci E, Bistoni F, et al. (2005) Influence of indinavir on virulence and growth of Cryptococcus neoformans. J Infect Dis 191: 307-311.
74. Naglik JR, Challacombe SJ, Hube B (2003) Candida albicans secreted aspartyl proteinases in virulence and pathogenesis. Microbiol Mol Biol Rev 67: 400-428, table of contents.
75. Byrnes EJ, 3rd, Li W, Lewit Y, Ma H, Voelz K, et al. (2010) Emergence and pathogenicity of highly virulent Cryptococcus gattii genotypes in the northwest United States. PLoS Pathog 6: e1000850.
76. Vickers I, Reeves EP, Kavanagh KA, Doyle S (2007) Isolation, activity and immunological characterisation of a secreted aspartic protease, CtsD, from Aspergillus fumigatus. Protein Expr Purif 53: 216-224.
77. Stewart K, Abad-Zapatero C (2001) Candida proteases and their inhibition: prospects for antifungal therapy. Curr Med Chem 8: 941-948.
78. Pozio E, Morales MA (2005) The impact of HIV-protease inhibitors on opportunistic parasites. Trends Parasitol 21: 58-63.
79. Dos Santos AL (2011) Protease expression by microorganisms and its relevance to crucial physiological/pathological events. World J Biol Chem 2: 48-58.
80. Mehra T, Koberle M, Braunsdorf C, Mailander-Sanchez D, Borelli C, et al. (2012) Alternative approaches to antifungal therapies. Exp Dermatol 21: 778-782.
81. Olsen I, Potempa J (2014) Strategies for the inhibition of gingipains for the potential treatment of periodontitis and associated systemic diseases. J Oral Microbiol 6.
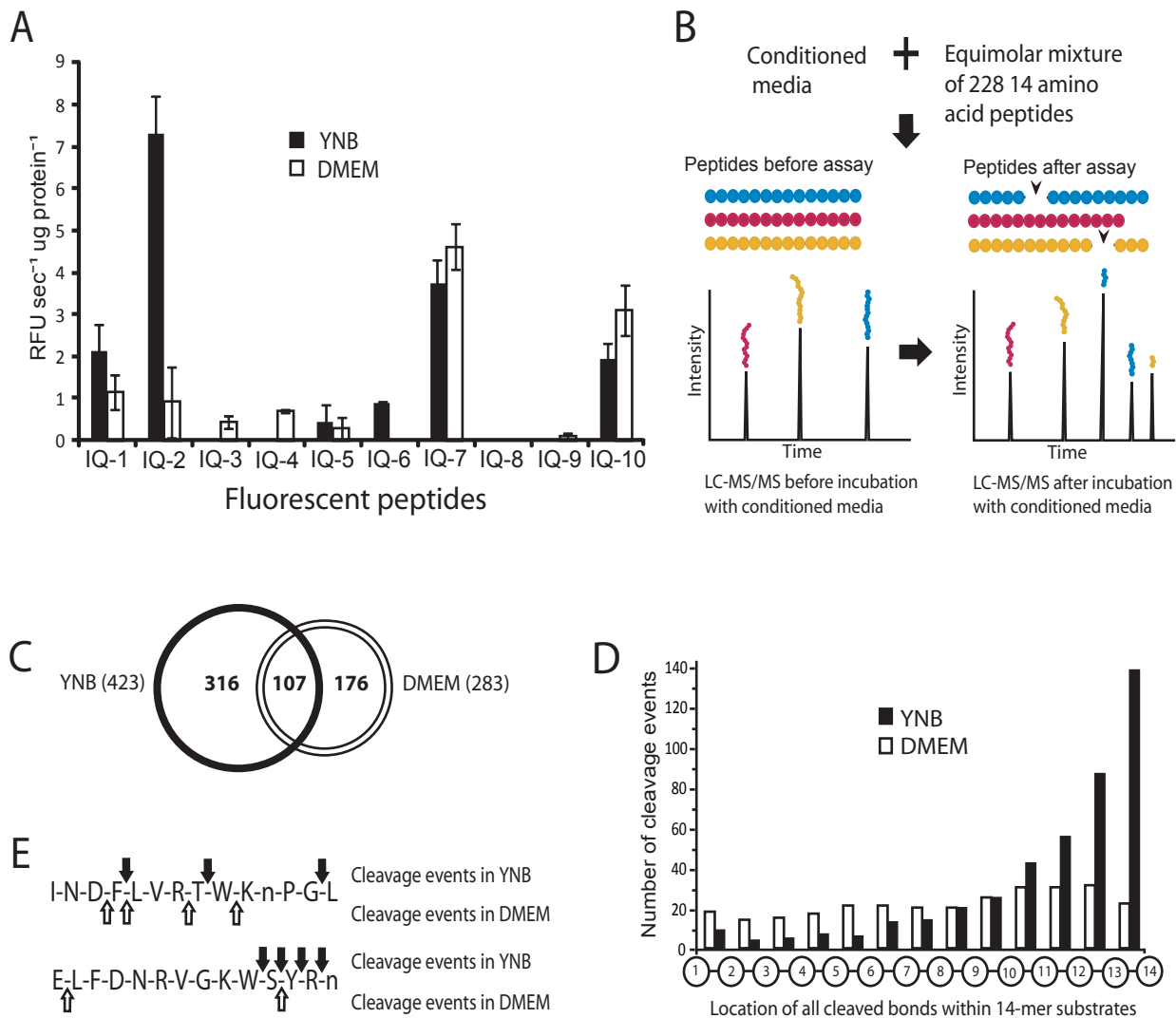
Fig 4.1 Profiling of *C. neoformans* **conditioned media reveals abundant secreted peptidase activity with environment-specific regulation**. (A) Profiling of secreted peptidase activity present in YNB or DMEM media conditioned by wild-type cells using a panel of internally quenched (IQ) fluorescent peptides. (B) Schematic of Multiplex Substrate Profiling by Mass Spectrometry (MSP-MS). Conditioned media is combined with a 228-member peptide library and mass spectrometry analysis is run before and after incubation to identify cleavage events. Norleucine replaces methionine in the MSP-MS library and is indicated by an "n" in representations of the peptides. (C) Several hundred cleavage events were detected in both YNB and DMEM conditioned media profiled by MSP-MS. (D) Positional profiling of all cleaved bonds detected by MSP-MS in either media type. (E) Two representative examples of peptides cleaved in the MSP-MS assay by both media conditions. Arrows indicate the location of cleavage sites.
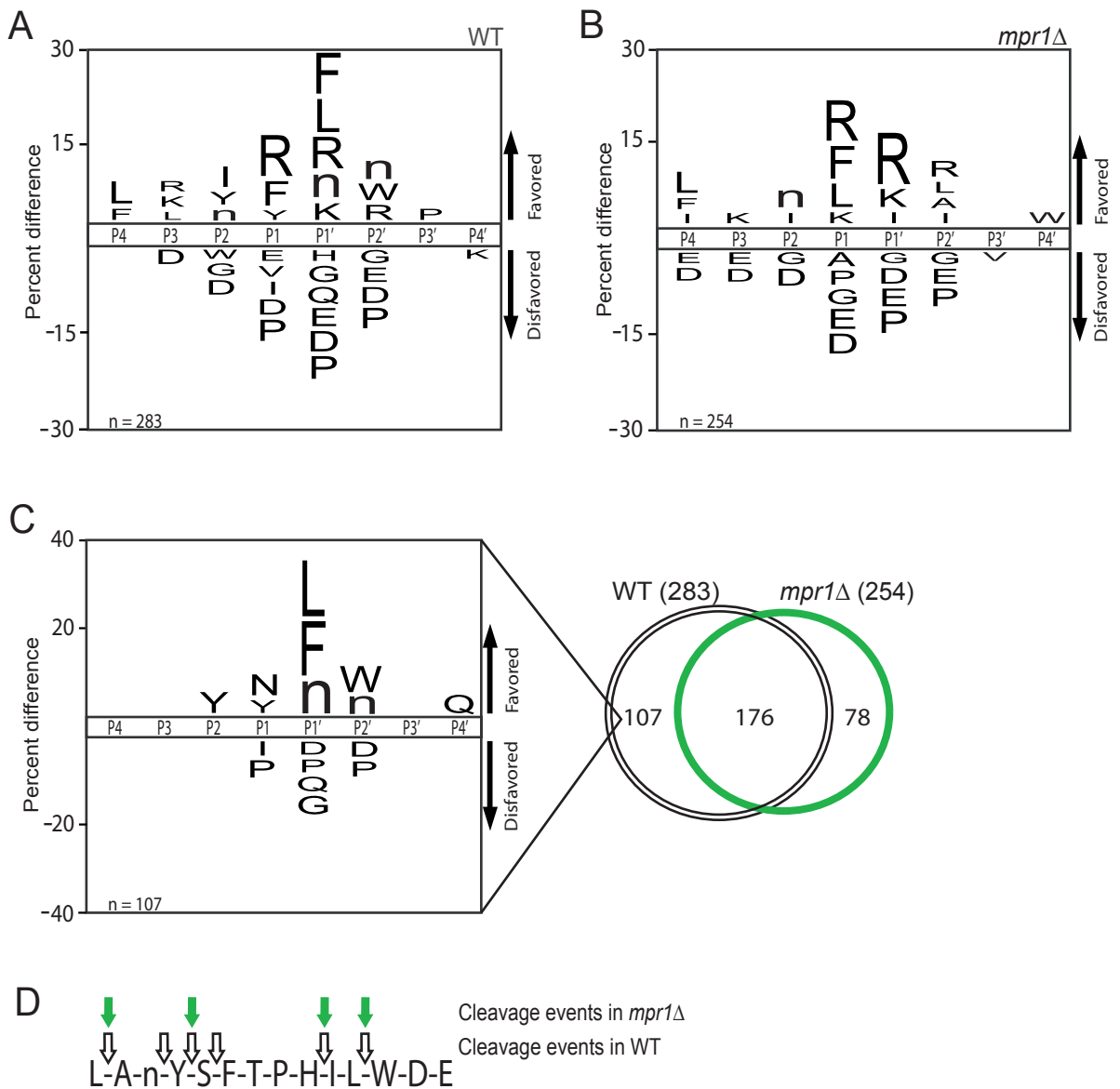
Fig 4.2 DMEM conditioned media contains a metallopeptidase and trypsin-like endopeptidase activity. (A) The peptidase substrate specificity profile of DMEM media conditioned by wild type. Residues are significantly favored or disfavored as determined by the frequency of detection in substrates versus the frequency of the residue in the peptide library, $p = 0.05$. (B) Substrate specificity profile of peptidase activity from mpr1Δ conditioned media, $p = 0.05$. (C) Peptidase substrate specificity profile constructed from cleavage events detected in wild type but not mpr1Δ, $p = 0.05$. (D) A representative peptide cleaved by peptidases in both wild type and mpr1Δ conditioned media.
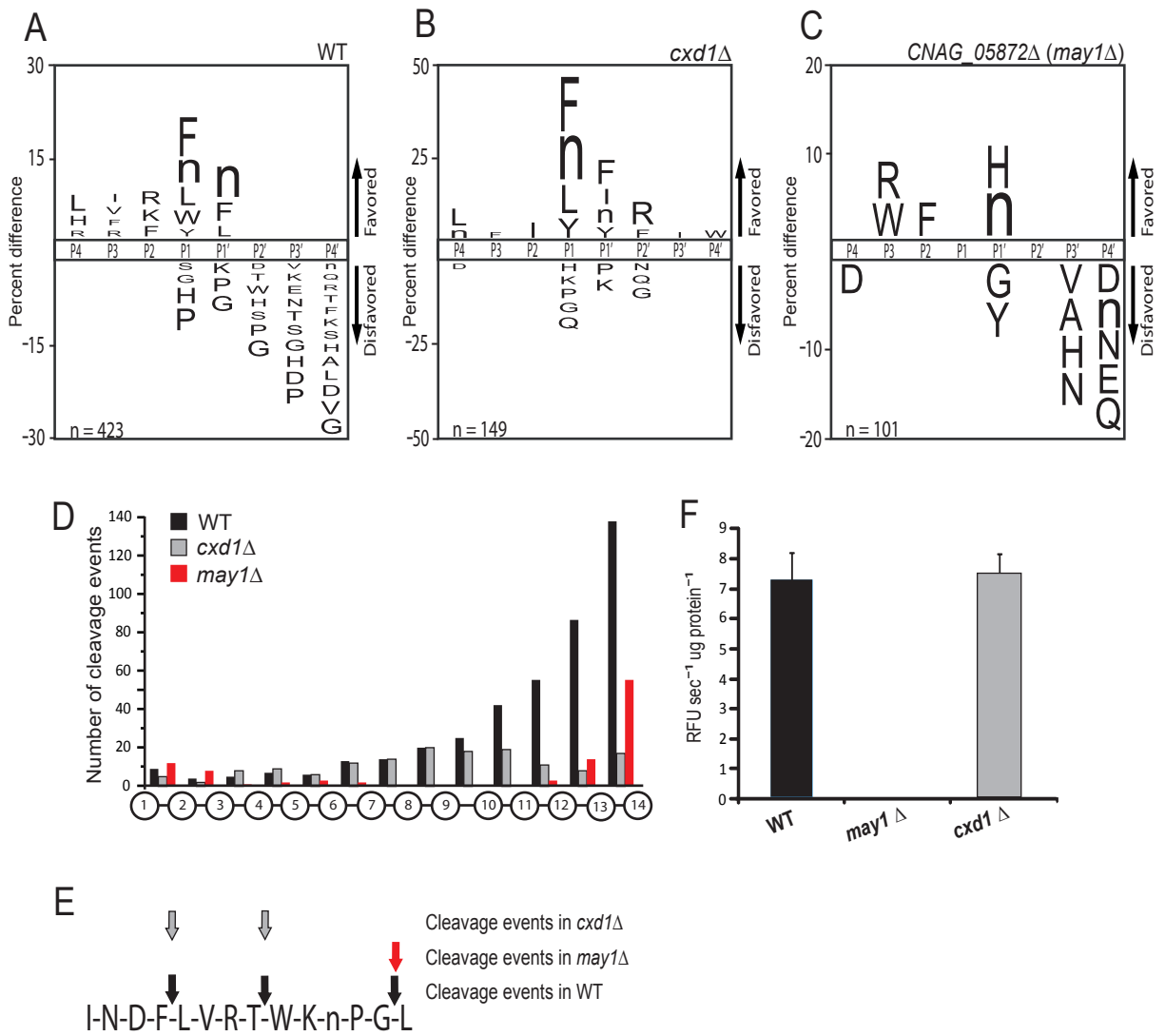
Fig 4.3 May1 and Cxd1 are the major proteolytic activities in YNB conditioned media.
(A) The substrate specificity profile of YNB media conditioned by wild type, p = 0.05. (B) The substrate specificity profile of the carboxypeptidase D deletion strain cxd1Δ, p = 0.05. (C) The substrate specificity profile of the aspartyl peptidase deletion strain may1Δ, p = 0.05. (D) Positional profiling of all bonds cleaved within the tetradecapeptides of the MSP-MS library. (E) A representative example of a peptide in the MSP-MS library cleaved by wild type and both deletion strains. (F) Deletion of May1 abolishes endopeptidase activity against IQ-2.
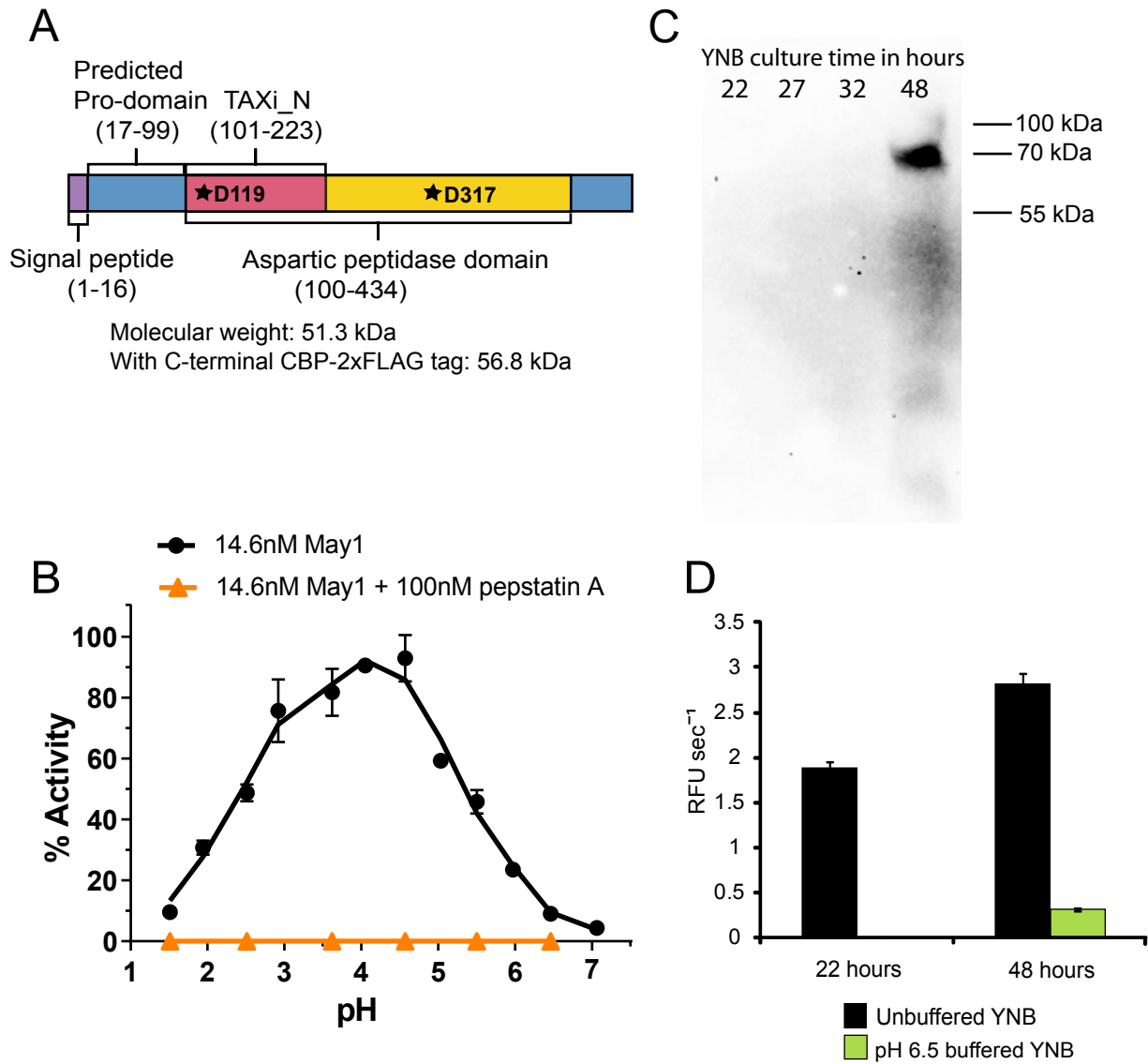
Fig 4.4 May1 is a member of the pepsin-like aspartyl peptidase family with optimal expression and activity at acidic pH. (A) The domain architecture of May1. The catalytic aspartic acids are indicated by stars. (B) pH titration of May1 activity using IQ-2. Averages and standard deviation (S.D.) of triplicates are shown. (C) Immunoblot detection of May1 tagged with a CBP-2xFLAG tag in supernatants after culturing in YNB for different lengths of time. (D) Activity against IQ-2 in conditioned media from wild type C. neoformans grown in unbuffered or pH 6.5 buffered YNB. All samples were adjusted to pH 4.5 prior to being assayed.
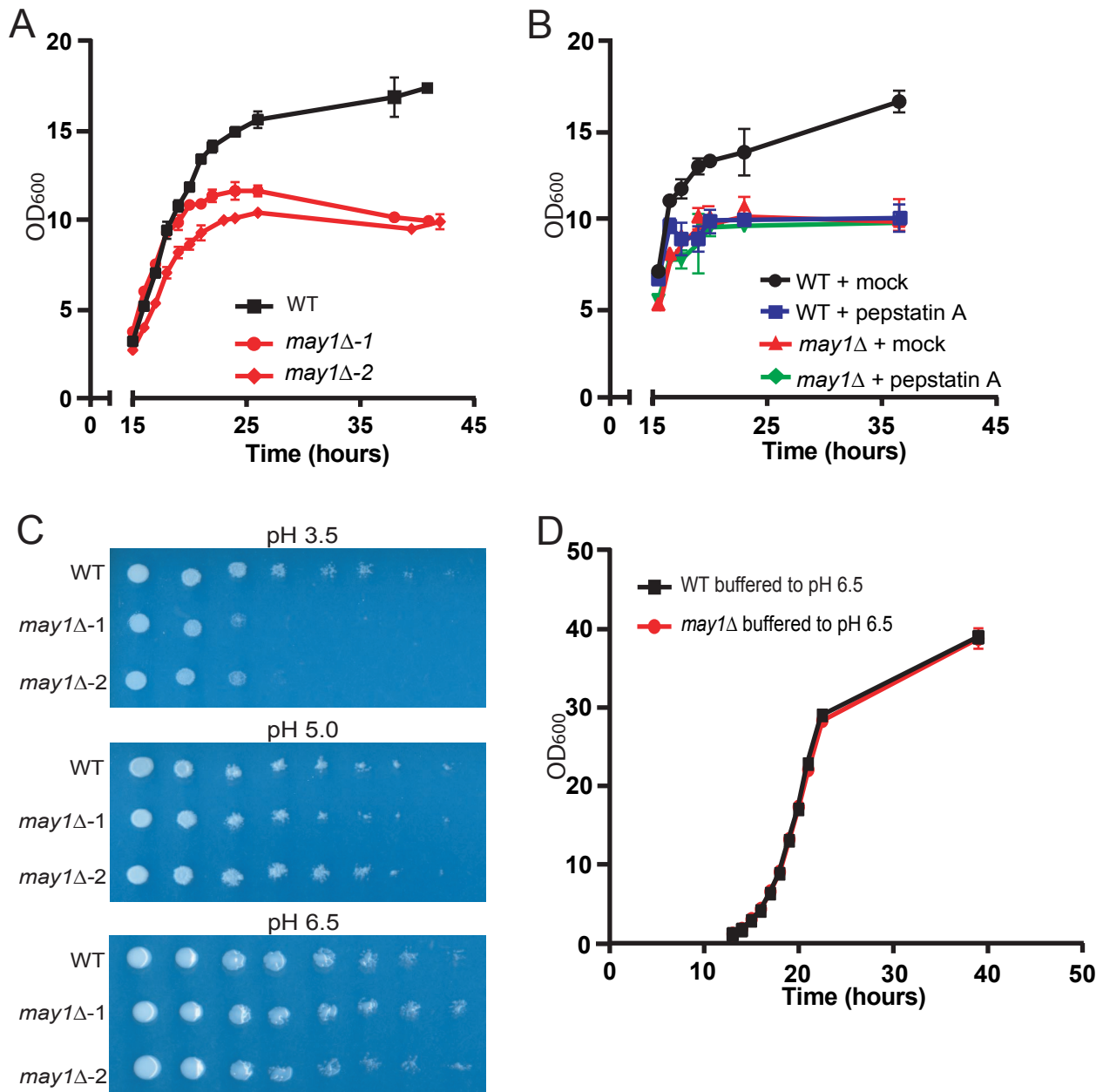
Fig 4.5 May1 activity is required for wild-type growth at acidic pH. (A) Culture density was recorded over time by measuring OD600 of cultures grown in YNB. The average density and S.D. of triplicates are shown. (B) Culture density of wild-type C. neoformans grown in YNB containing 2μM pepstatin A was recorded. Mock indicates DMSO treatment. The average and S.D. of cultures grown in triplicate are shown. (C) Mutant characterization assays were conducted using YNB agar plates. Cultures were adjusted to an OD600 of 5 and then spotted in 10-fold serial dilutions on plates. (D) Full rescue of may1Δ saturation density in YNB was observed after buffering to pH 6.5. Average density and S.D. of triplicates are shown.
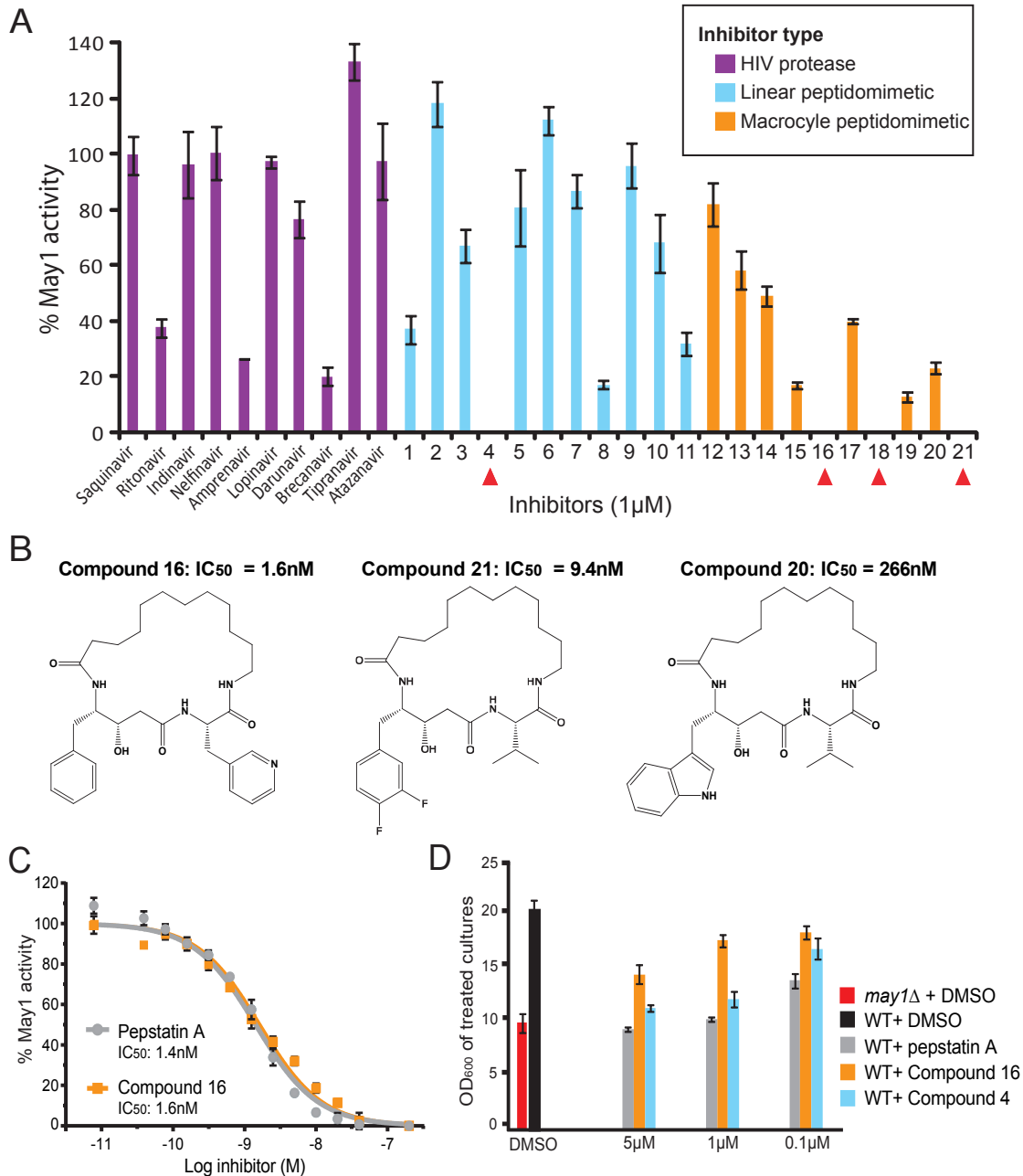
Fig 4.6. A screen of aspartyl peptidase inhibitors uncovers compounds antagonistic to May1. (A) Three groups of compounds were screened for inhibition of May1 activity using IQ-2. Compounds completely inhibiting May1 at 1µM are denoted with red triangles. Averages and S.D. of triplicates are shown. (B) The structures of three macrocyclic compounds screened for inhibition of May1. (C) The IC50 for the most potent May1 inhibitor, (compound 16) was found to be 1.6nM, while peptstatin A had an IC50 of 1.4nM. The average and S.D. of measurements in triplicate are shown. (D) Density at saturation (after 48 hours of growth) is shown for YNB cultures of wild type C. neoformans treated with May1 inhibitors. Average values and S.D. of triplicates are shown.
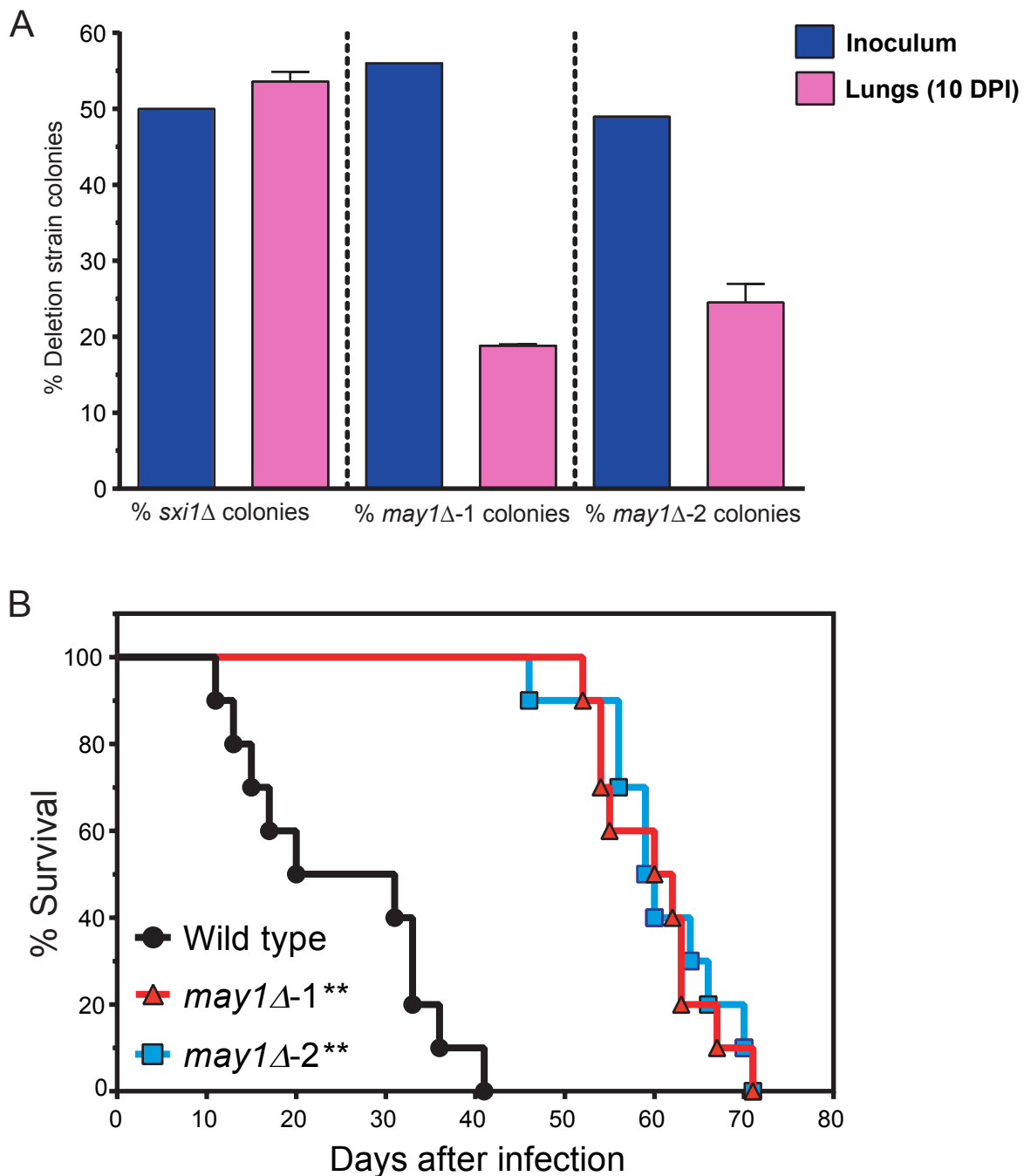
Fig 4.7 May1 is required for virulence in a mouse inhalation model of infection. (A) Three mice per group were infected with a 1:1 ratio of wild type to mutant cells using the mutant strains may1Δ-1, may1Δ-2 or sxi1Δ. The 1:1 ratio of wild type to mutant strain was confirmed by plating the inoculums on plates containing nourseothricin as a selection agent. Ten days after infection, lungs were harvested and plated to determine the ratio of wild type to deletion strain. Average values and S.D. are shown. DPI stands for days post infection. (B) Ten mice per group were infected with wild type, may1Δ-1 or may1Δ-2 cells. Significance was determined by a log-rank test, ** indicates p < 10-5.

Table 1. Peptidase deletion strains generated in this study. Gene names were determined where possible by following the recommended naming guidelines for C. neoformans [50]. NatR is nourseothricin resistance. An asterisk indicates the observation of a phenotype in subsequent mutant characterization studies (S7-8 Fig). Evidence for activity in YNB or DMEM conditioned media was determined in subsequent experiments analyzing proteolytic activity in media conditioned by the peptidase deletion strains (Fig 2-3, S3-4 Fig).

Genotype	Name	Peptidase Type	Identification method	Evidence for activity		Citation
				in YNB	in DMEM	
CNAG_05973Δ::Nat^R	SCX1	Serine carboxypeptidase	Literature			[27]
CNAG_06640Δ::Nat^R	PRC1	Serine carboxypeptidase	This study			This study
CNAG_00919Δ::Nat^R	CXD1	Carboxypeptidase D	This study	**+**		[27]
CNAG_01040Δ::Nat^R	CXD2	Carboxypeptidase D	Predicted homolog			[44]
CNAG_02966Δ::Nat^R	CXD3	Carboxypeptidase D	Predicted homolog			[44]
CNAG_00150Δ::Nat^R	-	Serine endopeptidase	Literature			[27]
CNAG_04625Δ::Nat^R	PRB1*	Serine endopeptidase	This study			[27]
CNAG_00581Δ::Nat^R	PEP4*	Aspartyl endopeptidase	This study			[27]
CNAG_05872Δ::Nat^R	MAY1*	Aspartyl endopeptidase	This study	**+**		This study
CNAG_04735Δ::Nat^R	MPR1	Metallo endopeptidase	This study		**+**	[27]
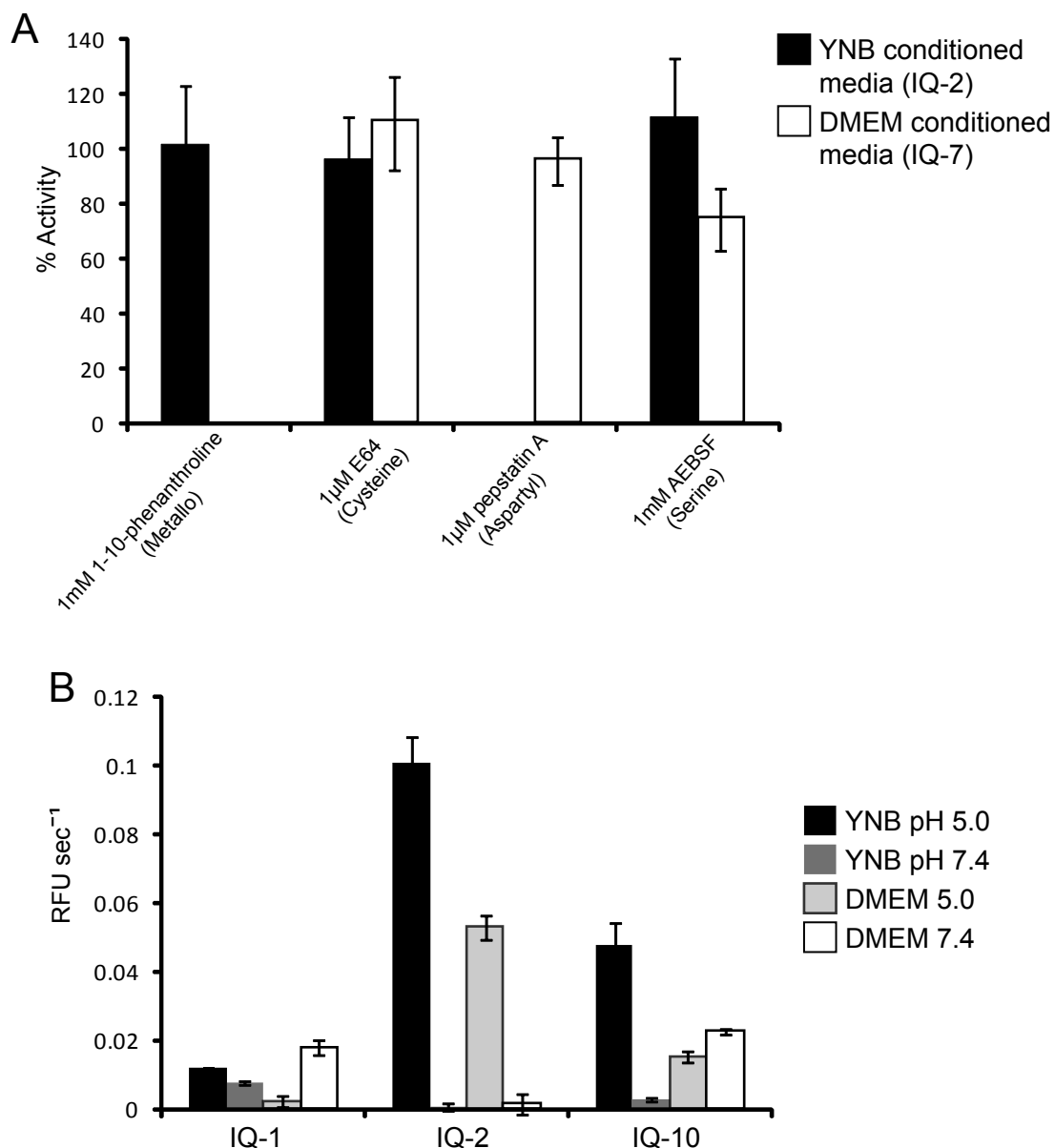
S 4.1 Fig. Fluorogenic assays indicate the endopeptidase class present in conditioned media and the optimal pH for detection of activity. (A) The impact of class specific peptidase inhibitors on peptidase activity. The class of enzyme inhibited by each compound is indicated in parenthesis. Averages and S.D. are shown for triplicates. The substrates cleaved most efficiently by peptidases in each media condition are shown (IQ-2 and IQ-7 for YNB and DMEM, respectively). Cleavage of all other IQ substrates by conditioned YNB media were also inhibited by pepstatin A, while cleavage of all other IQ substrates by DMEM conditioned media were also inhibited by 1-10-phenanthroline. (B) Screen of the affect of pH on proteolytic activity in YNB and DMEM supernatants. Three efficiently cleaved IQ substrates were chosen for this analysis. This assay was conducted on a SpectraMax Gemini plate reader (Molecular Devices) although conditions were otherwise equivalent. Averages and S.D. are shown for triplicates.

A

Peptidase Type	Gene	Uniprot
Serine carboxypeptidase	*CNAG_06640*	J9VTZ8
Carboxypeptidase D	*CNAG_00919*	J9VTZ8
Serine endopeptidase	*CNAG_04625*	J9W0L4
Aspartyl endopeptidase	*CNAG_00581*	J9VH59
Aspartyl endopeptidase	*CNAG_05872*	J9VS02
Metallo endopeptidase	*CNAG_04735*	J9VXZ9
GPI-anchored aspartyl endopeptidase	*CNAG_04380*	J9VUL8
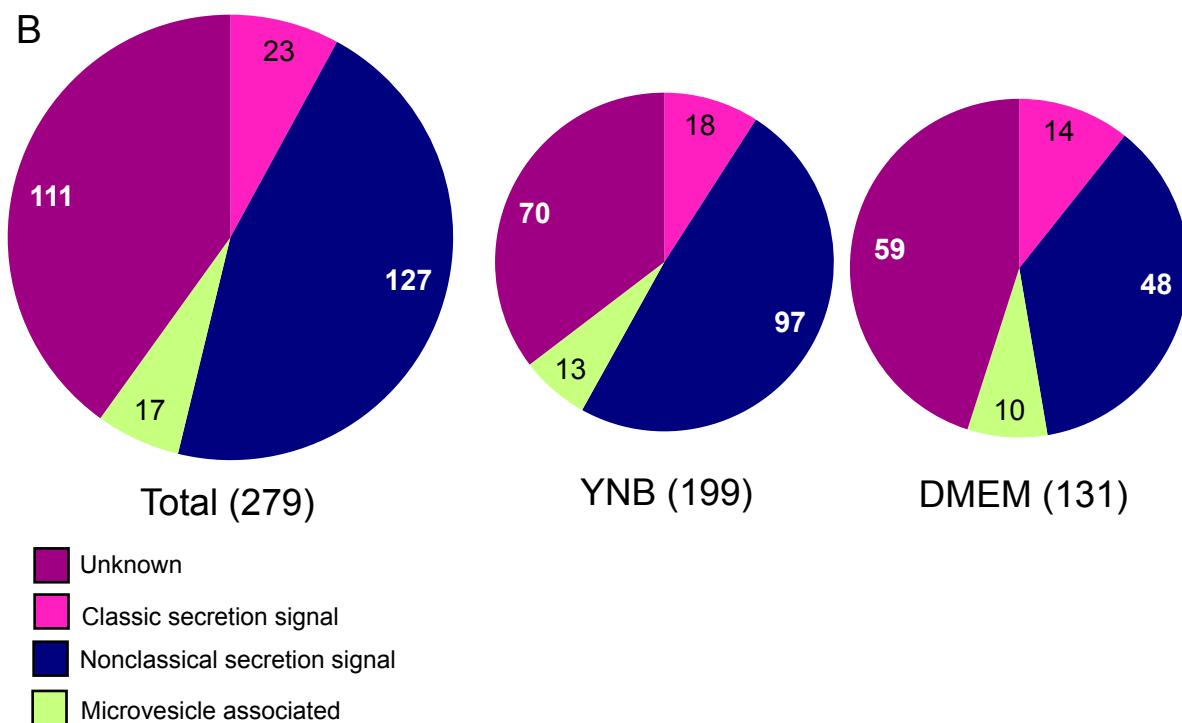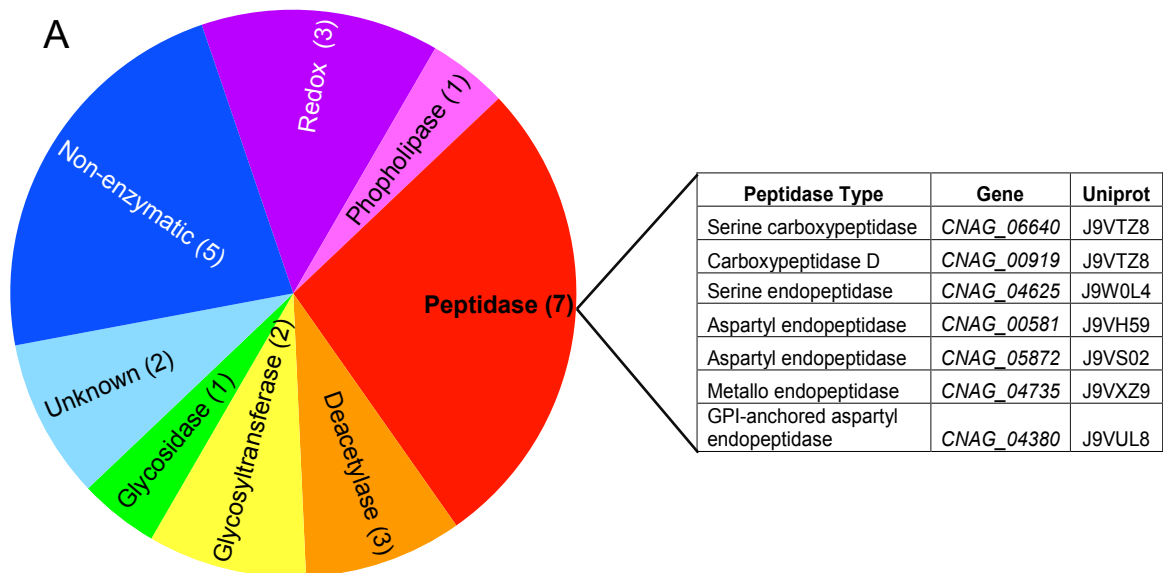
S 4.2 Fig. Functional categorization and analysis of secretion mechanism for proteomics results. (A) Functional categorization of all 23 proteins predicted to have a secretion signal. Functions were determined for unannotated proteins by the closest annotated protein after conducting a Blastp search. (B) Analysis of predicted secretion method for all proteins detected in YNB or DMEM conditioned media by proteomics.
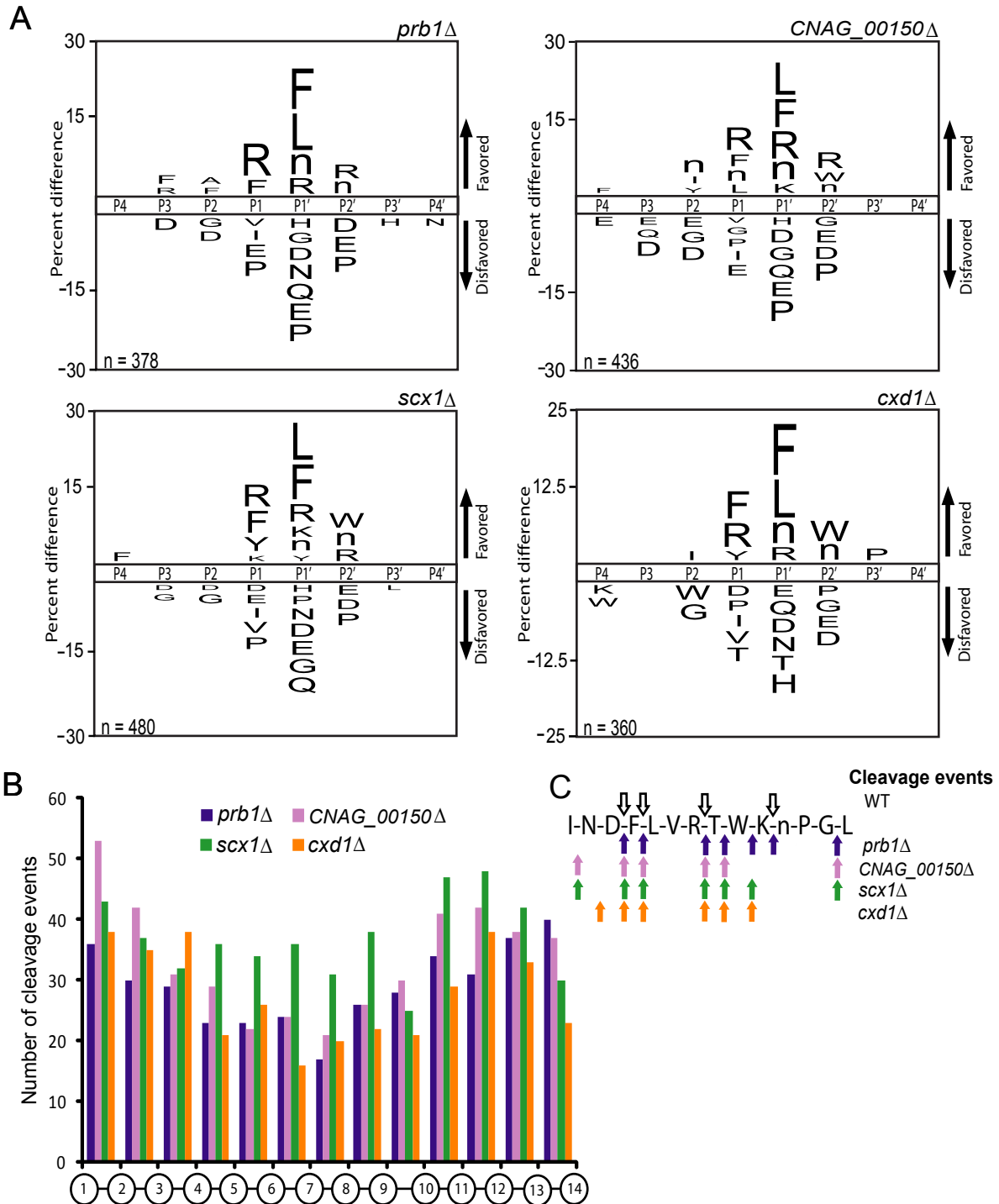
S 4.3 Fig. MSP-MS analysis of secreted peptidase activity in *prb1Δ, CNAG_00150Δ, scx1Δ* and *cxd1Δ* strains cultured in DMEM. (A) Substrate specificity profiles of the serine peptidase deletion strains *prb1Δ* and *CNAG_00150 Δ* and the carboxypeptidase deletion strains *scx1Δ* and *cxd1Δ* grown in DMEM, (p = 0.05). (B) Positional analysis of the bonds cleaved in the four deletion strains. (C) Representative example of a peptide cleaved by peptidases in media conditioned by each of the four deletion strains.
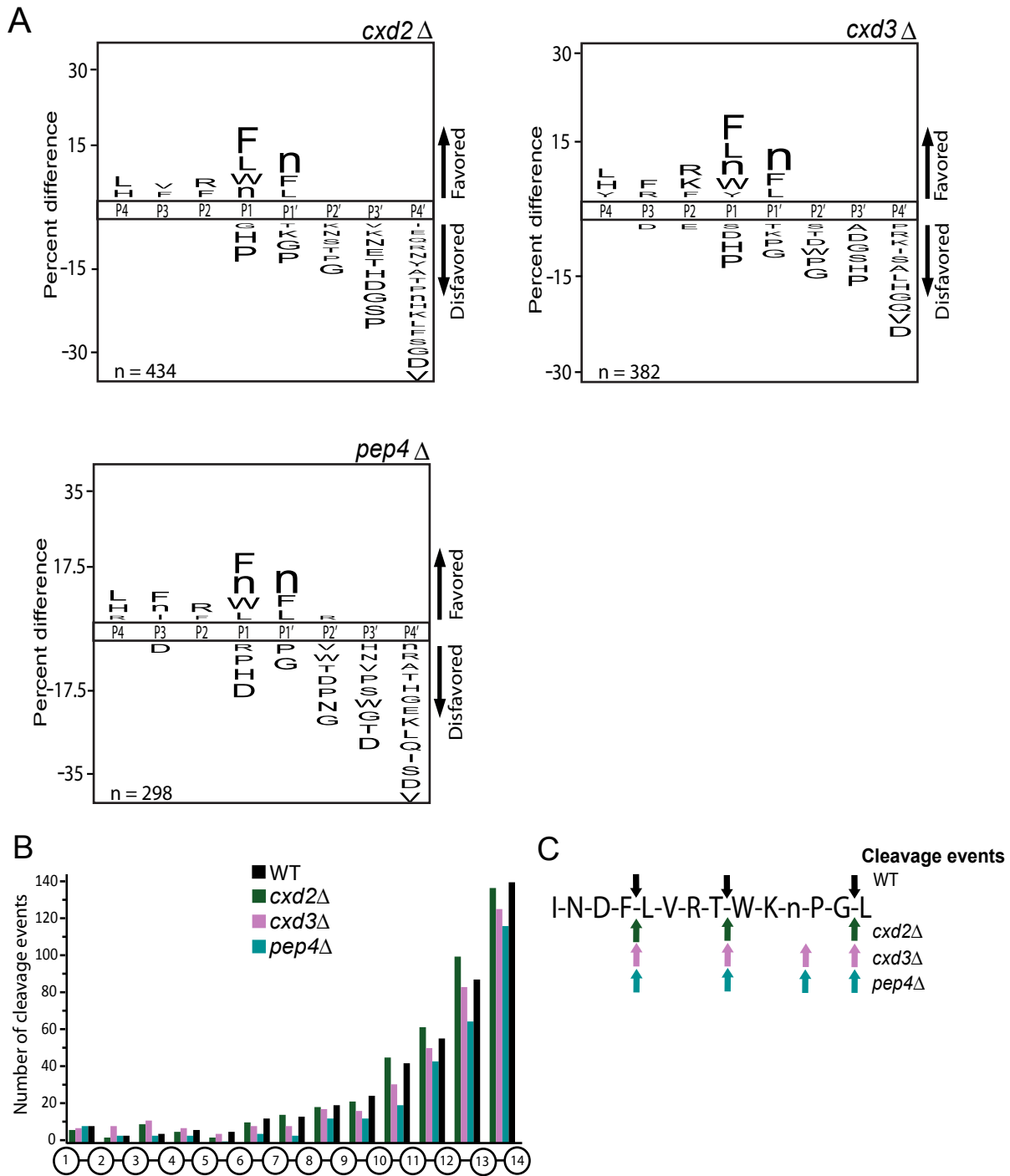
S 4.4 Fig. MSP-MS analysis of secreted peptidase activity in *cxd2Δ, cxd3Δ* and *pep4Δ* strains cultured in YNB media. (A) Substrate specificity profiles of the carboxypeptidase deletion strains *cxd2Δ* and *cxd3Δ* as well as the aspartyl peptidase deletion strain *pep4Δ* grown in YNB, (p = 0.05). (B) Positional analysis of the bonds cleaved in the four deletion strains. (C) An example of a representative peptide cleaved by conditioned media from each deletion strain.

A

B

Start of reaction

333 mV[sum= 26314 mV] Profiles 1-79 Smooth Av 1

1327.4
1311.3
1313.4

4247 mV Profiles 1-79: Threshold 25% Centroid

Full length peptide: (mca)GSPAFLAK(dnp)dR
[M+H] Calculated: 1327.42
[M+H] Observed: 1327.4

After 24 hour incubation

37 mV[sum= 7325 mV] Profiles 1-200 Smooth Av 1

715.8
716.9
1327.2
1328.1
1312.2
1313.2

5245 mV Profiles 1-200: Threshold 25% Centroid

Expected cleavage product: (mca)GSPAF
[M+H] Calculated: 693.7
[M+Na] Calculated: 715.6
[M+Na] Observed: 715.8

S 4.5 Fig. IQ-2 is cleaved by May1. (A) Proteolysis of IQ-2 was measured in a fluorogenic assay of YNB supernatants from all peptidase deletion strains. Deletion of MAY1 led to complete loss of cleavage of IQ-2. (B) May1 was diluted to 14.6nM in 100mM MES pH 4.5, 100mM NaCl and incubated with IQ-2. At the start of the reaction and after 24 hours of incubation at room temperature, samples were collected and analyzed by Matrix Assisted Laser Desorption Ionization-Time of Flight (MALDI-TOF). Based on analysis of its substrate specificity, it was hypothesized that May1 would cleave between the phenylalanine and leucine in IQ-2. The sodium adduct was observed for the N-terminal fragment of the expected cleavage product, confirming the site of cleavage.
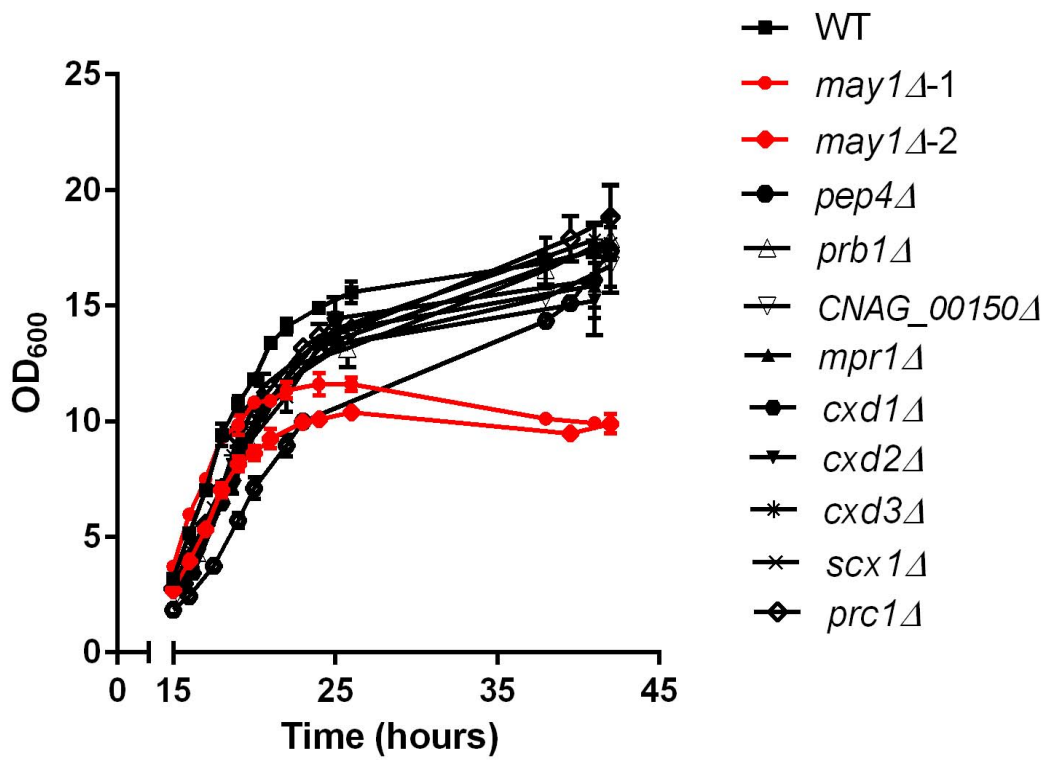
S 4.6 Fig. Growth curves for all peptidase deletion strains. OD_{600} measurements were recorded for cultures grown in triplicate. Averages and S.D. of triplicates are shown.
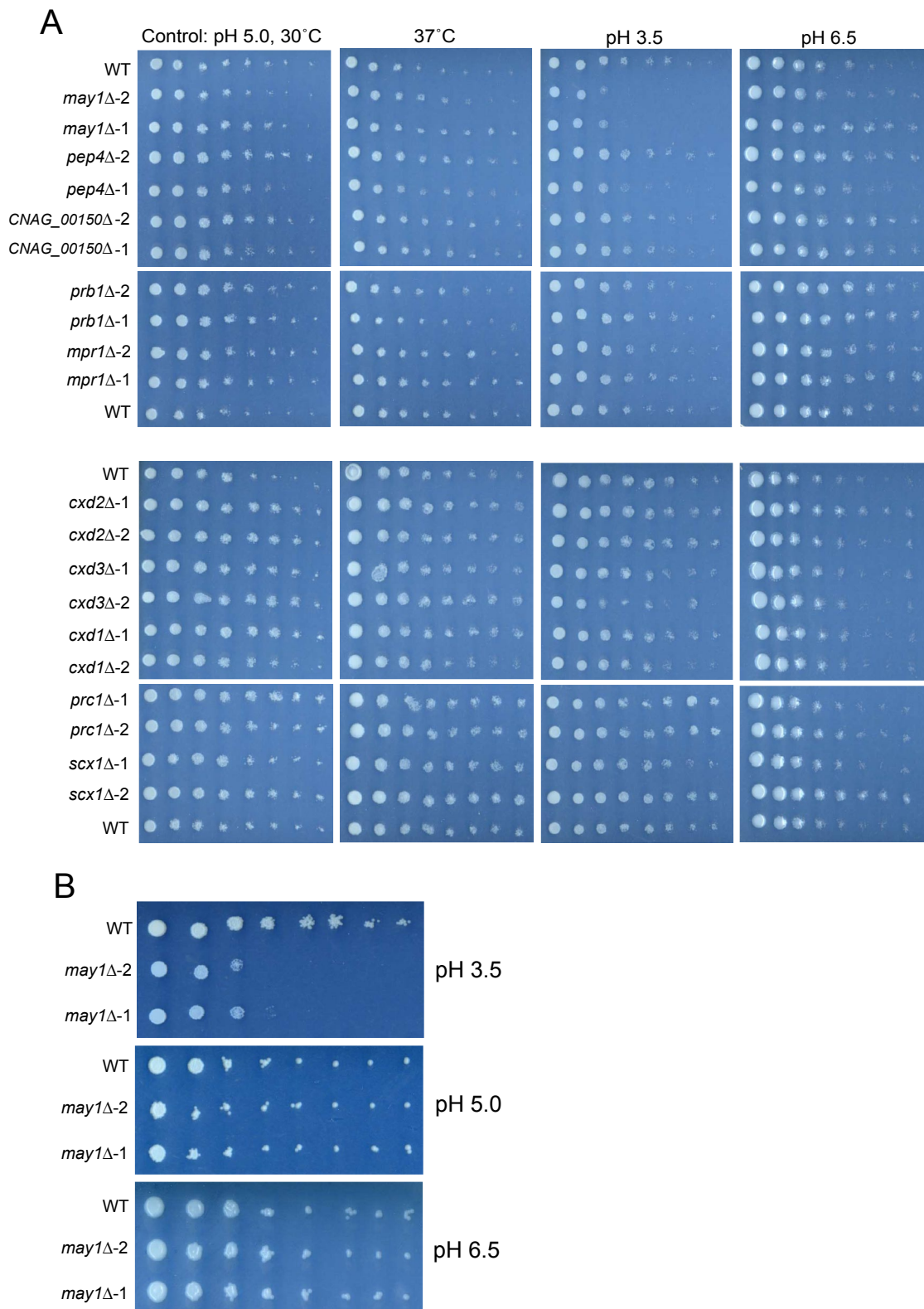
S 4.7 Fig. Temperature and pH tolerance of peptidase deletion strains. (A) Two independent isolates of each peptidase deletion strain were spotted in a 10-fold dilution series on YNB agar plates and grown for 48 hours before imaging. (B) pH tolerance of may1Δ strains after 72 hours of growth.
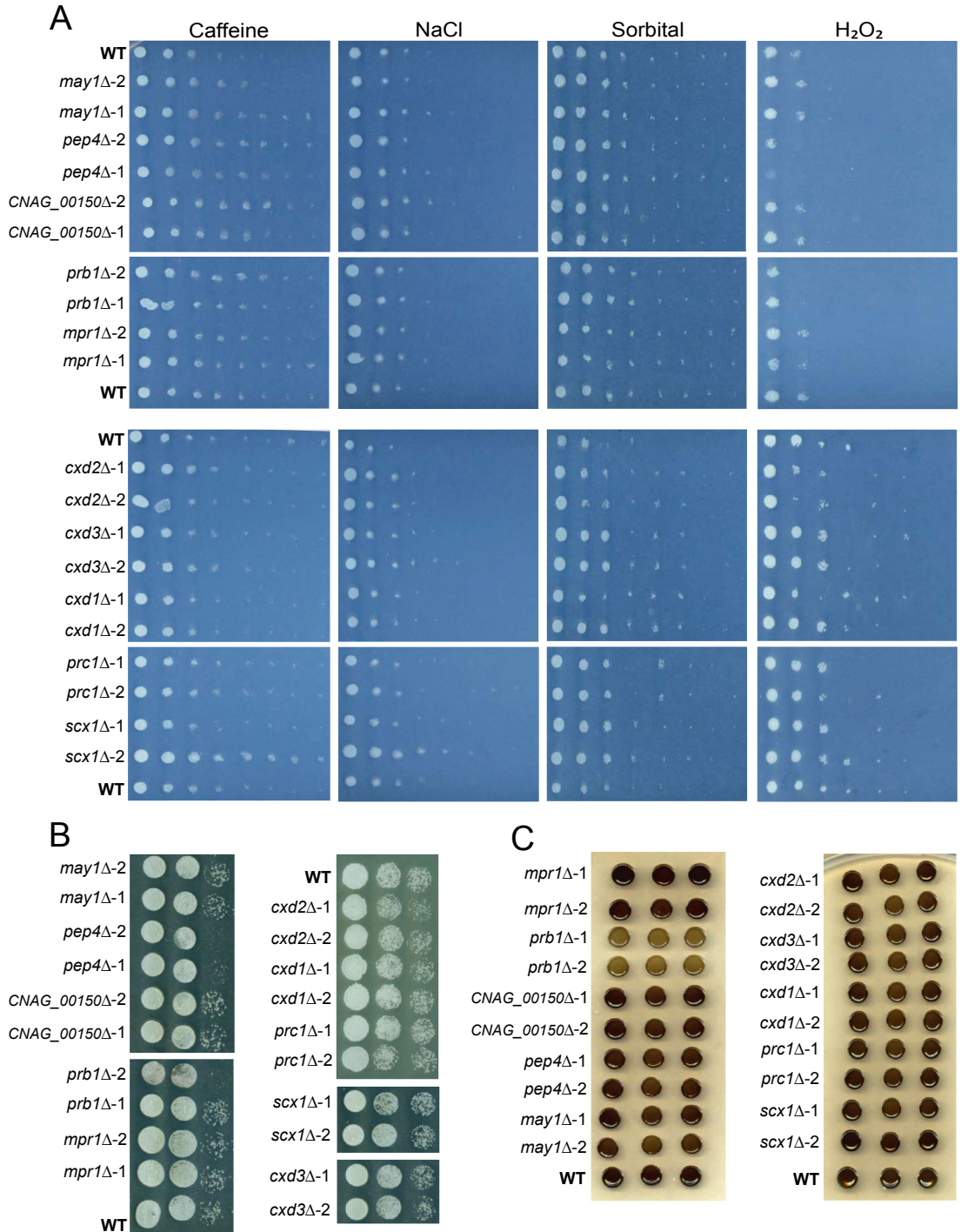
S 4.8 Fig. Tolerance to solute, peroxide and cell wall stress and production of melanin of peptidase deletion strains. (A) 10-fold dilution series of all peptidase deletion strains were spotted on YNB agar plates containing the indicated stress and grown for 48 hours, except for H$_2$O$_2$ plates which were grown for four days before imaging. (B) 10-fold dilution series of peptidase deletion strains grown on rich media plates (YPAD) containing 0.02% SDS and imaged after four days of growth. (C) Melanin production in the presence of L-DOPA. Strains were spotted in triplicate and images were taken after 72 hours of growth.

From left to right

WT
mpr1Δ
cxd2Δ

prb1Δ
scx1Δ
pep4Δ

cxd3Δ
may1Δ
prc1Δ

cxd1Δ
CNAG_00150Δ
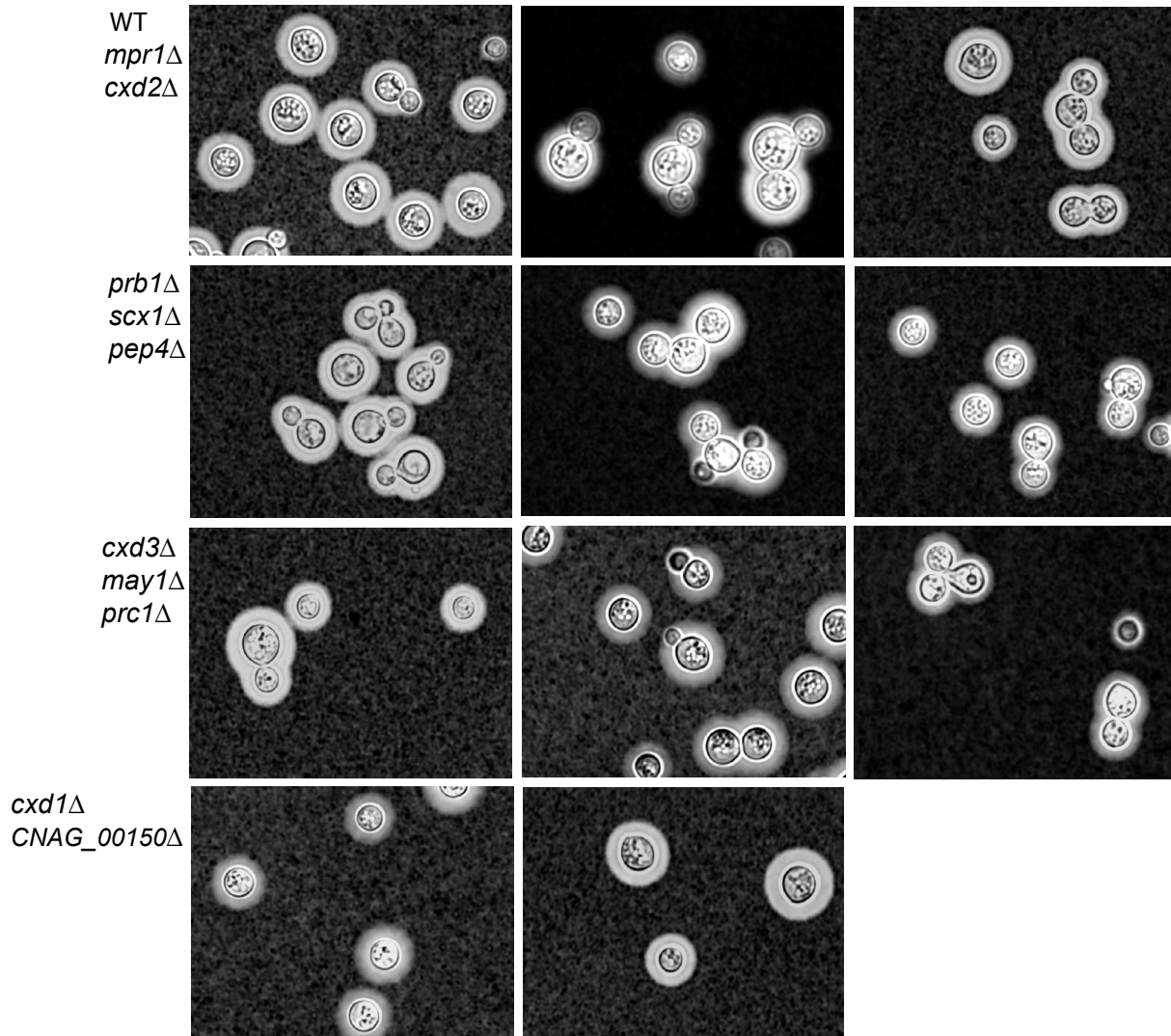
S 4.9 Fig. Capsule production of peptidase deletion strains. Representative images for each deletion strain are shown after 48 hours of capsule induction in Sabourad's media. Contrast staining with India ink was used to visualize capsule.
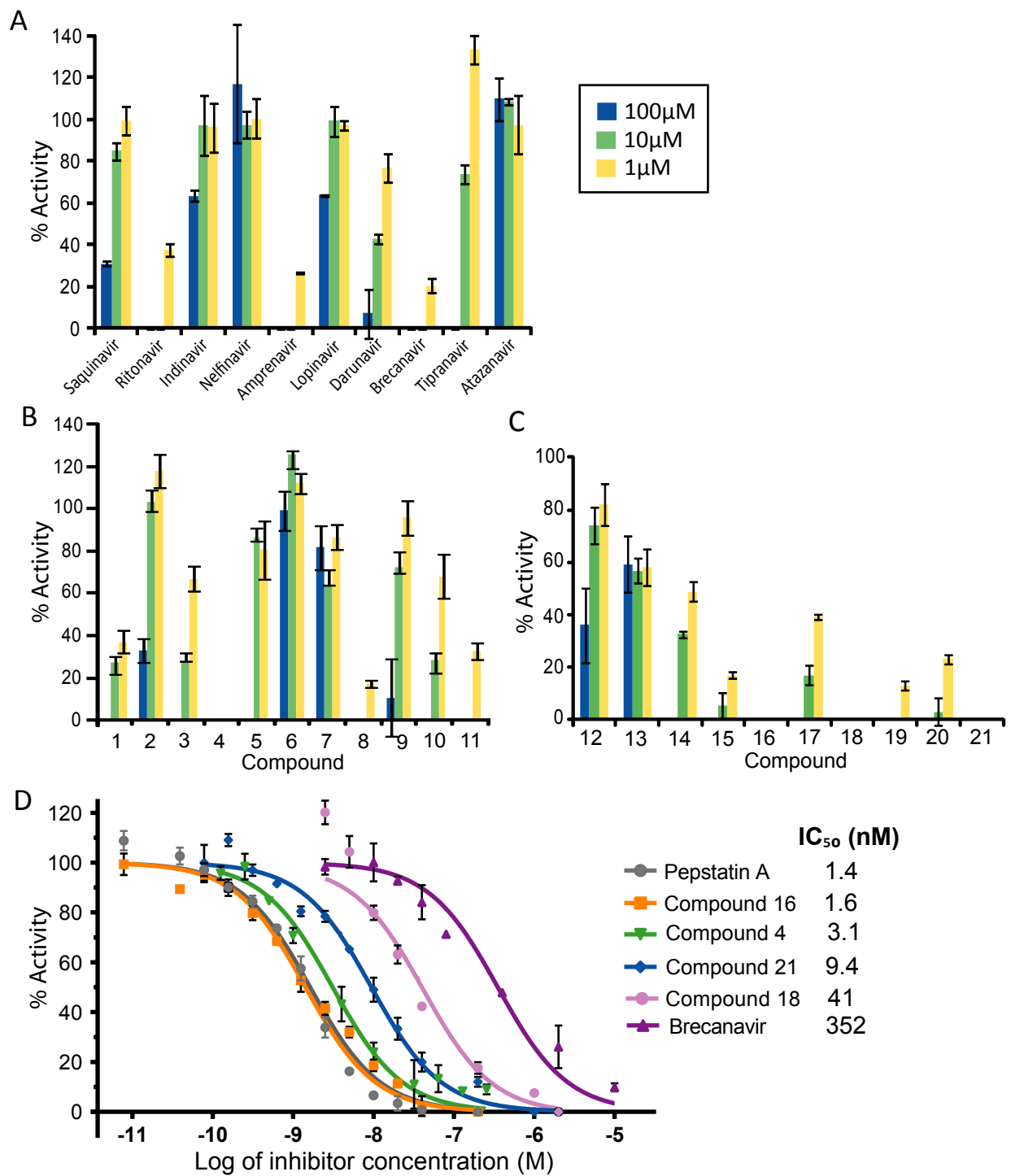
S 4.10 Fig. Screen of aspartyl peptidase inhibitors. Panels (A), (B) and (C) show the results of each inhibitor compound tested in triplicate at 100µM, 10µM and 1µM. May1 activity against IQ-2 was measured. The average value and S.D. of triplicates are shown. (D) IC50 values were calculated for Brecanavir, pepstatin A and compounds 4, 16, 18 and 21. Values are averaged from triplicates and S.D. is shown by error bars.
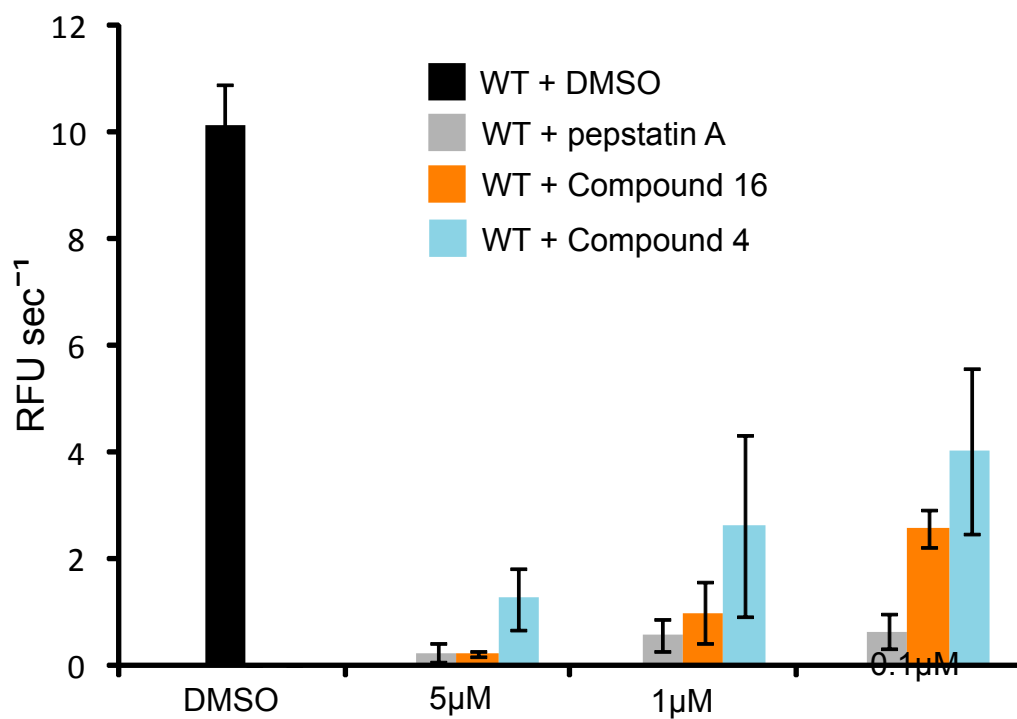
S 4.11 Fig. May1 activity in cultures treated with aspartyl peptidase inhibitors. Activity was recorded against the substrate IQ-2. Average values and S.D. of triplicate measurements are shown.

S 4.1 Table. Sequences of internally quenched fluorogenic substrates.

Name	N-terminus	Sequence	C-terminus
IQ-1	K-AMC	PLGKQVEY	K-DNP
IQ-2	AMC	GSPAFLA	K-DNP, dR
IQ-3	K-AMC	PVYtQASS	K-DNP
IQ-4	K-AMC	KCACSNHE	K-DNP
IQ-5	K-AMC	SRKDKATC	K-DNP
IQ-6	AMC	VGKWSYRn	K-DNP
IQ-7	AMC	PKRLSALL	K-DNP
IQ-8	AMC	IRnQKIET	DNP, dRdR
IQ-9	AMC	WKLLAKWV	K-DNP
IQ-10	AMC	PSLIAKWV	K-DNP

Legend: AMC: aminomethylcoumarin.
K-AMC: AMC bound to the side chain of lysine.
DNP: di-nitrophenol.
K-DNP: DNP bound to the side chain of lysine.
t: tert-butyl glycine
n: norleucine
dR: D form of arginine

Proteins identified in YNB conditioned media

Rank	Unique peptides	%Coverage	Best E value	C. neoformans var grubii Uniprot #	Protein name or putative function	Secretion signal	Non-classic secretion	Microvesicle associated
[1]	7	8.3	1.20E-08	J9VZG6	Expressed protein		*	
[2]	6	58.4	8.20E-08	J9VLJ9	Superoxide dismutase [Cu-Zn]		*	
[3]	6	69.4	2.00E-09	J9VPP8	Putative uncharacterized protein		*	
[4]	11	13	2.00E-06	J9VLH9	glyoxyl oxidase putative	*		
[5]	5	7.5	1.40E-07	J9VWK6	Plasma membrane H(+)-ATPase		*	*
[6]	6	56.7	5.20E-06	J9VTJ3	Multiprotein-bridging factor 1			
[7]	5	21.5	6.50E-07	J9VM09	Inorganic diphosphatase, putative			
[8]	5	17.5	6.30E-08	J9VTK1	Glutamate dehydrogenase (NADP+), putative			*
[9]	5	46	3.80E-07	J9VSL1	Ribosomal protein, putative			
[10]	5	59.3	6.30E-07	O94746	FK506-binding protein 1		*	
[11]	6	9.3	1.90E-06	J9VVQ4	glyoxyl oxidase putative	*		
[12]	4	13.5	1.80E-06	J9VWJ8	ATP synthase, putative			
[13]	3	30.9	1.00E-08	J9W3S6	Nascent polypeptide-associated complex subunit beta		*	
[14]	6	37.4	2.40E-06	J9VZD2	60s ribosomal protein l17, putative		*	
[15]	4	22.9	3.60E-06	J9VHP1	Voltage-dependent ion-selective channel, putative			*
[16]	4	8.5	6.90E-08	Q9P8P2	Phospholipase B	*		
[17]	8	52.7	1.50E-04	B0L2Y2	Translation elongation factor 1a (Fragment)		*	
[18]	5	54.9	5.70E-08	J9VPH6	Putative uncharacterized protein		*	
[19]	4	31	1.60E-07	J9VPA5	Putative uncharacterized protein		*	
[20]	3	14.9	9.30E-08	J9VJ71	Putative uncharacterized protein		*	
[21]	3	11.8	1.90E-06	J9VND2	Chitin deacetylase-like mannoprotein		*	
[22]	4	14.2	2.40E-05	J9VV89	Transaldolase			
[23]	2	11.9	2.70E-09	J9VF55	Putative uncharacterized protein	*		
[24]	2	14.6	2.90E-07	J9W2G5	Ribosomal protein L13, putative		*	
[25]	4	7.9	4.80E-05	J9VY34	Eukaryotic ADP/ATP carrier		*	
[26]	3	31	7.50E-09	J9VQ24	Initiation factor 5a (Eif-5a), putative			
[27]	3	32.7	5.20E-07	J9VPH6	Hmp1 protein, putative		*	
[28]	3	7.9	2.30E-06	J9VTH3	Putative uncharacterized protein		*	
[29]	4	6.1	8.00E-06	J9VEL7	Heat shock protein 70			*
[30]	4	18.5	9.50E-07	J9W0B1	Ras2, putative		*	
[31]	3	10.3	3.90E-06	J9W3S8	Inorganic phosphate transporter, putative		*	
[32]	5	3.1	2.00E-04	T2BMZ0	Alpha-1,3-glucan synthase, putative	*		
[33]	3	32.5	1.30E-04	J9VRA3	Gal4 DNA-binding enhancer protein 2, putative		*	
[34]	4	57.4	7.20E-05	J9VPT7	Putative uncharacterized protein			
[35]	3	9.7	1.10E-06	J9VVB2	Adenylate kinase 1		*	

[36]	3	20.1	5.60E-07	J9VW33	Putative uncharacterized protein			
[37]	3	14.3	1.30E-06	J9VTW1	Expressed protein		*	
[38]	3	44.8	2.10E-05	J9VF41	Putative uncharacterized protein		*	
[39]	4	15.8	9.40E-06	J9VV35	Putative uncharacterized protein			
[40]	2	3.8	5.20E-07	J9VL23	Putative uncharacterized protein		*	*
[41]	4	38.5	1.20E-04	J9VKF2	Thioredoxin			*
[42]	2	9.9	1.40E-07	Q9P8W9	Peptidyl-prolyl cis-trans isomerase			
[43]	2	29.3	1.90E-08	J9VJC5	60s ribosomal protein l37a, putative		*	
[44]	2	17.5	6.20E-08	J9VQ69	Putative uncharacterized protein			
[45]	2	14	2.40E-06	J9VTE3	Putative uncharacterized protein		*	
[46]	2	31.3	6.30E-08	J9VUT1	Putative uncharacterized protein		*	
[47]	2	4.4	1.10E-06	J9VTZ8	Carboxypeptidase D	*		
[48]	4	19.7	1.70E-04	J9VMQ3	Putative uncharacterized protein			
[49]	2	22	3.60E-07	J9VF13	Ribosomal protein L35, putative			*
[50]	3	29.6	4.60E-07	J9W025	Expressed protein			
[51]	3	3.4	2.50E-05	J9VR37	Putative uncharacterized protein			
[52]	4	39.6	2.20E-05	J9VHI6	Electron carrier, putative		*	
[53]	2	14.3	3.30E-06	J9VQK8	Putative uncharacterized protein		*	
[54]	2	32.5	1.60E-06	J9VUT1	Putative uncharacterized protein		*	
[55]	2	5.4	8.90E-07	J9VU43	Putative uncharacterized protein			
[56]	3	18.6	9.10E-07	J9VJH1	Putative uncharacterized protein		*	
[57]	2	20.9	5.20E-06	J9VTV6	6,7-dimethyl-8-ribityllumazine synthase			
[58]	2	13	3.50E-05	J9VLM0	Structural molecule, putative			
[59]	3	14.3	1.30E-05	J9VL75	40S ribosomal protein S6			
[60]	2	48.6	5.00E-07	J9VNN9	Ribosomal protein P2, putative			*
[61]	3	20.6	2.70E-05	J9VFP1	Expressed protein			
[62]	2	39.1	1.00E-05	J9VY76	Long-chain fatty acid transporter, putative			*
[63]	2	18.3	3.00E-06	J9VQS6	Putative uncharacterized protein		*	
[64]	2	18.2	3.70E-06	J9VT78	Nonhistone protein 6, putative			
[65]	2	5.4	4.40E-07	J9VHM5	Putative uncharacterized protein			
[66]	3	14.4	2.20E-06	J9VSC8	Ribosomal protein L19			
[67]	1	11.9	1.30E-08	J9VFG4	PRCDNA87, putative		*	
[68]	2	4.5	2.00E-05	J9VGV0	Enolase		*	
[69]	2	13.5	9.80E-07	J9VF18	Mannoprotein MP88, putative	*		*
[70]	2	23	5.80E-06	J9VK98	Putative uncharacterized protein		*	
[71]	2	11.1	2.10E-06	J9VIF7	Cytoplasm protein, putative		*	
[72]	1	14.5	3.00E-08	J9W225	60s ribosomal protein l23, putative		*	
[73]	4	28.8	8.20E-05	J9VHE4	Ribosomal protein S11, putative		*	
[74]	3	12.7	5.40E-05	J9VXN5	Malate dehydrogenase		*	
[75]	2	24.8	2.60E-05	J9VF52	60s ribosomal protein l34-b, putative		*	

No.	n	Score	E-value	Accession	Description			
[76]	3	35.8	2.10E-05	J9VLT3	Suppressor of initiator codon mutations, putative			*
[77]	2	6.2	3.90E-06	J9VVV4	Stearoyl-CoA 9-desaturase, putative			*
[78]	2	11.8	5.60E-06	J9VUE7	Putative uncharacterized protein			*
[79]	3	17	4.40E-05	J9VPZ7	Rho GDP-dissociation inhibitor 1, putative			*
[80]	2	29.9	4.90E-05	J9VK51	40S ribosomal protein S21			*
[81]	1	15.6	1.20E-07	J9VNR0	Putative uncharacterized protein			*
[82]	2	0.9	2.90E-06	J9VVH4	Fatty acid synthase beta subunit			
[83]	3	14.1	8.50E-05	J9VTL4	Mannitol-1-phosphate dehydrogenase			
[84]	2	10	2.50E-06	J9VI11	6-phosphogluconate dehydrogenase, decarboxylating			
[85]	2	7.7	1.80E-04	J9VKH4	Cytoplasm protein, putative		*	*
[86]	1	2.5	6.90E-06	J9VZ45	1,3-beta-glucanosyltransferase	*		
[87]	2	38.1	8.00E-05	J9VS14	Expressed protein			
[88]	2	19.3	6.90E-06	J9VFY8	PRCDNA95, putative			*
[89]	1	9.6	9.50E-08	J9W241	Cytochrome c oxidase subunit V, putative			*
[90]	2	16.8	1.60E-04	J9VU89	Ribosomal protein S20, putative			*
[91]	2	7.7	3.80E-05	J9VET5	Inosine-5'-monophosphate dehydrogenase	*		
[92]	1	2.3	8.30E-06	J9VXZ9	**llular elastinolytic metalloproteinase, putative, zinc**	*		*
[93]	2	15.3	9.00E-05	J9VMM1	Putative uncharacterized protein			*
[94]	1	12.6	1.40E-06	J9VJD3	Profilin			*
[95]	3	19.3	2.00E-04	J9VUD6	Structural constituent of ribosome, putative			
[96]	1	6	5.20E-07	J9W0T8	Glucan 1,3 beta-glucosidase protein, putative			*
[97]	1	4.2	3.90E-06	J9W2Y1	Rds1 protein, putative	*		
[98]	1	7	3.20E-07	J9VUB8	Cytokine inducing-glycoprotein			
[99]	4	18.6	2.80E-04	J9VUF9	Aldehyde reductase i, putative			
[100]	1	4	2.90E-06	J9VX05	Putative uncharacterized protein			
[101]	1	8.7	4.90E-05	J9VQI0	60S ribosomal protein L24 (L30), putative		*	*
[102]	1	2.7	1.40E-06	J9W0L4	**Serine-type endopeptidase, putative**			
[103]	1	3	5.80E-07	J9VP88	ATP synthase subunit alpha	*		
[104]	1	3.4	8.40E-07	J9VI89	RNA-binding protein sce3, putative			
[105]	2	9.8	2.50E-05	J9VPV1	L-malate dehydrogenase, putative			*
[106]	2	3.2	1.50E-04	J9VPM4	Putative uncharacterized protein		*	
[107]	3	6.8	0.0024	J9VXF1	ahydropteroyltriglutamate-homocysteine S-methyltransferase, putative			*
[108]	1	6	4.70E-06	J9VN13	Allergen, putative			*
[109]	2	15.1	3.00E-04	J9VRE8	Putative uncharacterized protein			
[110]	1	13.7	2.10E-05	J9VUN4	Copper chaperone, putative			*
[111]	2	13.7	2.90E-04	J9VZS2	40s ribosomal protein s5-1, putative			*
[112]	1	9.4	5.60E-06	J9VYI4	Expressed protein			
[113]	2	15.9	8.00E-05	J9VD88	Ribosomal protein S3, putative			
[114]	1	7.3	7.60E-06	J9VW34	Putative uncharacterized protein			
[115]	1	19.5	1.00E-05	J9VI59	Expressed protein			

No.	n	Score	E-value	Accession	Description	*	*
[116]	1	18.6	5.90E-05	J9VG91	60s ribosomal protein l30-1 (L32), putative		*
[117]	1	19.6	8.70E-06	J9VDK8	Chaperone, putative		*
[118]	2	7.2	5.10E-05	J9VSJ2	60S ribosomal protein L6		
[119]	1	9.3	1.20E-04	J9VUU8	Ribosomal protein, putative		
[120]	1	6.3	6.90E-07	J9VKI7	Inorganic phosphate transporter pho88, putative		
[121]	1	4.3	7.40E-06	J9VII7	Putative uncharacterized protein		*
[122]	1	10.3	1.80E-05	J9VVW7	Putative uncharacterized protein		*
[123]	1	3	1.00E-05	J9VPU9	Putative uncharacterized protein		
[124]	2	39.4	6.80E-04	J9VQ34	Putative uncharacterized protein		*
[125]	1	14	3.40E-05	J9VD54	Putative uncharacterized protein		*
[126]	1	11.3	1.20E-05	J9VRI4	Expressed protein		
[127]	1	5.3	2.50E-06	J9VJV4	Ribose-5-phosphate isomerase		*
[128]	1	2.1	1.50E-05	J9VJF3	Putative uncharacterized protein		
[129]	1	3.2	2.00E-05	J9VJK0	Putative uncharacterized protein		*
[130]	1	3.2	6.00E-06	J9VRH0	Glucosidase, putative		*
[131]	1	5.5	1.40E-04	J9VN50	Putative uncharacterized protein		
[132]	1	3.1	1.70E-04	J9VNK3	Putative uncharacterized protein		
[133]	1	8.2	8.30E-05	J9VRP1	Putative uncharacterized protein		*
[134]	1	12.7	1.50E-05	J9VFV3	Structural constituent of ribosome, putative		*
[135]	1	5.3	4.50E-04	J9VPD7	Chitin deacetylase, putative	*	
[136]	1	7.6	3.30E-05	J9VPE0	40s ribosomal protein s23, putative		*
[137]	1	14.5	4.10E-05	J9VKN9	Structural constituent of ribosome, putative		
[138]	1	2.2	9.90E-05	J9VD75	Chaperone, putative		*
[139]	1	3	2.10E-04	J9VJM1	Exo-beta-1,3-glucanase		*
[140]	2	6.6	3.40E-04	J9VR33	60S ribosomal protein L36		
[141]	1	7	2.10E-04	J9VXQ3	Putative uncharacterized protein		
[142]	1	11.5	5.40E-05	J9VLO3	Putative uncharacterized protein		
[143]	1	5.2	2.80E-05	J9VQJ6	Expressed protein		*
[144]	1	7.7	1.30E-04	J9VKW7	Putative uncharacterized protein		
[145]	2	12.7	4.40E-05	J9VQF1	Thioredoxin-dependent peroxide reductase, putative		*
[146]	1	2.2	2.70E-04	J9VS18	CRYNB Putative uncharacterized protein		
[147]	2	26.8	0.0033	J9VTH9	Putative uncharacterized protein		*
[148]	2	10.2	5.50E-04	J9VS17	14-3-3 protein, putative		*
[149]	1	3.5	8.80E-06	J9VPP7	ATP synthase subunit beta		
[150]	3	19.6	0.0015	J9W0I6	Structural constituent of ribosome, putative		
[151]	1	10.1	1.90E-04	J9VL49	Ribosomal chaperone, putative		*
[152]	1	5.1	5.30E-05	J9VPL3	Putative uncharacterized protein		
[153]	1	31.7	7.70E-05	J9VJW7	Expressed protein		
[154]	1	9.3	4.80E-05	J9W2U5	Putative uncharacterized protein		
[155]	1	6.7	3.60E-04	J9VUB8	Cytokine inducing-glycoprotein, putative		
[156]	1	6	1.00E-04	J9VWT4	Cytochrome-b5 reductase, putative		*
[157]	1	10.3	9.20E-05	J9VHC1	60s ribosomal protein, putative		*
[158]	1	3	2.60E-04	J9VMB5	ARF GTPase activator, putative		*
[159]	1	12.7	4.60E-05	J9VE62	Expressed protein		*
[160]	1	23.7	0.002	J9VG26	Histone H2A		

108

#								
[161]	1	7.9	0.0037	T2BPD2	RNA binding protein, putative		*	
[162]	1	1	5.50E-04	J9VQZ9	Elongation factor 3			
[163]	1	7.7	0.0013	J9VWW9	Superoxide dismutase		*	
[164]	1	3.4	6.40E-04	J9VLU5	Glucosidase, putative		*	
[165]	3	6.2	0.0014	J9VHV5	Citrate synthase		*	
[166]	1	3.9	2.70E-04	J9VHW3	Disulfide-isomerase, putative	*		
[167]	1	1	0.0012	J9VT51	Structural molecule, putative			
[168]	1	2	5.50E-04	J9VJW5	Glutamate decarboxylase, putative			
[169]	1	5.3	8.30E-04	J9VW36	Cytoplasm protein, putative		*	
[170]	1	3.1	7.80E-04	J9VZ98	Ubiquitin-protein ligase, putative			
[171]	1	4.3	0.0011	J9VST6	Putative uncharacterized protein			
[172]	1	11.7	5.50E-04	J9VUB1	Histone H4			*
[173]	1	2.2	9.90E-04	J9VVE7	Putative uncharacterized protein			
[174]	2	2.6	0.0033	J9VSF6	Putative uncharacterized protein			
[175]	1	1.5	0.0018	J9VW13	Eukaryotic translation initiation factor 2C 2, putative			
[176]	1	10.8	0.0024	J9VKB8	Histone H2B		*	
[177]	1	2.7	7.40E-04	J9VI89	Putative uncharacterized protein			
[178]	1	8.2	0.0022	J9VXL7	40S ribosomal protein S8			
[179]	1	18.9	0.0034	J9VVL6	13 kDa ribonucleoprotein-associated protein			
[180]	1	4.9	0.0021	J9VR46	all wall organization and biogenesis-related protein, putati	*		
[181]	2	9.8	0.0088	J9VZT6	Putative uncharacterized protein		*	
[182]	1	4.6	0.0022	J9VF99	Putative uncharacterized protein		*	
[183]	1	1.5	0.0067	na	Putative uncharacterized protein			
[184]	1	7.5	0.0017	J9VH03	Ribosomal protein S19, putative			
[185]	1	10	0.0034	J9VT96	Ribosomal protein L18			
[187]	1	19.5	0.0069	J9VM09	Inorganic diphosphatase (Fragment)			
[188]	1	5.2	0.0011	J9VRX9	Thioredoxin reductase		*	
[189]	1	2.4	2.50E-04	J9VK44	Putative uncharacterized protein		*	
[190]	1	5.8	0.0035	J9VH48	Putative uncharacterized protein		*	
[191]	1	7.7	0.0036	J9VXH5	Ribosomal protein S18, putative		*	
[192]	1	5	5.40E-04	J9VQ17	DNA-binding protein hexbp, putative		*	
[193]	1	5.3	0.0049	J9VZY1	Putative uncharacterized protein	*		
[194]	1	1.5	0.0062	J9VUL8	**Peptidase, putative, aspartyl endoprotease**	*		
[195]	1	13.5	0.0076	J9VZF4	Transcriptional elongation regulator, putative		*	
[196]	1	7	0.0032	J9VEV6	Mitochondrion protein, putative		*	
[197]	1	9.2	0.0019	J9VQQ2	Putative uncharacterized protein		*	
[198]	1	2.1	0.0048	J9VH59	**Endopeptidase, putative aspartyl**	*		
[199]	1	20.5	0.0055	J9VVA3	Ribosomal protein S27		*	
[200]	1	3.9	0.0057	J9VQK9	UDP-xylose synthase			*

Proteins identified in DME M conditioned media

Rank	Unique peptides	% Cov	Best E val	C. neoformans var grubii Uniprot number	Protein name	Secretion signal	Non-classic secretion	Microvesicle associated	Observed in YNB
[1]	12	71.6	2.10E-09	J9VNS6	Peptidyl-prolyl cis-trans isomerase				yes
[2]	9	64.9	2.50E-09	J9VLJ9	Superoxide dismutase [Cu-Zn]		*		yes
[3]	11	63.1	3.10E-07	J9VHI6	Electron carrier, putative		*		yes
[4]	7	69.4	1.10E-08	J9VPP8	Putative uncharacterized protein		*	*	yes
[5]	11	42.4	3.20E-08	J9VE48	3-isopropylmalate dehydrogenase				
[6]	7	49.3	1.60E-08	J9VJU4	Cofilin		*		
[7]	9	71.2	2.20E-06	J9VKF2	Thioredoxin				yes
[8]	6	36	1.70E-06	J9VPA7	Carbonic anhydrase 2				
[9]	4	7	2.50E-07	J9VZ70	Heat shock protein 70, putative			*	
[10]	3	11.5	2.90E-06	J9VNT0	Putative uncharacterized protein				
[11]	4	36.9	3.40E-07	J9VNN9	Ribosomal protein P2, putative			*	yes
[12]	4	30.3	9.60E-07	Q9P8P2	Phospholipase (Fragment)	*			yes
[13]	6	15.5	2.60E-05	J9VRJ9	Adenosylhomocysteinase		*		
[14]	4	13.7	9.30E-07	J9VQI1	Ribosomal protein		*		
[15]	3	3.1	6.20E-07	J9VPF5	Heat shock protein, putative				
[16]	7	17	8.30E-05	J9VVA4	Heat-shock protein 90			*	
[17]	4	23.8	3.80E-07	J9VTL4	Mannitol-1-phosphate dehydrogenase				
[18]	3	24.4	3.90E-05	J9VQF1	Thioredoxin-dependent peroxide reductase, putative		*		yes
[19]	3	11.1	1.00E-07	J9VTK1	Glutamate dehydrogenase			*	yes
[20]	4	13.5	4.60E-05	J9VUR9	Polyadenylate-binding protein, cytoplasmic and nuclear				
[21]	6	9.4	2.30E-05	J9VPD7	Putative uncharacterized protein	*			yes
[22]	5	40	3.40E-06	J9VTQ3	Expressed protein		*		
[23]	3	35.4	9.80E-08	J9VJD3	Profilin		*		yes
[24]	3	5.3	1.20E-05	J9VZ45	1,3-beta-glucanosyltransferase, putative	*			yes
[26]	1	2.1	3.50E-08	J9VZG6	Expressed protein		*		
[27]	2	14.3	2.70E-07	J9VDT9	Putative uncharacterized protein			*	
[28]	2	43.1	5.10E-07	J9VJJ1	PRCDNA35, putative				
[29]	3	19.5	3.40E-06	J9VS17	14-3-3 protein, putative				yes
[30]	2	10.7	9.70E-08	J9VGH4	Cytoplasm protein, putative			*	
[31]	2	18.6	9.40E-04	J9VLE6	Ubiquitin-60S ribosomal protein L-40				
[32]	3	9.3	5.30E-07	J9VVQ4	Putative uncharacterized protein	*			yes
[33]	4	11.8	1.40E-06	J9V1J2	Putative uncharacterized protein				
[34]	3	3.7	1.90E-06	J9VK61	Putative uncharacterized protein				
[35]	6	10.1	1.00E-04	J9VXF1	tetrahydropteroyltriglutamate-homocysteine S-methyltransferase, putative		*		yes
[36]	2	28	5.90E-07	J9W3S6	Nascent polypeptide-associated complex subunit beta		*		yes
[37]	3	8.7	1.10E-05	J9VX99	Pyruvate kinase				
[38]	3	11.5	0.0021	J9VJ43	Glucose-6-phosphate isomerase				
[39]	2	7.1	1.30E-06	J9VPG5	Putative uncharacterized protein				
[40]	2	20.8	9.60E-06	J9VTH9	Putative uncharacterized protein				yes
[41]	2	19.7	4.50E-06	J9VM42	Nucleoside diphosphate kinase				
[42]	2	26.2	1.70E-05	J9W1N8	Putative uncharacterized protein		*		
[43]	2	10	3.00E-05	J9VM14	Ubiquinol-cytochrome c reductase iron-sulfur subunit		*		yes
[44]	2	11.8	3.80E-06	J9VGY3	Putative uncharacterized protein				
[45]	4	28.7	4.60E-04	J9VJP0	Putative uncharacterized protein		*		
[46]	2	8.4	2.50E-06	J9VRA3	Nascent polypeptide-associated complex subunit alpha		*		yes
[47]	2	6.6	3.10E-06	J9VHW3	Disulfide-isomerase, putative	*			yes
[48]	1	3.5	2.10E-07	J9VPN6	Putative uncharacterized protein	*			
[49]	1	7	5.30E-08	J9VUB8	Cytokine inducing-glycoprotein				yes
[50]	2	23.3	1.00E-05	J9VTJ3	Multiprotein-bridging factor 1		*		yes

[51]	2	4.5	7.20E-06	Homocitrate synthase, putative		*		
[52]	1	25	1.60E-07	Nuclear transport factor 2 (Ntf2), putative		*		
[53]	2	8.4	1.40E-04	Putative uncharacterized protein				yes
[54]	2	4.1	5.10E-05	Plasma membrane H(+)-ATPase			*	yes
[55]	1	2.5	1.90E-07	Putative uncharacterized protein				
[56]	1	6	5.40E-07	Glucan 1,3 beta-glucosidase protein, putative		*		
[57]	2	7.4	4.80E-05	Adenylate kinase, putative		*		
[58]	1	3.7	1.50E-07	Putative uncharacterized protein				
[59]	2	8.1	5.90E-04	Glucan 1,3-beta-glucosidase	*			
[60]	2	8.1	5.20E-06	Putative uncharacterized protein		*		
[61]	2	11.8	2.00E-04	Putative uncharacterized protein		*		
[62]	1	4.2	3.90E-06	Rds1 protein, putative	*			yes
[63]	2	9.7	2.90E-04	Putative uncharacterized protein				yes
[64]	1	4.1	5.60E-07	Phosphomannomutase, putative		*		
[65]	1	3	2.60E-06	Putative uncharacterized protein	*			yes
[66]	1	2.3	2.20E-06	Putative uncharacterized protein				yes
[67]	1	5.4	3.70E-05	Putative uncharacterized protein		*		yes
[68]	1	6.4	1.20E-05	Cysteine synthase				
[69]	1	13.6	4.10E-06	Nonhistone protein 6, putative		*		yes
[70]	2	6.9	8.50E-05	Uv excision repair protein rhp23, putative				
[71]	1	27.1	1.30E-06	Putative uncharacterized protein				
[72]	1	5.5	8.10E-06	Hydroxyacylglutathione hydrolase, putative		*		
[73]	1	16.4	1.10E-05	Mini-intein (Fragment)				
[74]	1	15.7	3.30E-05	Putative uncharacterized protein		*		
[75]	1	3.6	9.20E-06	Enolase			*	yes
[76]	1	12.8	1.90E-06	Putative uncharacterized protein				
[77]	1	1.3	1.30E-05	Putative uncharacterized protein				
[78]	1	21.2	2.50E-06	Expressed protein				yes
[79]	1	6.8	4.30E-06	Putative uncharacterized protein				
[80]	1	7	1.60E-04	Cytochrome-c oxidase chain VI, putative		*		
[81]	1	6	3.40E-06	Triosephosphate isomerase		*		
[82]	1	2	2.70E-06	Meiotic recombination-related protein, putative				
[83]	1	1.9	1.70E-05	Cytoplasm protein, putative				
[84]	1	5.5	3.00E-05	Wos2 protein (P21), putative				
[85]	1	3.6	8.10E-05	FACT complex subunit POB3				
[86]	2	7.9	0.001	Ribosomal protein S3, putative				yes
[87]	1	1.7	9.80E-06	Heat shock protein, putative			*	
[88]	1	5.4	2.40E-05	Cytoplasm protein, putative		*		
[89]	1	4.5	1.50E-04	Chaperone regulator, putative				
[90]	1	2.9	1.20E-04	Chaperone, putative		*		yes
[91]	1	29.3	2.10E-04	Expressed protein		*		
[92]	4	10.7	0.0078	Elongation factor 1-alpha			*	
[93]	1	3.2	4.30E-04	Peptidyl-prolyl cis-trans isomerase D				
[94]	1	9.6	3.90E-05	Expressed protein		*		yes
[95]	1	2.5	3.30E-04	Putative uncharacterized protein				

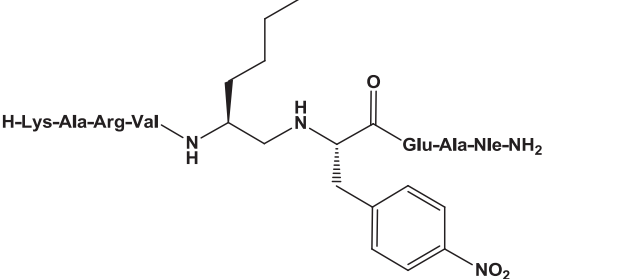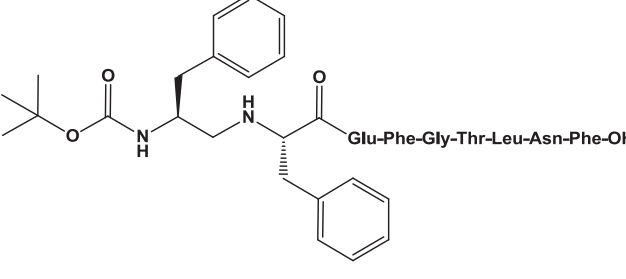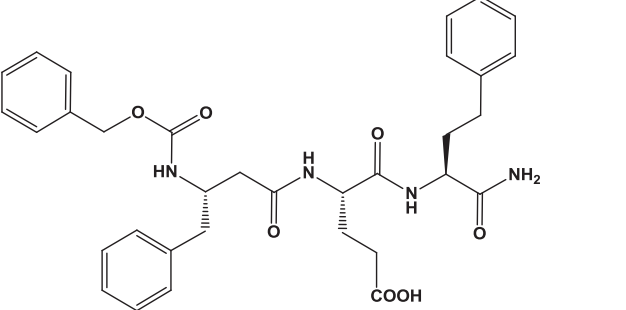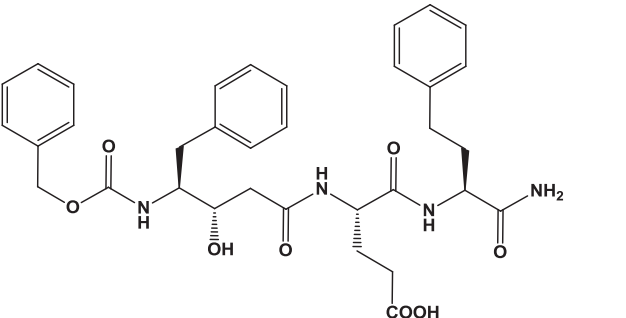|								
[96]	1	5.9	2.20E-04	J9VV67	Putative uncharacterized protein		*	
[97]	1	2	2.20E-05	J9VWV9	Putative uncharacterized protein		*	
[98]	1	2	5.50E-05	J9VTZ8	**Carboxypeptidase D, putative**	*		yes
[100]	1	2.8	1.80E-04	J9VIP1	Putative uncharacterized protein			
[101]	1	8.7	1.30E-04	J9VK98	Putative uncharacterized protein	*	*	yes
[102]	1	3.2	2.60E-04	J9VPM4	Putative uncharacterized protein	*		yes
[103]	1	32.5	0.0021	J9VUT1	heat shock protein (Glucose and lipid-regulated protein), putative		*	yes
[104]	1	1.6	1.20E-04	J9VE82	Putative uncharacterized protein			
[105]	1	10.1	0.0089	J9VMN8	Mitochondrion protein, putative		*	
[106]	3	19.3	0.0028	J9VH03	Ribosomal protein S19, putative			yes
[107]	1	3.1	4.20E-05	J9VJS8	Dihydroorotase, putative			
[108]	1	5.9	6.80E-04	J9VV89	Transaldolase			yes
[109]	1	1.1	7.30E-04	J9VV31	CRYNB Putative uncharacterized protein			
[110]	1	2.6	7.00E-04	J9VEL7	Heat shock protein 70		*	yes
[111]	1	2.7	0.0072	J9W2T4	Heat shock protein, putative			
[112]	1	2.6	1.80E-04	J9VS02	**Endopeptidase, putative**	*		
[114]	1	5.3	0.002	J9VQN8	Phosphoglycerate kinase			
[115]	1	31.3	0.0012	J9VUT1	Putative uncharacterized protein		*	yes
[116]	1	2.2	0.0027	J9VQG1	Putative uncharacterized protein	*		
[117]	1	2.5	0.002	J9VJ75	Putative uncharacterized protein			
[118]	1	2.3	0.002	J9VHM5	Putative uncharacterized protein			
[119]	1	3.9	6.00E-04	J9VRV8	Putative uncharacterized protein		*	yes
[120]	1	6.6	0.0038	J9V0F1	Cytoplasm protein, putative		*	
[121]	1	5.1	0.0015	J9VNU4	Putative uncharacterized protein		*	
[122]	1	2.8	0.0026	J9VM09	Inorganic diphosphatase, putative			yes
[123]	1	9.5	0.0011	J9VWK2	Expressed protein		*	
[124]	1	5.9	0.0028	J9VDD8	ethylthioadenosine phosphorylase (MTAP), putative; Meu1p		*	
[125]	1	4.9	0.0034	J9VV69	Putative uncharacterized protein			
[126]	1	7.7	0.0048	J9VTE3	Putative uncharacterized protein		*	yes
[127]	1	6.7	0.0051	J9VSC8	Ribosomal protein L19			yes
[128]	1	7.3	0.0038	J9W2G5	Ribosomal protein L13, putative		*	yes
[129]	1	2.2	0.0078	na	Putative uncharacterized protein			
[130]	1	5.3	0.0044	J9VZY1	Putative uncharacterized protein	*		yes
[131]	1	2.3	0.0032	na	GPI-anchor transamidase, putative			
[132]	1	3.1	0.0064	J9VS68	Putative uncharacterized protein			
[136]	1	7.5	0.0064	J9VZ7	40S ribosomal protein S7, putative		*	
[138]	1	7	0.0041	J9W2U5	Putative uncharacterized protein			yes
[139]	1	2.3	0.0041	J9VNM1	Putative uncharacterized protein			
[140]	1	6.7	0.0089	J9VUB8	Cytokine inducing-glycoprotein, putative			yes

S 4.3 Table Expanded strain information. All strains used in this study are indicated, along with the CM number denoting their location in the Madhani laboratory strain database. "NatR" indicates nourseothricin resistance, under the column labeled source
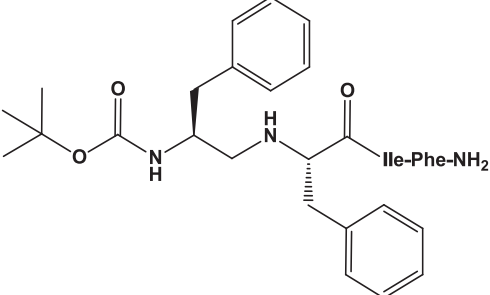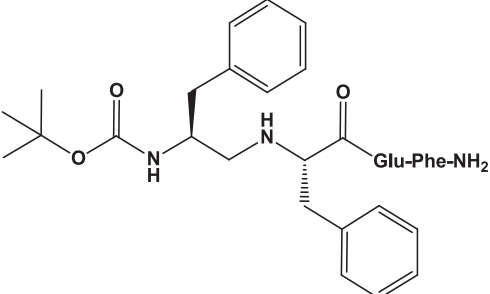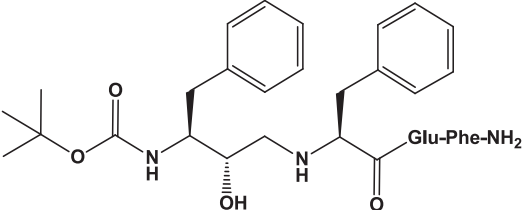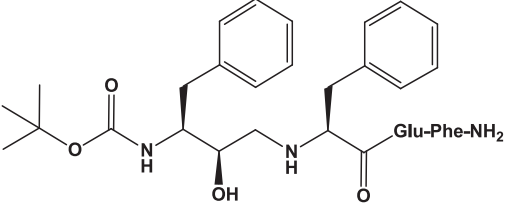"1" indicates a gift of J. Lodge, while "2" indicates this study.
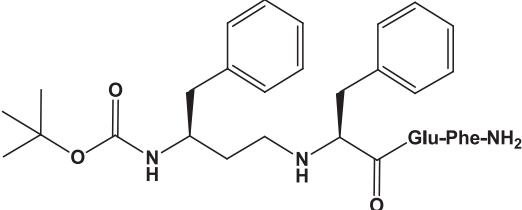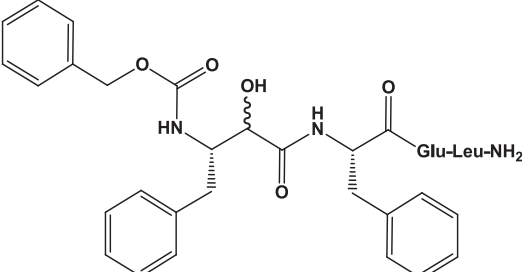
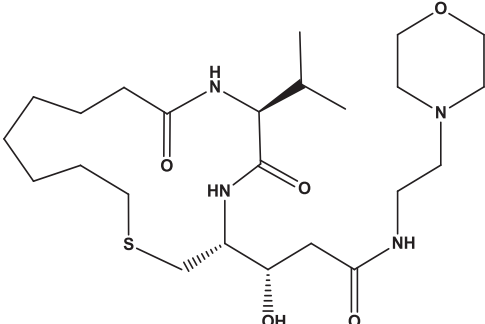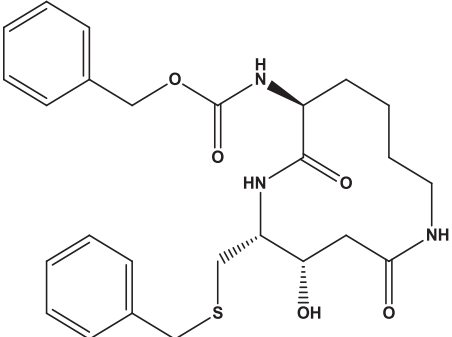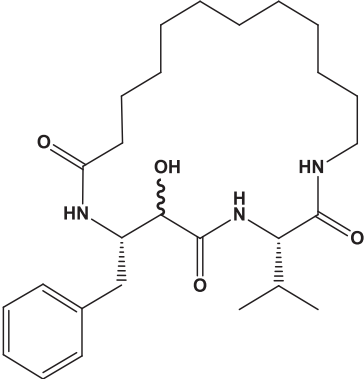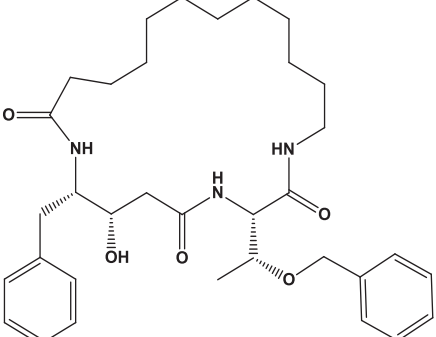| Genotype	CM#	Parent	Source	Peptidase Type	Citation
H99 (wild-type)	229	-	1	-	
CNAG_06814Δ::Nat^R	43	CM018		-	(59)
CNAG_05973Δ::Nat^R	1375	CM018	2	Serine carboxypeptidase	(27)
CNAG_05973Δ::Nat^R	1376	CM018	2		
CNAG_06640Δ::Nat^R	1386	CM018	2	Serine carboxypeptidase	This study
CNAG_06640Δ::Nat^R	1387	CM018	2		
CNAG_00919Δ::Nat^R	1398	CM018	2	Serine carboxypeptidase	(27)
CNAG_00919Δ::Nat^R	1399	CM018	2		
CNAG_01040Δ::Nat^R	1372	CM018	2	Serine carboxypeptidase	(27)
CNAG_01040Δ::Nat^R	1401	CM018	2		
CNAG_02966Δ::Nat^R	1380	CM018	2	Serine carboxypeptdase	(27)
CNAG_02966Δ::Nat^R	1381	CM018	2		
CNAG_00150Δ::Nat^R	1405	CM018	2	Serine endopeptidase	(27)
CNAG_00150Δ::Nat^R	1406	CM018	2		
CNAG_04625Δ::Nat^R	1373	CM018	2	Serine endopeptidase	(27)
CNAG_04625Δ::Nat^R	1374	CM018	2		
CNAG_00581Δ::Nat^R	1378	CM018	2	Aspartyl endopeptidase	(27)
CNAG_00581Δ::Nat^R	1379	CM018	2		
CNAG_05872Δ::Nat^R	1383	CM018	2	Aspartyl endopeptidase	This study
CNAG_05872Δ::Nat^R	1384	CM018	2		
CNAG_04735Δ::Nat^R	1369	CM018	2	Metallo endopeptidase	(27)
CNAG_04735Δ::Nat^R	1370	CM018	2		
CNAG_05872Flag::Nat^R	1669	CM026	2	Aspartyl endopeptidase	This study
		Sources:			
		1= Gift of J. Lodge			
		2= This study			
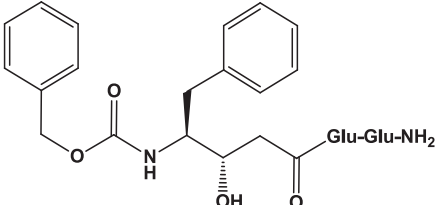
S 4.4 Table. Doubling times and saturation densities of strains grown in YNB. Values shown are averages of triplicates grown in 25-mL YNB cultures. The online doubling time calculator was used to estimate doubling times during the exponential growth phase [53].
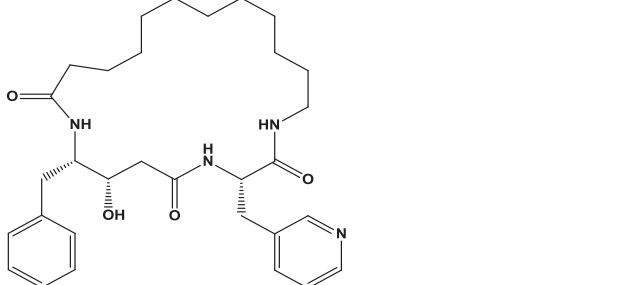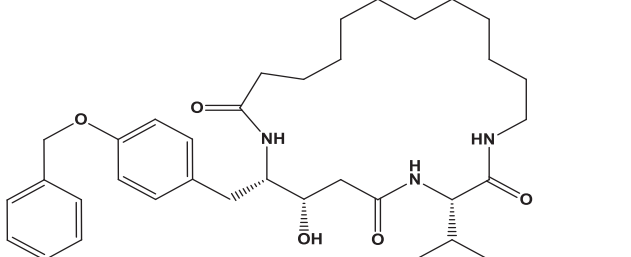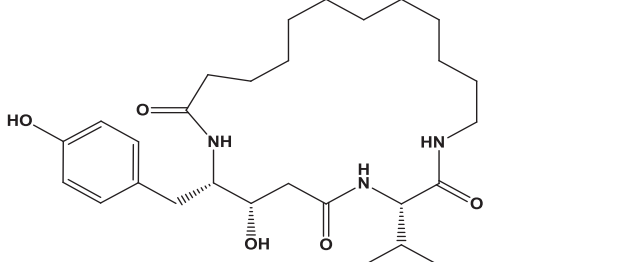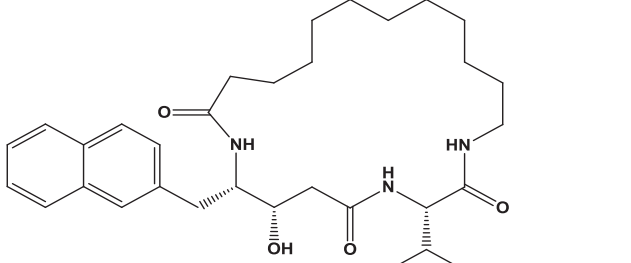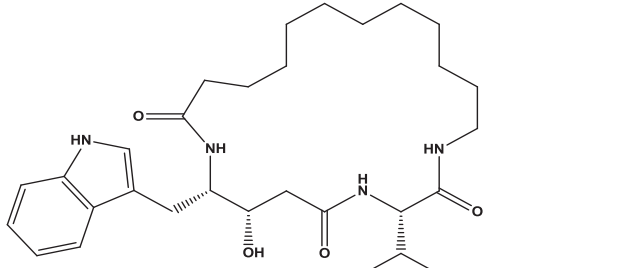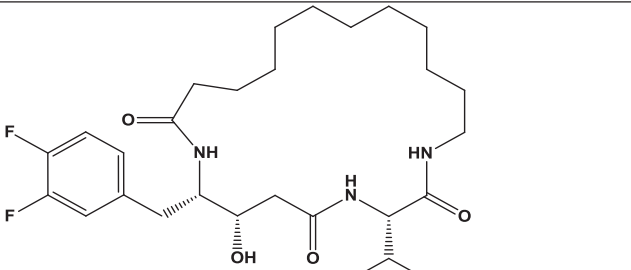
Strain	Doubling time (hours)	Final saturation density (OD_{600}) after 42 hours
Wild-type	2.03	17.40
may1Δ-1	2.48	9.92
may1Δ-2	2.23	9.88
scx1Δ	2.79	17.67
prc1Δ	2.25	18.83
cxd1Δ	2.16	17.45
cxd2Δ	2.78	15.58
cxd3Δ	2.41	17.85
pqp1Δ	2.37	16.72
prb1Δ	2.54	19.23
pep4Δ	2.61	18.25
mpr1Δ	2.72	15.87

S 4.6 Table. Structures of aspartyl peptidase inhibitors. The structures of the 21 peptidomimetic aspartyl peptidase inhibitors used in this study are shown, as well as their effectiveness at inhibiting May1 activity at a concentration of 1μM.

Compound	Structure	% May1 inhibition at 1uM
1		62
2		0
3		33
4		100

5	Boc-NH-CH₂-NH-...-Ile-Phe-NH₂	19
6	Boc-NH-CH₂-NH-...-Glu-Phe-NH₂	0
7	Boc-NH-...-OH...-NH-...-Glu-Phe-NH₂	13
8	Boc-NH-...-OH...-NH-...-Glu-Phe-NH₂	83
9	Boc-NH-...-NH-...-Glu-Phe-NH₂	4
10	Cbz-NH-...-OH...-NH-...-Glu-Leu-NH₂	32

11		68
12		18
13		42
14		52
15		83

16		100
17		61
18		100
19		87
20		77
21		100

Chapter 5. Convergent evolution of a peptide-based cell-cell signaling system required for the virulence of a eukaryotic pathogen

5.1 Forward

This publication represents the first observation of a peptide-based quorum sensing pathway in a eukaryotic organism, the opportunistic fungal pathogen, *Cryptococcus neoformans.* Quorum sensing has been well studied in bacteria and is important for regulating numerous virulence characteristics. In *C. neoformans* it appears that quorum sensing also impacts virulence. In studies using a mouse model of *C. neoformans* infection, fungal strains lacking the quorum sensing peptide were significantly less virulent. Quorum sensing appears to regulate multiple characteristics of *C. neoformans* cells including colony morphology and cell wall structure. However, it is unclear how these factors contribute to the hypovirulent phenotype seen in the mouse infection model.

My contribution to this report was the finding that secreted proteolytic activity is aberrant in strains unable to produce the quorum sensing peptide. These strains have decreased expression of a secreted aspartyl peptidase, (May1), which I identified and characterized as a significant contributor to virulence in chapter 4. Additionally, I determined that strains lacking the quorum sensing peptide over express metallo and serine peptidase activities when cultured in human cell-culture conditions. Although these proteolytic enzymes are thought to enhance virulence, we hypothesize that over expression at a high cell concentration could be deleterious to the fungal population.

Additionally, I conducted Multi-Plex Substrate Profiling by Mass Spectrometry (MSP-MS) to analyze the secreted proteolytic activity in supernatants from cultures lacking the quorum sensing peptide. Although these data were not included in the final manuscript, they supported

our conclusion that the level of peptidase secretion was affected by the quorum sensing pathway but not the type of secreted peptidase.

5.2 Abstract

Qsp1 is a peptide of previously unknown function secreted by the fungal meningitis pathogen *Cryptococcus neoformans*. We identified *QSP1* as a target of three transcription factors required for virulence. Indeed, mutants lacking Qsp1 are attenuated for infection, pulmonary accumulation and growth within macrophages. Qsp1 signaling modulates the activities of multiple secreted proteases and promotes cell wall organization and function at high cell densities. Production of the peptide requires its C-terminal release from a secreted precursor by a cell-associated protease, while sensing of Qsp1 requires an oligopeptide importer. The effects of cytoplasmic expression of Qsp1 suggest that it has at least one intracellular target. These features of biogenesis and reception closely mirror the quorum sensing systems of gram-positive bacteria in which peptide precursors are exported, processed and then imported to act intracellularly. Despite these remarkable similarities, the components of the respective systems are not ancestrally related, implying convergent evolution. Our studies reveal an unexpectedly bacterial-like signaling system that promotes the virulence of a eukaryotic pathogen.

5.3 Introduction

Disseminated fungal infections are among the most difficult to diagnose and treat, resulting in death in 30-50% of cases (Brown et al., 2012; Denning and Bromley, 2015; Ostrosky-Zeichner, 2012; Perfect, 2005). Compounding this challenge, antifungal drug development has been slow: only a single new class of antifungal agents, the echinocandins, has been introduced in the last two decades (Arathoon, 2001). While there are more than a

million fungal species estimated to exist (Gauthier and Keller, 2013), the majority of

disseminated infections are caused by a limited number of species, notably *Candida albicans*,

Aspergillus fumigatus, and *Cryptococcus neoformans* (Brown et al., 2012; Enoch et al., 2006).

Our understanding of the specializations that allow them to thrive within mammalian hosts is

primitive compared to our sophisticated understanding of bacterial pathogenesis.

Many bacterial pathogens require quorum sensing (QS) systems for virulence (Antunes

et al., 2010; de Kievit and Iglewski, 2000). In classical bacterial QS paradigms, cells secrete a

signal and monitor its concentration to sense cell density and/or diffusion (Fuqua and

Greenberg, 2002; Miller and Bassler, 2001; Ng and Bassler, 2009), and use this information to

regulate diverse biological functions (Whitehead et al., 2001). In gram-negative bacteria, the

most common secreted signals are acyl homoserine lactones (Eberhard, 1972; Fuqua et al.,

2001; Nealson et al., 1970), whereas gram-positive bacteria generally utilize peptides (Ng and

Bassler, 2009). These peptides are typically synthesized as precursors that are matured by

proteolytic processing and/or circularization (Claverys and Havarstein, 2002; Pomerantsev et

al., 2009; Tortosa et al., 2001). They are often sensed by transmembrane histidine kinases

(Rutherford and Bassler, 2012). However, in a number of gram-positive species (including

several *Bacillus* species, *Staphylococcus aureus, Streptococcus pneumoniae* and *Enteroccus

faecalis*), the response to the signal requires an oligopeptide permease (Lazazzera and

Grossman, 1998; Rutherford and Bassler, 2012). In systems for which mechanistic details are

available, it has been shown that the permease allows import of the peptide, which then binds

and modulates the activity of a transcription factor or intracellular phosphatase (Rocha-Estrada

et al., 2010; Rutherford and Bassler, 2012).

No AHL- or peptide-based QS systems have been identified in eukaryotes. However,

alcohols have been proposed as a fungal QS signal (Nickerson et al., 2006). *Saccharomyces

cerevisiae* cultures are well known for accumulating ethanol during fermentation, but also

accumulate isoamyl alcohol, isobutyl alcohol, and tyrosol (Hazelwood et al., 2008). Additionally, commercial strains of *S. cerevisiae* produce nerolidol and farnesol (Carrau et al., 2005). Cultures of *Candida albicans* also accumulate farnesol and tyrosol, and applying these alcohols directly to *C. albicans* cultures inhibits and promotes, respectively, the switch from yeast to hyphal growth (Chen et al., 2004; Hornby et al., 2001). Farnesol has additional effects on *C. albicans*, including growth inhibition, triggering cell death, and biofilm suppression (Kruppa, 2009). A mutant lacking a diacylglycerol pyrophosphatase homolog, Dpp3, displays dramatically reduced farnesol production and is attenuated for infection in mice (Navarathna et al., 2007). However, the *dpp3Δ* mutant did not display a change in hyphal switching during plate culture and it was not reported whether the mutant impacted density-dependent inhibition of filamentous growth (Navarathna et al., 2007). Furthermore, whether the *dpp3Δ* mutant has altered levels of other lipids predicted to be products of reactions catalyzed by this type of enzyme (phosphatidic acid, lysophosphatidic acid and diacylglycerol) (Toke et al., 1998a; Toke et al., 1998b) remains unknown. Additional effects of conditioned media consistent with QS-like phenomena in fungi have been reported, but the mechanistic details and functions remain to be understood (Albuquerque and Casadevall, 2012).

Cryptococcus neoformans, one of the key pathogens described above, is the most common cause of fungal meningitis. This single species causes 950,000 cases and 625,000 deaths annually, an estimated 40% of all HIV-related mortality (Armstrong-James et al., 2014). *C. neoformans* is part of a pathogen species complex that includes *C. neoformans var grubii* (serotype A), *C. neoformans var. neoformans* (serotype D), and *C. gattii* (serotypes B and C) (Kwon-Chung et al., 2014). *C. gattii* is thought to infect individuals who are less immunocompromised (Chen et al., 2014). Our studies focus on *C. neoformans var grubii*, which is responsible for ~90% of infections in the HIV/AIDS patient population (Mitchell and Perfect, 1995; Steenbergen and Casadevall, 2000). Investigations of *C. neoformans* pathogenesis have

focused on three primary virulence traits: melanin production (Nosanchuk and Casadevall, 2003), growth at human body temperature (Kraus et al., 2004), and the presence of a polysaccharide capsule with antiphagocytic and immunomodulatory properties (Doering, 2009).

A candidate QS signal in *C. neoformans* was identified in studies of a mutant *C. neoformans* Serotype D strain lacking the co-repressor Tup1 (Lee et al., 2007). The authors observed that *tup1Δ* strains, but not wild-type strains, failed to grow when a low number of cells were spread on agar plates. Purification of supernatants revealed an 11 amino acid peptide that was able to rescue this low-density plating defect. They suggested that this peptide, which was active at nanomolar concentrations, could be part of a QS system in *C. neoformans* and named it Qsp1 (Quorum Sensing-Like Peptide 1). Unfortunately, no phenotype for a *qsp1Δ* strain was reported and a similar low-density plating defect was not observed in a Serotype A strain lacking *TUP1* (Lee et al., 2009)*.* Thus, the normal function of Qsp1 remained unknown.

Through studies aimed at extending a virulence-associated regulatory network, we identified *QSP1* as a direct target of three transcription factors implicated in virulence. This finding, in addition to the importance of QS in bacterial virulence, prompted us to address the following outstanding questions: 1) Is Qsp1 produced in wild-type cultures? 2) Does Qsp1 accumulate with cell density, as is typical for bacterial QS signals? 3) Does Qsp1 have density-dependent functions? 4) Is Qsp1 required for virulence? and 5) What factors are required for cells to produce and respond to Qsp1? We show that Qsp1 is indeed produced by wild-type cells and accumulates with increasing cell density. We demonstrate that Qsp1 is required for virulence, pulmonary accumulation, and growth within macrophages. Furthermore, we find Qsp1 promotes normal colony morphology, cell wall ultrastructure and stress resistance, and modulates the activities of multiple secreted proteases. To identify additional candidate components of the Qsp1 pathway, we constructed a large library of targeted gene deletion strains and screened them using colony morphology as a phenotype. One hit corresponds to a

cell-associated serine protease that we demonstrate to be required for the production of Qsp1 from a secreted precursor. The screen further identified a predicted oligopeptide transporter, which we find is required for cells to respond to Qsp1, raising the possibility that Qsp1 might function after being imported into cells. Indeed, cytoplasmic expression of Qsp1 suggests that it has at least one intracellular target. These studies reveal a peptide-based QS system required for the virulence of an important human fungal pathogen, with remarkable similarity to a class of QS systems in gram-positive bacteria described above. However, the components identified here are unrelated to their bacterial counterparts, providing a remarkable example of the convergent evolution of an unusual cell-cell signaling paradigm.

5.4 Methods

Yeast strains Yeast strains used in this study are listed in Table S6. All *C. neoformans* strains were derived from the described parent using published procedures (Chun and Madhani, 2010). Gat201, Gat204, and Liv3 strains are derived from the CM18 parent for consistency with previous publications. The rest of the strains reported are derived from the KN99α wild type since that isolate is more closely related to the H99 clinical isolate (Janbon et al., 2014).

Chromatin immunoprecipitation   *C. neoformans* cultures were grown and immunoprecipitation was performed as previously published (Chun et al., 2011). Libraries were prepared for high throughput sequencing and analyzed using standard tools as described in Extended Experimental Procedures.

RNA expression profiling Total RNA was isolated and prepared as described previously (Dumesic et al., 2015). Analysis was performed as described in Extended Experimental Procedures.

Mass spectrometry Peptides were enriched in samples from culture supernatants by precipitation with 90% methanol and 1% acid. LC-MS/MS identification was performed as described in Extended Experimental Procedures.

Colony morphology patch complementation assays All assays were performed at room temperature unless otherwise specified. Peptides were synthesized by Peptide 2.0.

ELISA and immunoblotting assays Cell-free supernatants were analyzed by ELISA or immunoblotting as described in the Extended Experimental Procedures using 1:1000 diluted anti-Qsp1 serum.

Fluorescent substrate protease activity assay Protease assays were performed with the specified substrates as previously described (Jambunathan et al., 2012).

Growth, viability, capsule and melanin assays Growth on various medias, viability after incubation with stressors, capsule and melanin were analyzed as previously described (Liu et al., 2008) with minor changes described in the Extended Experimental Procedures.

Cell wall analysis 24-hr DMEM cultures were analyzed as previously described (Reese et al., 2007).

Intranasal infection model Mouse lung infections were performed as previously described (Chun and Madhani, 2010).

Rabbit meningitis model Rabbit intrathecal infections were performed as previously described (Perfect et al., 1980).

Leukocyte recruitment and cytokine analysis Lung leukocytes and cytokines were isolated and analyzed as previously described (Wiesner et al., 2015).

Macrophage interaction analysis Macrophage uptake and intracellular proliferation assays were performed as previously described (Nicola and Casadevall, 2012) with minor changes described in the Extended Experimental Procedures

5.5 Results

QSP1 is a direct target of three virulence regulators

We reasoned that virulence factors would be enriched in the regulatory targets of Gat201, a key transcription factor required for virulence and the inhibition of phagocytosis by macrophages (Chun et al., 2011; Liu et al., 2008). Our previous ChIP-Chip studies of Gat201 identified targets important for phagocytosis inhibition and infection (Chun et al., 2011). To develop an extended network, we employed the more sensitive and higher resolution ChIP-Seq method (Figure 5.1A). Previous studies of fungal pathogens suggest that genes whose promoters are bound by more than one transcription factor regulating a particular process often share that function (Beyhan et al., 2013; Nobile et al., 2012; Perez et al., 2013). Thus, we selected two additional transcription factors previously identified as Gat201 targets and implicated in virulence, Gat204 and Liv3 (Chun et al., 2011; Liu et al., 2008), for ChIP-Seq. All ChIP-Seq experiments were done in tissue culture conditions in which Gat201 has been shown to control transcription and to bind DNA (Chun et al., 2011).

Analysis revealed that Gat201, Gat204, and Liv3 constitute a transcription factor network, sharing characteristics of other published regulatory networks (Hughes and de Boer, 2013; Sorrells and Johnson, 2015). Gat201, Gat204, and Liv3 targets (Table S1) overlap significantly ($p < 10^{-300}$ by chi-squared test), with 225 shared by all three regulators, a group we refer to as 'central targets' (Figure 5.1B). Using RNA-Seq analysis of deletion mutants, we found that 42% of all binding events cause an increase or decrease in target transcript levels,

consistent with reported properties of other networks (Fisher et al., 2012; Nobile et al., 2012; Whitfield et al., 2012) (Figure S1A&B, Table S2). Each transcription factor binds its own promoter and the promoter of the other two regulators (Figure 5.1C), a characteristic common to many networks (Beyhan et al., 2013; Hughes and de Boer, 2013; Sorrells and Johnson, 2015). As predicted by their DNA-binding domains, de novo motif finding using the ChIP-Seq data (Figure 5.1D) shows that Gat201 and Gat204 have typical specificities of GATA family transcription factors (Ko and Engel, 1993) (Figure S1C). Liv3, a Wor1-ortholog, produced a DNA binding motif reminiscent of the *Candida albicans* Wor1 motif (Lohse et al., 2010) (Figure S1C).

Based on the reasoning above, we anticipated the central targets of the Gat201-Gat204-Liv3 network would be enriched for genes encoding virulence factors. A prominent target, *QSP1,* is a gene encoding a small secreted peptide described in the Introduction (Lee et al., 2007) (Figure 5.1E, Figure S1D). Strikingly, our RNA-Seq analysis revealed that *QSP1* produces the single most abundant transcript in cells grown in tissue culture conditions (Figure 5.1F). *QSP1* is predicted to encode a peptide precursor containing a signal sequence and a 24 amino acid pro-peptide (Qsp24) (Figure 5.1G). To determine whether this abundant transcript templates the production of high amounts of the mature peptide, we raised polyclonal rabbit antisera against Qsp1, the previously identified C-terminal 11 amino acids of the predicted precursor (Lee et al., 2007) (Figure S1E). Qsp1 is readily detected in wild-type culture supernatants by immunoblotting, producing a signal that comigrates with the synthetic peptide (Figure S1F). We next developed a quantitative ELISA to measure the concentration of Qsp1 in culture supernatants, and this analysis revealed that the peptide accumulates with cell density to micromolar concentrations (Figure 5.1H).

QSP1 is required for virulence

As a highly expressed, central target of the Gat201-Gat204-Liv3 network, *QSP1* was a strong candidate for a virulence factor. We first assessed whether the *qsp1Δ* mutant has a growth defect in vitro. However, consistent with prior work (Lee et al., 2007), we observed no defects in the growth of *qsp1Δ* cells after plating on media containing oxidative, nitrosative, pH, iron starvation, osmotic, or cell wall stressors (Figure S1G&H). Next, we assessed virulence of two independently derived *qsp1Δ* mutants using a well-established intranasal infection model in A/J mice (Chun and Madhani, 2010; Lim et al., 1980). We found that *qsp1Δ* mutants are significantly attenuated compared to wild type. Mean survival time for mice infected with wild-type *C. neoformans* was 14.6 days whereas mean survival times for mice infected with *qsp1Δ*-1 or *qsp1Δ*-2 mutants were 24.2 and 23.2 days, respectively (Figure 5.2A). Since A/J mice have a C5 complement deficiency (Cinader et al., 1964; Nilsson and Muller-Eberhard, 1967), we tested whether *qsp1Δ* mutants were also attenuated in an immunocompetent mouse strain, C57Bl/6 (Huffnagle et al., 1998). We obtained similar results in this mouse background: wild-type-infected mice displayed a mean survival time of 17.6 days, while *qsp1Δ*-1- and *qsp1Δ*-2-infected mice had mean survival times of 27.1 days and 25.4 days, respectively (Figure S2A).

In *C. neoformans*, virulence-attenuated mutants fall into several classes. One group, including *gat201Δ* and *gat204Δ* (Chun et al., 2011), comprises mutants deficient in early pulmonary accumulation. Another group, including a urease-deficient strain, *ure1Δ,* corresponds to mutants that accumulate normally in the lung but are deficient in dissemination (Olszewski et al., 2004). We quantified the cryptococcal viable cell counts (CFUs) in lung homogenates obtained from mice infected with either wild-type or *qsp1Δ* strains. Compared to wild type, *qsp1Δ* mutants accumulated to approximately three-fold fewer CFUs at day four after infection and to 20-fold fewer CFUs at day 13 (Figure 5.2B). Control experiments demonstrate that these defects were not due to a defect in initial seeding of the tissue (Figure S2B). In addition to fitness during early infection, we examined how *qsp1Δ* strains behaved during central nervous

system infection, using an intracisternal infection model in immunosuppressed rabbits (Perfect et al., 1980). This model has been used to identify mutants unable to establish chronic meningitis, which are rapidly cleared from the meninges (Lee et al., 2010; Petzold et al., 2006). However, no significant difference in cerebrospinal fluid CFUs was detected between wild type and *qsp1Δ* (Figure S2C), indicating similar fitness in the central nervous system.

Investigations of the cryptococcal *rim101Δ* mutant demonstrate that it alters the host immune response, thereby impacting its virulence (O'Meara et al., 2013; O'Meara et al., 2010). Cryptococcal infections that result in a Th2-dominant T-cell response tend to favor *C. neoformans* (Herring et al., 2004; Piehler et al., 2011), while Th1-dominant responses more effectively clear the fungus (Zhang et al., 2009). We therefore examined the inflammatory response during infection. We determined the number and type of leukocytes recruited to lungs after infection with wild-type or *qsp1Δ* strains using flow cytometry-based profiling of tissue homogenate (Figure 5.2C). Consistent with the lower pulmonary burden found after infection with *qsp1Δ* mutants, there were fewer leukocytes recruited. However, the majority of leukocytes were eosinophils, as in samples obtained from lungs of mice infected with wild type, indicating a Th2-polarized infection (Akuthota et al., 2008). We also assessed the level of cytokines in lungs and found no differences between wild-type and mutant infections at either early or late time points (Figure 5.2D). As predicted by the leukocyte recruitment data, the type 2 cytokines (IL-4, IL-5, and IL-13) (Herring et al., 2004) dominated at 3.5 days after infection and IL-13 was the most abundant cytokine 13 days after infection. These data are consistent with studies of T cell polarization during *C. neoformans* infections (Qiu et al., 2013; Zhang et al., 2010).

While the general Th1/Th2 polarization is unchanged, this finding does not preclude Qsp1 impacting the interaction between *C. neoformans* and specific leukocytes. In particular, we examined the interaction between the *qsp1Δ* mutant and macrophages, the most abundant immune cells in the alveolar space (Gordon and Read, 2002). Hosts susceptible to cryptococcal

infection display decreased survival when depleted of macrophages (Osterholzer et al., 2009; Shao et al., 2005) and macrophages contribute to cryptococcal dissemination (Charlier et al., 2009; Coelho et al., 2014). First, we measured phagocytosis rates of wild-type and *qsp1Δ* strains using bone marrow-derived macrophages (BMDMs) as described previously (Chun and Madhani, 2010; Nicola and Casadevall, 2012) (Figure S2D). We observed no difference between wild type and *qsp1Δ* mutants when cryptococcal cells were unopsonized (Figure 5.2E, Figure S2E) or opsonized with a monoclonal antibody against the capsule (Figure 5.2F, Figure S2E). We reasoned that *qsp1Δ* cells may have altered fitness after phagocytosis in the high-density environment of the phagolysosome (Carnes et al., 2010). Therefore, we measured the ability of the *qsp1Δ* mutant to accumulate within macrophages (Davis et al., 2015; Nicola and Casadevall, 2012) (Figure S2F). We observed that wild type accumulated intracellularly significantly faster than *qsp1Δ* cells, which displayed approximately two-fold fewer CFUs 24- and 48-hrs after infection (Figure 5.2G). These observations indicate that *qsp1Δ* cells are more readily controlled by BMDMs than wild type.

Intercellular signaling by Qsp1

Since *qsp1Δ* strains are attenuated for virulence, we continued to seek in vitro phenotypes for the mutant. We serendipitously observed that *qsp1Δ* mutants form dry, wrinkled colonies at room temperature, in contrast to smooth, mucoid colonies produced by wild type (Figure 5.3A). This colony phenotype provided an opportunity to assay Qsp1 function. Specifically, we developed a confrontation assay in which recipient strains were grown as colonies near a patch of donor strain cells. We found that when *qsp1Δ* cells are confronted by a wild-type patch, they produce smooth colonies (Figure 5.3B). We reasoned that this effect was due to Qsp1 diffusing from the donor. Indeed, synthetic Qsp1 was sufficient to cause *qsp1Δ*

cells to produce smooth colonies (Figure S3A). To determine whether this effect was specific to Qsp1, we synthesized and tested the predicted proQsp1 (Qsp24) and pro-region (Qsp13) peptides. We also synthesized a peptide in which the 11 amino acids of Qsp1 are scrambled (Scrambled Qsp1). We found that Qsp24 is sufficient to complement the *qsp1Δ* colony phenotype but Qsp13 and scrambled Qsp1 are not (Figure S3A), suggesting Qsp1 is sufficient and necessary for normal colony morphology. To assess structure-function relationships in Qsp1, we synthesized mutant peptides, including all possible single amino acid substitutions to alanine/glycine for large side-chains or to tryptophan for small side-chains. We also synthesized N- and C- terminal truncations of Qsp1 (Extended Experimental Procedures). N- and C-terminal truncations displayed no colony morphology-altering activity (data not shown). All but three positions in Qsp1 are sensitive to substitution (Figure S3B), indicating a strong requirement for specific residues for function.

Modulation of canonical virulence factors by Qsp1

Other cryptococcal mutants have been reported to form dry colonies, mostly mutants lacking capsule (Chang and Kwon-Chung, 1994, 1998) or mutants that are hyphal (Walton et al., 2006). However, *qsp1Δ* cells are exclusively yeast form and they stain normally with an anti-capsular monoclonal antibody (mAb339) (data not shown). *C. neoformans* cells display much larger capsules during mammalian infection (Rivera et al., 1998) than those seen under standard laboratory conditions (Zaragoza et al., 2003), and large capsules can be induced by specific in vitro conditions (Zaragoza and Casadevall, 2004). Therefore, we subjected wild-type and *qsp1Δ* cells to capsule-inducing conditions and quantified capsule size. At 37ºC, *qsp1Δ* mutants formed capsules that were similar to wild type. However, at room temperature, *qsp1Δ* mutant cells had larger capsules than wild type (Figure 5.3C, Figure S3C). While the capsule

enlargement of *qsp1Δ* mutants at room temperature is potentially informative for understanding their colony phenotype, we note that capsule size appears normal at 37ºC, the temperature of the mammalian host.

Cell wall-associated melanin is the second well-studied virulence factor of *C. neoformans*, and is typically assayed by spotting a dense patch of cells from a liquid culture on an agar plate containing L-DOPA, a phenolic precursor to melanin. *lac1Δlac2Δ* mutants lack both laccase enzymes responsible for melanin and therefore remain light-colored on L-DOPA plates. We found that *qsp1Δ* mutants displayed altered melanization at 37ºC, but the phenotype curiously depended on the condition of cultures used to inoculate the assay (Figure 5.3D). Using cells from saturated liquid cultures, we observed that *qsp1Δ* cells are hypomelanized and that this defect could be rescued by addition of synthetic Qsp1. Conversely, *qsp1Δ* cells grown on solid medium and then transferred to L-DOPA plates are hypermelanized (Figure 5.3D). Thus, Qsp1 is not essential for melanization and, under some conditions, negatively regulates this virulence factor. *lac1Δ* cells are attenuated for virulence but mouse lungs infected with this mutant show altered leukocyte recruitment and cytokine levels compared to wild type (Qiu et al., 2012). Therefore, we find it unlikely that *qsp1Δ* mutants' attenuated virulence is due solely to its melanin deficiency since it does not display an altered immune response.

qsp1Δ mutants display density-dependent cell wall defects

Changes in the cryptococcal cell wall can affect colony morphology (Siafakas et al., 2007), melanization (Baker et al., 2007; Banks et al., 2005), and capsule behavior (Baker et al., 2007; Reese and Doering, 2003; Reese et al., 2007). As described, *qsp1Δ* mutants do not display plating defects on agar containing the cell wall probes Calcofluor White, caffeine, or SDS. However, we reasoned that this assay, where single cells are deposited and then allowed

to form colonies, may not capture phenotypes that occur at high cell density. Therefore, we analyzed cell wall function in dense conditions by incubating saturated cultures with cell wall stressors for three hours and examining the viability of cells by plating. We found that *qsp1Δ* mutants display a dramatic loss of viability relative to wild type when exposed to high concentrations of SDS and caffeine (Figure 5.4A, Figure S4A). This phenotype could be rescued by growing cells with synthetic Qsp1 before the assay. In contrast, exponentially growing cultures display no difference in viability between the two genotypes. This density-dependent effect is specific, as we did not observe increased killing of saturated *qsp1Δ* cultures upon acidified nitrite or hydrogen peroxide treatment (Figure 5.4A, Figure S4A).

Since *qsp1Δ* mutants are sensitive to cell wall probes, we assessed cell wall structure of wild-type and *qsp1Δ* cells grown in tissue culture conditions by transmission electron microscopy as described previously (O'Meara et al., 2013) (Figure 5.4B). We observed that *qsp1Δ* cells have significantly thinner cell walls than wild type (Figure 5.4C). Previous studies have shown that wild-type cell walls are organized into distinct layers, typical of fungi (Doering, 2009; Free, 2013). The number and width of layers appears to be dynamic and dependent on conditions, as in other fungi (Sakaguchi et al., 1993; Shepardson and Cramer, 2013). We observed distinct layers in 90% of the wild-type cell walls examined. In contrast, we observed layers in about 60% of *qsp1Δ* cell walls (Figure 5.4D). Since the *qsp1Δ* phenotypes are reminiscent of mutants lacking chitin or chitosan, we examined chitin and chitosan content of cell walls biochemically but found that wild type and *qsp1Δ* mutants maintain similar profiles (Figure S4B).

Qsp1 impacts multiple secreted protease activities

Since numerous bacterial QS systems regulate secreted protease activities (Miller and Bassler, 2001; Redfield, 2002; Rutherford and Bassler, 2012), we examined the impact of Qsp1 on such activities in *C. neoformans*. In work that will be described elsewhere, we developed fluorogenic peptide substrates (Figure S4C) to assay secreted serine, aspartyl, and metalloprotease activities in *C. neoformans* and we identified the genes encoding each activity (S.C.C., Phillip A. Dumesic, C.M.H., H.D.M. and C.C., manuscript in preparation). In wild-type culture supernatants incubated in tissue culture conditions, secreted metalloprotease and serine endoprotease activity dominate while secreted aspartyl endoprotease activity dominates when cultures are grown in minimal media (S.C.C. et al, manuscript in preparation). *qsp1*Δ mutants have decreased secreted aspartyl endoprotease activity in minimal media and increased secreted serine endoprotease and metalloprotease activities in tissue culture conditions compared to wild type (Figure 5.4E). Growing cultures with 1 μM synthetic Qsp1 restored the aberrant altered serine and aspartyl endoprotease activities (Figure S4E).

Forward genetic screen for mutants in the *QSP1* pathway

We sought additional components of the Qsp1 signaling pathway. Since *QSP1* is part of the Gat201-Gat204-Liv3 network, we hypothesized that other network members may encode such components. Therefore, we attempted to construct a deletion mutant for each of the 1010 direct targets of Gat201, Gat204, and/or Liv3. We screened the 705 mutants successfully created for a dry colony phenotype and profiled melanin, capsule, and cell morphology for the nine mutants identified (Table 1). We screened an additional 1407 mutants produced while constructing a complete deletion collection for *C. neoformans* (all 2112 strains screened have been deposited to the Fungal Genetics Stock Center). Of these, we identified six additional mutants with the dry colony phenotype (Table S3), including three known acapsular mutants

[*cap10Δ, cap60Δ,* and *cap64Δ* (Chang and Kwon-Chung, 1998, 1999; Chang et al., 1996)] and one mutant defective in capsule attachment [*pbx1Δ* (Kumar et al., 2014; Liu et al., 2007)].

Pqp1 is a protease required for Qsp1 processing

Among the mutants with a dry colony phenotype was a predicted secreted subtilisin-like serine protease encoded by *CNAG_00150* (Figure 5.5A). Since only a peptide corresponding to the C-terminal 11 amino acids of the larger predicted proQsp1 peptide was detected in culture supernatants in prior studies (Lee et al., 2007), we hypothesized that *CNAG_00150* encodes the protease responsible for cleaving the predicted proQsp1 (Qsp24) species. Furthermore, we found that the colony phenotype of *qsp1Δpqp1Δ* mutant could be complemented by synthetic Qsp1 but not by synthetic Qsp24 (Figure 5.5B), suggesting that cells lacking Pqp1 are defective in processing synthetic Qsp24 into Qsp1 but are able to respond to Qsp1 itself. Below, we refer to this protease as proQsp1 Protease 1 (Pqp1).

Using synthetic Qsp1 and Qsp24, we determined that, while both peptides are recognized by the anti-Qsp1 antibody, they cannot be distinguished by mobility after SDS-PAGE and immunoblotting. To resolve these species, we used one-dimensional isoelectric focusing gel electrophoresis (IEFGE). In both rich and minimal media, we identified a band with identical mobility as synthetic Qsp1 in wild-type supernatants but that band was not detected in *pqp1Δ* or *qsp1Δ* supernatants. In *pqp1Δ* supernatants from rich media, we observed a band migrating at a more basic pI but its mobility differed from synthetic Qsp24 (Figure 5.5C). We hypothesize that the larger band appearing in *pqp1Δ* supernatants may be a modification to the predicted precursor. Since we observed that synthetic Qsp24 complements the colony morphology defect of *qsp1Δ* cells, this modification may not be necessary for processing. To test this explicitly, we added synthetic Qsp24 to *qsp1Δ* cultures and analyzed the products by IEFGE and

immunoblotting (Figure S5A). Consistent with processing, we observed the appearance of a band that comigrated with synthetic Qsp1. This band did not appear when synthetic Qsp24 was added to *qsp1Δpqp1Δ* cultures (Figure 5.5D).

To confirm these results, we analyzed peptide species in culture supernatants by LC-MS/MS and identified peaks (Figure 5.5E) and spectra (Figure 5.5F) corresponding to the Qsp1 species in samples generated from the supernatants of wild-type cultures. However, we did not observe peaks or spectra corresponding to Qsp1 in *qsp1Δ* or *pqp1Δ* culture supernatants (Figure 5.5E), supporting the model that Pqp1 is required for maturation of the endogenous precursor. No peak corresponding to Qsp24 was identified in culture supernatants from any genotype examined, again consistent with a modification of this putative precursor that has yet to be defined.

Pqp1 was previously identified as an extracellular but surface-associated protein (Eigenheer et al., 2007). Therefore, we quantified proQsp1 processing activity by cells and cell-free supernatants using a proQsp1-like substrate that contained the presumptive cleavage site between a fluorophore and a quencher (Figure S5B). We found that approximately 90% of protease activity against this substrate is cell-associated, although a small amount of activity accumulates in supernatants over time (Figure 5.5G). A small amount of residual cleavage (10-20% of wild type) was found in *pqp1Δ* mutant cells, presumably due to one or more unknown proteases that cleave at some position in the peptide (this assay does not measure cleavage at a specific site). To test whether any other proteases are essential for Qsp1 biogenesis or action, we analyzed nine additional mutants lacking predicted secreted proteases (Table S4), and found that none display a dry colony phenotype.

A predicted oligopeptide transporter required for peptide sensing: evidence for intracellular action

As described above, in gram-positive bacterial QS systems, the secreted peptide either binds a cell surface receptor or is imported back into the cell by multisubunit oligopeptide permeases (Opp), where it acts through an intracellular receptor (Lazazzera, 2001; Miller and Bassler, 2001). We began our search for downstream components by analyzing all published mutants in G protein coupled receptors and two-component receptors (Table S4), but none had the dry colony phenotype (data not shown). It was striking, however, that our screen identified a gene encoding a predicted oligopeptide transporter, which we named *OPT1* (Figure S6A). Although *OPT1* is phylogenetically unrelated to bacterial Opps, we hypothesized that it might play a similar role. Further suggesting a connection, the *QSP1* and *OPT1* genes are located next to each other in the genome and Gat201, Gat204, and Liv3 bind their promoters (Figure 5.6A). Significantly, *opt1Δ* mutants display nearly all of the phenotypes of *qsp1Δ* mutants. Like *qsp1Δ* mutants, *opt1Δ* mutants form dry colonies and are hypercapsular at room temperature (Figure 5.6B&C, Figure S6B). *opt1Δ* mutants display the same melanization pattern as *qsp1Δ* mutants at 37˚C (Figure S6C). Additionally, *opt1Δ* mutants display viability defects in saturated cultures exposed to SDS and caffeine that are indistinguishable from those of *qsp1Δ* mutants (Figure 5.6D, Figure S6D), despite no plating defects on any stress conditions tested (Figure S6E&F). Finally, *opt1Δ* mutants showed no difference in unopsonized (Figure 5.6E) or opsonized (Figure 5.6F) uptake by BMDMs compared to wild type, but have a two-fold intracellular proliferation defect after phagocytosis by BMDMs, similar to *qsp1Δ* mutants (Figure 5.6G).

Although the phenotypes of *opt1Δ* mutants closely mirror those of *qsp1Δ* mutants, *opt1Δ* mutants differ in two important respects: 1) They can produce Qsp1 and 2) They fail to respond to Qsp1. *opt1Δ* dry colonies cannot be complemented by wild-type patches (Figure 5.6H).

However, *opt1Δ* patches can complement nearby *qsp1Δ* colonies, indicating that *opt1Δ* mutants still produce Qsp1. Using our ELISA assay, we found that *opt1Δ* mutants produced wild-type amounts of Qsp1 in culture supernatants at cell densities up to OD 4.5. At higher densities, we observed less accumulation (Figure S6G). However, the *opt1Δ* colony morphology and the sensitivity of saturated *opt1Δ* cultures to cell wall stressors are not phenotypes due to a defect in Qsp1 production since they cannot be complemented by addition of synthetic Qsp1 (Figure 5.6I & Figure S6H).

In *B. subtilis*, the QS peptide CSF was initially inferred to have an intracellular site of action by the finding that expressing the mature peptide intracellularly could complement the phenotype of the mutant gene (Lazazzera et al., 1997). Therefore, we replaced the *QSP1* gene with an allele missing the signal sequence and the predicted pro region, but retaining an initiator methionine upstream of the mature Qsp1 sequence (iMet-QSP1) (Figure 5.6J). In addition, we created a strain harboring an N-terminal fusion of ubiquitin to the mature Qsp1 peptide, which is predicted to be cleaved by endogenous ubiquitin proteases to release the mature peptide (iUbi-Qsp1) (Varshavsky, 2000) (Figure 5.6J). As expected, these strains did not secrete Qsp1 in confrontation assays by immunoblotting of culture supernatants (data not shown), yet they displayed a shiny colony phenotype (Figure 5.6K). While additional studies (including identification of its receptor) will be required to definitively understand the reception of the Qsp1 signal, the requirement of *OPT1* for Qsp1 action, together with the result of the iMet-Qsp1 and iUbi-Qsp1 experiments, strongly suggest at least one of its receptors will be an intracellular protein.

Although *QSP1* and *OPT1* form a two-gene cluster in *C. neoformans*, they do not appear to be a recently acquired cluster produced by horizontal gene transfer. Our phylogenetic analysis of *OPT1* (Figure S6I) reveals its presence in the ancestor of present day *Tremellaceae*, the basidiomycete family that harbors the *Cryptococous* species complex. Moreover, we have

identified candidate Qsp1 precursors adjacent to *OPT1* in *Filobasidiella depauparata,*

Tsuchiyaea wingfieldii, and Cryptococcus amylolentus (Table S5)*.* These findings suggest that

the Qsp1 system evolved before the emergence of the pathogenic *Cryptococcus*

neoformans/*Cryptococcus gattii* species complex.

5.6 Discussion

Specialized biosynthetic pathways to produce specific signaling molecules and

dedicated mechanisms to detect signals characterize classical bacterial QS paradigms. The

cellular programs controlled by QS in prokaryotic systems are numerous and, in many

pathogens, include those required to cause disease. The latter feature has made them targets

for the development of new therapies. The widespread presence of prokaryotic QS systems

contrasts with our limited understanding of QS in eukaryotic microbes. If QS were generally

adaptive, one predicts systems of bacterial-level sophistication would have evolved in unicellular

eukaryotes and that these QS systems would promote their virulence.

Here, we characterize a peptide-based autoregulatory system in the fungal pathogen

Cryptococcus neoformans, a major driver of mortality in HIV/AIDS. We demonstrate that this

system impacts virulence, growth within macrophages, fungal cell wall integrity, and exoenzyme

activities. Our further analysis revealed remarkable similarities to an independently-evolved type

of QS found in *B. subtilis* and other gram-positive bacteria, involving the actions of an

extracellular protease and a peptide importer. These studies uncover a bacterial-like, peptide-

based signaling system required for virulence in a fungal pathogen.

A secreted peptide required for virulence in a eukaryotic pathogen

The original functional isolation of Qsp1 relied on a serotype-specific phenotype, namely that deletion of the Tup1 corepressor created a dependency on the secreted peptide for growth (Lee et al., 2007; Lee et al., 2009). However, phenotypes for the *qsp1Δ* mutant or whether wild-type cells produce Qsp1 was not reported. Our studies of three regulators implicated in virulence prompted us to further investigate this peptide since each of these regulators bound to the *QSP1* promoter. Strikingly, we found that *qsp1Δ* cells are substantially attenuated in the murine inhalation model and show a dramatic reduction in tissue burden in the lungs, the initial site of infection. These studies demonstrate a role for *QSP1* in virulence. Our initial work ruled out several trivial explanations for the attenuated virulence phenotype of *qsp1Δ* cells. They had no in vitro plating defects under any condition tested and accumulation in the host is not generally defective as *qsp1Δ* cell counts reached wild-type levels in the central nervous system infection. Additionally, our studies of the immune response to infection did not reveal any obvious change in the Th2 polarization, which is characteristic of cryptococcal disease and known to promote pathogen success.

Because macrophages are key determinants of the outcome of *C. neoformans* infection, we compared the interactions of wild-type and *qsp1Δ* cells with primary macrophages and observed a two-fold defect in intracellular accumulation of *qsp1Δ* cells compared to wild type. We observed no difference in phagocytosis rates, suggesting a specific role for Qsp1 after uptake. Macrophages are thought to have multiple functions in cryptococcal infection, including a protective role in killing *C. neoformans* as well as contributing to its dissemination. For both functions, the pathogen would presumably gain an advantage by optimizing survival in this cell type. The *qsp1Δ* mutant's intracellular accumulation defect could explain its virulence defect as other cryptococcal mutants with decreased accumulation inside macrophages are attenuated in an intranasal model of infection (Evans et al., 2015; Liu and Xue, 2014; Ma et al., 2009). An

intriguing possibility is that confinement in the phagolysosome leads to the accumulation of high

local concentrations of Qsp1, thereby triggering its action (Carnes et al., 2010).

Qsp1 impacts multiple phenotypes at high cell densities

As mentioned above, our analysis revealed no plating defects for *qsp1Δ* cells on

numerous stressors. These findings are anticipated if Qsp1 functions specifically at high cell

densities since colony plating assays, by definition, involve the placement of cells onto a surface

at low density. Indeed, a clue that *qsp1Δ* cells might have density-dependent phenotypes came

from visual analysis of yeast colonies and patches, where cells are packed at near-maximal

density. We used this colony phenotype to demonstrate that Qsp1 produced by cells could

complement recipient cells lacking Qsp1. The finding that synthetic peptides are sufficient to

complement the dry colony phenotype enabled structure-function analysis that defined key

functional residues.

This colony phenotype also suggested that Qsp1 mediated an alteration in the cell

surface. The outer surface of fungi is the cell wall, a dynamic and highly organized protective

organelle that impacts virtually all transactions with the environment. These findings led us to

examine cell wall ultrastructure, which revealed changes in thickness and organization in *qsp1Δ*

mutants. We observed a concomitant change in the resistance of *qsp1Δ* cells from saturated

cultures, but not those in lower-density cultures, to cell wall perturbants. These density-

dependent phenotypes of *qsp1Δ* cells are consistent with a model in which signaling occurs

primarily in high-density environments.

Finally, we observed Qsp1-dependent changes in all three major protease activities that

we have characterized in saturated cultures of *C. neoformans* (S.C. et al., manuscript in

preparation). Qsp1 promotes an aspartyl protease activity that we have found in other studies to

be required for virulence (S.C. et al., manuscript in preparation). In contrast, Qsp1 limits

metalloprotease and serine protease activities, which we speculate could be deleterious at high

cell densities. The mechanism of these changes in protease activities remains to be determined

as Qsp1 could impact the production of proteases, their activators, or their inhibitors.

Furthermore, understanding which Qsp1-dependent cellular processes are responsible for the

role of the peptide in virulence will require developing a detailed analysis of the underlying

mechanisms. As a first step towards this goal, we have begun to isolate factors involved in the

production and reception of Qsp1, as discussed below.

Production of the signal requires a cell-associated serine protease

We used the *qsp1Δ* rough colony phenotype to perform a forward genetic screen for

additional components of the pathway. As part of an ongoing effort to construct a complete gene

deletion collection, we prioritized targets of the Gat201-Gat204-Liv3 network but also included a

large number of additional genes. We ultimately constructed and screened 2112 gene deletion

strains and identified 15 mutants that form dry colonies. Among the corresponding genes was

one that encoded a predicted subtilase-like serine protease *CNAG_00150*, which we named

PQP1. Many lines of evidence support the view that Pqp1 is the protease that cleaves a Qsp1

precursor to produce the 11 amino acid mature peptide, Qsp1: 1) *pqp1Δ* cells fail to produce

Qsp1. 2) While *qsp1Δ*'s dry colony phenotype can be complemented by addition of either Qsp1

or a predicted precursor, proQsp1, only mature Qsp1 can complement *pqp1Δ* cells. 3) An

exogenously-added predicted Qsp1 precursor can be processed by a cell-associated activity

present in wild-type and *qsp1Δ* strains but not mutants lacking Pqp1. 4) Deletion of genes

coding for all other predicted secreted endoproteases did not yield any mutants with a dry

colony phenotype. Although reconstitution of Pqp1 activity using recombinant protein will

ultimately be required, our studies strongly imply that Pqp1 is the protease responsible for processing a Qsp1 precursor to the functional Qsp1 signaling molecule. Notably, in *B. subtilis*, the production of the mature QS peptide is mediated by three redundant cell-associated proteases, including a subtilisin superfamily member (Lanigan-Gerdes et al., 2007).

Reception of the signal requires a predicted oligopeptide transporter: evidence for intracellular action

The additional mutants we isolated by their dry colony phenotype were not complemented by synthetic Qsp1, making them candidates for factors involved in the response to Qsp1. Among these, one mutant *opt1Δ*, shares the full constellation of *qsp1Δ* phenotypes including temperature-dependent changes in colony morphology, capsule production, melanin changes, defective growth within macrophages, and density-dependent sensitivity to cell wall stressors. The notable difference is that, rather than being defective in producing the signal, the *opt1Δ* mutant is defective in responding to the signal. As the *opt1Δ* mutant precisely mirrors the *qsp1Δ* mutant but is defective in signal reception, we reason that Opt1 functions in the response to Qsp1. The predicted protein sequence of Opt1 places it in a well-characterized family of oligopeptide transporters (Gomolplitinant and Saier, 2011), leading us to hypothesize that Opt1 transports Qsp1 into cells. Work over the last two decades has shown that QS peptides in some gram-positive bacteria require a multisubunit oligopeptide permease (Opp) for their actions and, in these instances, Opp imports QS peptides into cells where they bind intracellular receptors. The model that Opt1 transports Qsp1 would explain why no candidates for extracellular receptors that we have tested display the dry colony phenotype. As an initial test of this hypothesis, we constructed strains to express the mature Qsp1 peptide intracellularly (iQsp1), analogous to what has been done previously in *B. subtilis* and *B. thurengiensis* (Lazazzera et

al., 1997; Slamti and Lereclus, 2002). As expected, these iQsp1 strains did not secrete Qsp1. Nonetheless, they displayed a wild-type colony morphology, indicating that the iQsp1 variants complement the null colony phenotype. Together with the requirement of a predicted oligopeptide transporter for all actions of Qsp1, these data suggest that Qsp1 has at least one receptor that is located intracellularly. Definitive tests of this hypothesis will require the biochemical identification of this receptor. Candidates for downstream elements include those identified in our dry colony screen, such as the Liv3 transcription factor.

Evolution of peptide-based autoregulatory signaling

Although *QSP1* and *OPT1* form a two-gene cluster in *C. neoformans*, they appear to have evolved before the emergence of the pathogenic *Cryptococcus neoformans*/*Cryptococcus gatii* species complex. Like some gram-positive bacteria including *B. subtilis*, the Qsp1 system includes the production of a precursor whose C-terminus is cleaved by a cell wall-associated peptidase to produce the mature peptide, and the requirement of an oligopeptide transporter for cells to sense the peptide. As none of the components appear to have a proximal bacterial ancestor, we surmise that the eukaryotic system evolved independently. How might this have occurred? Both eukaryotic and prokaryotic microbes secrete proteases to obtain peptide nutrients that are then taken up by peptide transporters. Presumably, self-made secreted proteins are processed by the same proteases to some extent. If a peptide resulting from digestion of a self-made protein were produced that also had an intracellular activity (e.g. binding to a protein and modulation of its activity), this would effectively result in the birth of an autoregulatory system. In this regard, it is worth noting that the predicted Qsp1-like sequences encoded by the *Tsuchiyaea wingfieldii and Cryptococcus amylolentus* genomes appear to be part of larger proteins, suggesting that the ancestral *QSP1* system may have involved the

production of a signaling peptide from proteolysis of a larger secreted protein.

5.7 References

Akuthota, P., Wang, H.B., Spencer, L.A., and Weller, P.F. (2008). Immunoregulatory roles of eosinophils: a new look at a familiar cell. Clin Exp Allergy *38*, 1254-1263.

Albuquerque, P., and Casadevall, A. (2012). Quorum sensing in fungi--a review. Med Mycol *50*, 337-345.

Antunes, L.C., Ferreira, R.B., Buckner, M.M., and Finlay, B.B. (2010). Quorum sensing in bacterial virulence. Microbiology *156*, 2271-2282.

Arathoon, E.G. (2001). Clinical efficacy of echinocandin antifungals. Curr Opin Infect Dis *14*, 685-691.

Armstrong-James, D., Meintjes, G., and Brown, G.D. (2014). A neglected epidemic: fungal infections in HIV/AIDS. Trends Microbiol *22*, 120-127.

Baker, L.G., Specht, C.A., Donlin, M.J., and Lodge, J.K. (2007). Chitosan, the deacetylated form of chitin, is necessary for cell wall integrity in Cryptococcus neoformans. Eukaryot Cell *6*, 855-867.

Banks, I.R., Specht, C.A., Donlin, M.J., Gerik, K.J., Levitz, S.M., and Lodge, J.K. (2005). A chitin synthase and its regulator protein are critical for chitosan production and growth of the fungal pathogen Cryptococcus neoformans. Eukaryot Cell *4*, 1902-1912.

Beyhan, S., Gutierrez, M., Voorhies, M., and Sil, A. (2013). A temperature-responsive network links cell shape and virulence traits in a primary fungal pathogen. PLoS Biol *11*, e1001614.

Brown, G.D., Denning, D.W., Gow, N.A., Levitz, S.M., Netea, M.G., and White, T.C. (2012). Hidden killers: human fungal infections. Sci Transl Med *4*, 165rv113.

Carnes, E.C., Lopez, D.M., Donegan, N.P., Cheung, A., Gresham, H., Timmins, G.S., and Brinker, C.J. (2010). Confinement-induced quorum sensing of individual Staphylococcus aureus bacteria. Nat Chem Biol *6*, 41-45.

Carrau, F.M., Medina, K., Boido, E., Farina, L., Gaggero, C., Dellacassa, E., Versini, G., and Henschke, P.A. (2005). De novo synthesis of monoterpenes by Saccharomyces cerevisiae wine yeasts. FEMS Microbiol Lett *243*, 107-115.

Chang, Y.C., and Kwon-Chung, K.J. (1994). Complementation of a capsule-deficient mutation of Cryptococcus neoformans restores its virulence. Mol Cell Biol *14*, 4912-4919.

Chang, Y.C., and Kwon-Chung, K.J. (1998). Isolation of the third capsule-associated gene, CAP60, required for virulence in Cryptococcus neoformans. Infect Immun *66*, 2230-2236.

Chang, Y.C., and Kwon-Chung, K.J. (1999). Isolation, characterization, and localization of a capsule-associated gene, CAP10, of Cryptococcus neoformans. J Bacteriol *181*, 5636-5643.

Chang, Y.C., Penoyer, L.A., and Kwon-Chung, K.J. (1996). The second capsule gene of cryptococcus neoformans, CAP64, is essential for virulence. Infect Immun *64*, 1977-1983.

Charlier, C., Nielsen, K., Daou, S., Brigitte, M., Chretien, F., and Dromer, F. (2009). Evidence of a role for monocytes in dissemination and brain invasion by Cryptococcus neoformans. Infect Immun *77*, 120-127.

Chen, H., Fujita, M., Feng, Q., Clardy, J., and Fink, G.R. (2004). Tyrosol is a quorum-sensing molecule in Candida albicans. Proc Natl Acad Sci U S A *101*, 5048-5052.

Chen, S.C., Meyer, W., and Sorrell, T.C. (2014). Cryptococcus gattii infections. Clin Microbiol Rev *27*, 980-1024.

Chun, C.D., Brown, J.C., and Madhani, H.D. (2011). A major role for capsule-independent phagocytosis-inhibitory mechanisms in mammalian infection by Cryptococcus neoformans. Cell Host Microbe *9*, 243-251.

Chun, C.D., and Madhani, H.D. (2010). Applying genetics and molecular biology to the study of the human pathogen Cryptococcus neoformans. Methods Enzymol *470*, 797-831.

Cinader, B., Dubiski, S., and Wardlaw, A.C. (1964). Distribution, Inheritance, and Properties of an Antigen, Mub1, and Its Relation to Hemolytic Complement. J Exp Med *120*, 897-924.

Claverys, J.P., and Havarstein, L.S. (2002). Extracellular-peptide control of competence for genetic transformation in Streptococcus pneumoniae. Front Biosci *7*, d1798-1814.

Coelho, C., Bocca, A.L., and Casadevall, A. (2014). The intracellular life of Cryptococcus neoformans. Annu Rev Pathol *9*, 219-238.

Davis, M.J., Eastman, A.J., Qiu, Y., Gregorka, B., Kozel, T.R., Osterholzer, J.J., Curtis, J.L., Swanson, J.A., and Olszewski, M.A. (2015). Cryptococcus neoformans-induced macrophage lysosome damage crucially contributes to fungal virulence. J Immunol *194*, 2219-2231.

de Kievit, T.R., and Iglewski, B.H. (2000). Bacterial quorum sensing in pathogenic relationships. Infect Immun *68*, 4839-4849.

Denning, D.W., and Bromley, M.J. (2015). Infectious Disease. How to bolster the antifungal pipeline. Science *347*, 1414-1416.

Doering, T.L. (2009). How sweet it is! Cell wall biogenesis and polysaccharide capsule formation in Cryptococcus neoformans. Annu Rev Microbiol *63*, 223-247.

Dumesic, P.A., Homer, C.M., Moresco, J.J., Pack, L.R., Shanle, E.K., Coyle, S.M., Strahl, B.D., Fujimori, D.G., Yates, J.R., 3rd, and Madhani, H.D. (2015). Product binding enforces the genomic specificity of a yeast polycomb repressive complex. Cell *160*, 204-218.

Eberhard, A. (1972). Inhibition and activation of bacterial luciferase synthesis. J Bacteriol *109*, 1101-1105.

Eigenheer, R.A., Jin Lee, Y., Blumwald, E., Phinney, B.S., and Gelli, A. (2007). Extracellular glycosylphosphatidylinositol-anchored mannoproteins and proteases of Cryptococcus neoformans. FEMS Yeast Res *7*, 499-510.

Enoch, D.A., Ludlam, H.A., and Brown, N.M. (2006). Invasive fungal infections: a review of epidemiology and management options. J Med Microbiol *55*, 809-818.

Evans, R.J., Li, Z., Hughes, W.S., Djordjevic, J.T., Nielsen, K., and May, R.C. (2015). Cryptococcal phospholipase B1 is required for intracellular proliferation and control of titan cell morphology during macrophage infection. Infect Immun *83*, 1296-1304.

Fisher, W.W., Li, J.J., Hammonds, A.S., Brown, J.B., Pfeiffer, B.D., Weiszmann, R., MacArthur, S., Thomas, S., Stamatoyannopoulos, J.A., Eisen, M.B.*, et al.* (2012). DNA regions bound at low occupancy by transcription factors do not drive patterned reporter gene expression in Drosophila. Proc Natl Acad Sci U S A *109*, 21330-21335.

Free, S.J. (2013). Fungal cell wall organization and biosynthesis. Adv Genet *81*, 33-82.

Fuqua, C., and Greenberg, E.P. (2002). Listening in on bacteria: acyl-homoserine lactone signalling. Nat Rev Mol Cell Biol *3*, 685-695.

Fuqua, C., Parsek, M.R., and Greenberg, E.P. (2001). Regulation of gene expression by cell-to-cell communication: acyl-homoserine lactone quorum sensing. Annu Rev Genet *35*, 439-468.

Gauthier, G.M., and Keller, N.P. (2013). Crossover fungal pathogens: the biology and pathogenesis of fungi capable of crossing kingdoms to infect plants and humans. Fungal Genet Biol *61*, 146-157.

Gomolplitinant, K.M., and Saier, M.H., Jr. (2011). Evolution of the oligopeptide transporter family. J Membr Biol *240*, 89-110.

Gordon, S.B., and Read, R.C. (2002). Macrophage defences against respiratory tract infections. Br Med Bull *61*, 45-61.

Hazelwood, L.A., Daran, J.M., van Maris, A.J., Pronk, J.T., and Dickinson, J.R. (2008). The Ehrlich pathway for fusel alcohol production: a century of research on Saccharomyces cerevisiae metabolism. Appl Environ Microbiol *74*, 2259-2266.

Herring, A.C., Hernandez, Y., Huffnagle, G.B., and Toews, G.B. (2004). Role and development of TH1/TH2 immune responses in the lungs. Semin Respir Crit Care Med *25*, 3-10.

Hornby, J.M., Jensen, E.C., Lisec, A.D., Tasto, J.J., Jahnke, B., Shoemaker, R., Dussault, P., and Nickerson, K.W. (2001). Quorum sensing in the dimorphic fungus Candida albicans is mediated by farnesol. Appl Environ Microbiol *67*, 2982-2992.

Huffnagle, G.B., Boyd, M.B., Street, N.E., and Lipscomb, M.F. (1998). IL-5 is required for eosinophil recruitment, crystal deposition, and mononuclear cell recruitment during a pulmonary Cryptococcus neoformans infection in genetically susceptible mice (C57BL/6). J Immunol *160*, 2393-2400.

Hughes, T.R., and de Boer, C.G. (2013). Mapping yeast transcriptional networks. Genetics *195*, 9-36.

Jambunathan, K., Watson, D.S., Kodukula, K., and Galande, A.K. (2012). Proteolytic fingerprinting of complex biological samples using combinatorial libraries of fluorogenic probes. Curr Protoc Protein Sci *Chapter 21*, Unit21 22.

Janbon, G., Ormerod, K.L., Paulet, D., Byrnes, E.J., 3rd, Yadav, V., Chatterjee, G., Mullapudi, N., Hon, C.C., Billmyre, R.B., Brunel, F.*, et al.* (2014). Analysis of the genome and

transcriptome of Cryptococcus neoformans var. grubii reveals complex RNA expression and microevolution leading to virulence attenuation. PLoS Genet *10*, e1004261.

Ko, L.J., and Engel, J.D. (1993). DNA-binding specificities of the GATA transcription factor family. Mol Cell Biol *13*, 4011-4022.

Kraus, P.R., Boily, M.J., Giles, S.S., Stajich, J.E., Allen, A., Cox, G.M., Dietrich, F.S., Perfect, J.R., and Heitman, J. (2004). Identification of Cryptococcus neoformans temperature-regulated genes with a genomic-DNA microarray. Eukaryot Cell *3*, 1249-1260.

Kruppa, M. (2009). Quorum sensing and Candida albicans. Mycoses *52*, 1-10.

Kumar, P., Heiss, C., Santiago-Tirado, F.H., Black, I., Azadi, P., and Doering, T.L. (2014). Pbx proteins in Cryptococcus neoformans cell wall remodeling and capsule assembly. Eukaryot Cell *13*, 560-571.

Kwon-Chung, K.J., Fraser, J.A., Doering, T.L., Wang, Z., Janbon, G., Idnurm, A., and Bahn, Y.S. (2014). Cryptococcus neoformans and Cryptococcus gattii, the etiologic agents of cryptococcosis. Cold Spring Harb Perspect Med *4*, a019760.

Lanigan-Gerdes, S., Dooley, A.N., Faull, K.F., and Lazazzera, B.A. (2007). Identification of subtilisin, Epr and Vpr as enzymes that produce CSF, an extracellular signalling peptide of Bacillus subtilis. Mol Microbiol *65*, 1321-1333.

Lazazzera, B.A. (2001). The intracellular function of extracellular signaling peptides. Peptides *22*, 1519-1527.

Lazazzera, B.A., and Grossman, A.D. (1998). The ins and outs of peptide signaling. Trends Microbiol *6*, 288-294.

Lazazzera, B.A., Solomon, J.M., and Grossman, A.D. (1997). An exported peptide functions intracellularly to contribute to cell density signaling in B. subtilis. Cell *89,* 917-925.

Lee, A., Toffaletti, D.L., Tenor, J., Soderblom, E.J., Thompson, J.W., Moseley, M.A., Price, M., and Perfect, J.R. (2010). Survival defects of Cryptococcus neoformans mutants exposed to human cerebrospinal fluid result in attenuated virulence in an experimental model of meningitis. Infect Immun *78,* 4213-4225.

Lee, H., Chang, Y.C., Nardone, G., and Kwon-Chung, K.J. (2007). TUP1 disruption in Cryptococcus neoformans uncovers a peptide-mediated density-dependent growth phenomenon that mimics quorum sensing. Mol Microbiol *64,* 591-601.

Lee, H., Chang, Y.C., Varma, A., and Kwon-Chung, K.J. (2009). Regulatory diversity of TUP1 in Cryptococcus neoformans. Eukaryot Cell *8,* 1901-1908.

Lim, T.S., Murphy, J.W., and Cauley, L.K. (1980). Host-etiological agent interactions in intranasally and intraperitoneally induced Cryptococcosis in mice. Infect Immun *29,* 633-641.

Liu, O.W., Chun, C.D., Chow, E.D., Chen, C., Madhani, H.D., and Noble, S.M. (2008). Systematic genetic analysis of virulence in the human fungal pathogen Cryptococcus neoformans. Cell *135,* 174-188.

Liu, O.W., Kelly, M.J., Chow, E.D., and Madhani, H.D. (2007). Parallel beta-helix proteins required for accurate capsule polysaccharide synthesis and virulence in the yeast Cryptococcus neoformans. Eukaryot Cell *6*, 630-640.

Liu, T.B., and Xue, C. (2014). Fbp1-mediated ubiquitin-proteasome pathway controls Cryptococcus neoformans virulence by regulating fungal intracellular growth in macrophages. Infect Immun *82*, 557-568.

Lohse, M.B., Zordan, R.E., Cain, C.W., and Johnson, A.D. (2010). Distinct class of DNA-binding domains is exemplified by a master regulator of phenotypic switching in Candida albicans. Proc Natl Acad Sci U S A *107*, 14105-14110.

Ma, H., Hagen, F., Stekel, D.J., Johnston, S.A., Sionov, E., Falk, R., Polacheck, I., Boekhout, T., and May, R.C. (2009). The fatal fungal outbreak on Vancouver Island is characterized by enhanced intracellular parasitism driven by mitochondrial regulation. Proc Natl Acad Sci U S A *106*, 12980-12985.

Miller, M.B., and Bassler, B.L. (2001). Quorum sensing in bacteria. Annu Rev Microbiol *55*, 165-199.

Mitchell, T.G., and Perfect, J.R. (1995). Cryptococcosis in the era of AIDS--100 years after the discovery of Cryptococcus neoformans. Clin Microbiol Rev *8*, 515-548.

Navarathna, D.H., Hornby, J.M., Krishnan, N., Parkhurst, A., Duhamel, G.E., and Nickerson, K.W. (2007). Effect of farnesol on a mouse model of systemic candidiasis, determined by use of a DPP3 knockout mutant of Candida albicans. Infect Immun *75*, 1609-1618.

Nealson, K.H., Platt, T., and Hastings, J.W. (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. J Bacteriol *104*, 313-322.

Ng, W.L., and Bassler, B.L. (2009). Bacterial quorum-sensing network architectures. Annu Rev Genet *43*, 197-222.

Nickerson, K.W., Atkin, A.L., and Hornby, J.M. (2006). Quorum sensing in dimorphic fungi: farnesol and beyond. Appl Environ Microbiol *72*, 3805-3813.

Nicola, A.M., and Casadevall, A. (2012). In vitro measurement of phagocytosis and killing of Cryptococcus neoformans by macrophages. Methods Mol Biol *844*, 189-197.

Nilsson, U.R., and Muller-Eberhard, H.J. (1967). Deficiency of the fifth component of complement in mice with an inherited complement defect. J Exp Med *125*, 1-16.

Nobile, C.J., Fox, E.P., Nett, J.E., Sorrells, T.R., Mitrovich, Q.M., Hernday, A.D., Tuch, B.B., Andes, D.R., and Johnson, A.D. (2012). A recently evolved transcriptional network controls biofilm development in Candida albicans. Cell *148*, 126-138.

Nosanchuk, J.D., and Casadevall, A. (2003). Budding of melanized Cryptococcus neoformans in the presence or absence of L-dopa. Microbiology *149*, 1945-1951.

O'Meara, T.R., Holmer, S.M., Selvig, K., Dietrich, F., and Alspaugh, J.A. (2013). Cryptococcus neoformans Rim101 is associated with cell wall remodeling and evasion of the host immune responses. MBio *4*.

O'Meara, T.R., Norton, D., Price, M.S., Hay, C., Clements, M.F., Nichols, C.B., and Alspaugh, J.A. (2010). Interaction of Cryptococcus neoformans Rim101 and protein kinase A regulates capsule. PLoS Pathog *6*, e1000776.

Olszewski, M.A., Noverr, M.C., Chen, G.H., Toews, G.B., Cox, G.M., Perfect, J.R., and Huffnagle, G.B. (2004). Urease expression by Cryptococcus neoformans promotes microvascular sequestration, thereby enhancing central nervous system invasion. Am J Pathol *164*, 1761-1771.

Osterholzer, J.J., Milam, J.E., Chen, G.H., Toews, G.B., Huffnagle, G.B., and Olszewski, M.A. (2009). Role of dendritic cells and alveolar macrophages in regulating early host defense against pulmonary infection with Cryptococcus neoformans. Infect Immun *77*, 3749-3758.

Ostrosky-Zeichner, L. (2012). Invasive mycoses: diagnostic challenges. Am J Med *125*, S14-24.

Perez, J.C., Kumamoto, C.A., and Johnson, A.D. (2013). Candida albicans commensalism and pathogenicity are intertwined traits directed by a tightly knit transcriptional regulatory circuit. PLoS Biol *11*, e1001510.

Perfect, J.R. (2005). Cryptococcus neoformans: a sugar-coated killer with designer genes. FEMS Immunol Med Microbiol *45*, 395-404.

Perfect, J.R., Lang, S.D., and Durack, D.T. (1980). Chronic cryptococcal meningitis: a new experimental model in rabbits. Am J Pathol *101*, 177-194.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods *8*, 785-786.

Petzold, E.W., Himmelreich, U., Mylonakis, E., Rude, T., Toffaletti, D., Cox, G.M., Miller, J.L., and Perfect, J.R. (2006). Characterization and regulation of the trehalose synthesis pathway and its importance in the pathogenicity of Cryptococcus neoformans. Infect Immun *74*, 5877-5887.

Piehler, D., Stenzel, W., Grahnert, A., Held, J., Richter, L., Kohler, G., Richter, T., Eschke, M., Alber, G., and Muller, U. (2011). Eosinophils contribute to IL-4 production and shape the T-helper cytokine profile and inflammatory response in pulmonary cryptococcosis. Am J Pathol *179*, 733-744.

Pomerantsev, A.P., Pomerantseva, O.M., Camp, A.S., Mukkamala, R., Goldman, S., and Leppla, S.H. (2009). PapR peptide maturation: role of the NprB protease in Bacillus cereus 569 PlcR/PapR global gene regulation. FEMS Immunol Med Microbiol *55*, 361-377.

Qiu, Y., Davis, M.J., Dayrit, J.K., Hadd, Z., Meister, D.L., Osterholzer, J.J., Williamson, P.R., and Olszewski, M.A. (2012). Immune modulation mediated by cryptococcal laccase promotes pulmonary growth and brain dissemination of virulent Cryptococcus neoformans in mice. PLoS One *7*, e47853.

Qiu, Y., Dayrit, J.K., Davis, M.J., Carolan, J.F., Osterholzer, J.J., Curtis, J.L., and Olszewski, M.A. (2013). Scavenger receptor A modulates the immune response to pulmonary Cryptococcus neoformans infection. J Immunol *191*, 238-248.

Redfield, R.J. (2002). Is quorum sensing a side effect of diffusion sensing? Trends Microbiol *10*, 365-370.

Reese, A.J., and Doering, T.L. (2003). Cell wall alpha-1,3-glucan is required to anchor the Cryptococcus neoformans capsule. Mol Microbiol *50*, 1401-1409.

Reese, A.J., Yoneda, A., Breger, J.A., Beauvais, A., Liu, H., Griffith, C.L., Bose, I., Kim, M.J., Skau, C., Yang, S.*, et al.* (2007). Loss of cell wall alpha(1-3) glucan affects Cryptococcus neoformans from ultrastructure to virulence. Mol Microbiol *63*, 1385-1398.

Rivera, J., Feldmesser, M., Cammer, M., and Casadevall, A. (1998). Organ-dependent variation of capsule thickness in Cryptococcus neoformans during experimental murine infection. Infect Immun *66*, 5027-5030.

Rocha-Estrada, J., Aceves-Diez, A.E., Guarneros, G., and de la Torre, M. (2010). The RNPP family of quorum-sensing proteins in Gram-positive bacteria. Appl Microbiol Biotechnol *87*, 913-923.

Rutherford, S.T., and Bassler, B.L. (2012). Bacterial quorum sensing: its role in virulence and possibilities for its control. Cold Spring Harb Perspect Med *2*.

Sakaguchi, N., Baba, T., Fukuzawa, M., and Ohno, S. (1993). Ultrastructural study of Cryptococcus neoformans by quick-freezing and deep-etching method. Mycopathologia *121*, 133-141.

Shao, X., Mednick, A., Alvarez, M., van Rooijen, N., Casadevall, A., and Goldman, D.L. (2005). An innate immune system cell is a major determinant of species-related susceptibility differences to fungal pneumonia. J Immunol *175*, 3244-3251.

Shepardson, K.M., and Cramer, R.A. (2013). Fungal cell wall dynamics and infection site microenvironments: signal integration and infection outcome. Curr Opin Microbiol *16*, 385-390.

Siafakas, A.R., Sorrell, T.C., Wright, L.C., Wilson, C., Larsen, M., Boadle, R., Williamson, P.R., and Djordjevic, J.T. (2007). Cell wall-linked cryptococcal phospholipase B1 is a source of secreted enzyme and a determinant of cell wall integrity. J Biol Chem *282*, 37508-37514.

Slamti, L., and Lereclus, D. (2002). A cell-cell signaling peptide activates the PlcR virulence regulon in bacteria of the Bacillus cereus group. EMBO J *21*, 4550-4559.

Sorrells, T.R., and Johnson, A.D. (2015). Making Sense of Transcription Networks. Cell *161*, 714-723.

Steenbergen, J.N., and Casadevall, A. (2000). Prevalence of Cryptococcus neoformans var. neoformans (Serotype D) and Cryptococcus neoformans var. grubii (Serotype A) isolates in New York City. J Clin Microbiol *38*, 1974-1976.

Toke, D.A., Bennett, W.L., Dillon, D.A., Wu, W.I., Chen, X., Ostrander, D.B., Oshiro, J., Cremesti, A., Voelker, D.R., Fischl, A.S.*, et al.* (1998a). Isolation and characterization of the Saccharomyces cerevisiae DPP1 gene encoding diacylglycerol pyrophosphate phosphatase. J Biol Chem *273*, 3278-3284.

Toke, D.A., Bennett, W.L., Oshiro, J., Wu, W.I., Voelker, D.R., and Carman, G.M. (1998b). Isolation and characterization of the Saccharomyces cerevisiae LPP1 gene encoding a Mg2+-independent phosphatidate phosphatase. J Biol Chem *273*, 14331-14338.

Tortosa, P., Logsdon, L., Kraigher, B., Itoh, Y., Mandic-Mulec, I., and Dubnau, D. (2001). Specificity and genetic polymorphism of the Bacillus competence quorum-sensing system. J Bacteriol *183*, 451-460.

Varshavsky, A. (2000). Ubiquitin fusion technique and its descendants. Methods Enzymol *327*, 578-593.

Walton, F.J., Heitman, J., and Idnurm, A. (2006). Conserved elements of the RAM signaling pathway establish cell polarity in the basidiomycete Cryptococcus neoformans in a divergent fashion from other fungi. Mol Biol Cell *17*, 3768-3780.

Whitehead, N.A., Barnard, A.M., Slater, H., Simpson, N.J., and Salmond, G.P. (2001). Quorum-sensing in Gram-negative bacteria. FEMS Microbiol Rev *25*, 365-404.

Whitfield, T.W., Wang, J., Collins, P.J., Partridge, E.C., Aldred, S.F., Trinklein, N.D., Myers, R.M., and Weng, Z. (2012). Functional analysis of transcription factor binding sites in human promoters. Genome Biol *13*, R50.

Wiesner, D.L., Specht, C.A., Lee, C.K., Smith, K.D., Mukaremera, L., Lee, S.T., Lee, C.G., Elias, J.A., Nielsen, J.N., Boulware, D.R.*, et al.* (2015). Chitin recognition via chitotriosidase promotes pathologic type-2 helper T cell responses to cryptococcal infection. PLoS Pathog *11*, e1004701.

Zaragoza, O., and Casadevall, A. (2004). Experimental modulation of capsule size in Cryptococcus neoformans. Biol Proced Online *6*, 10-15.

Zaragoza, O., Fries, B.C., and Casadevall, A. (2003). Induction of capsule growth in Cryptococcus neoformans by mammalian serum and CO(2). Infect Immun *71*, 6155-6164.

Zhang, Y., Wang, F., Bhan, U., Huffnagle, G.B., Toews, G.B., Standiford, T.J., and Olszewski, M.A. (2010). TLR9 signaling is required for generation of the adaptive immune protection in Cryptococcus neoformans-infected lungs. Am J Pathol *177*, 754-765.

Zhang, Y., Wang, F., Tompkins, K.C., McNamara, A., Jain, A.V., Moore, B.B., Toews, G.B., Huffnagle, G.B., and Olszewski, M.A. (2009). Robust Th1 and Th17 immunity supports pulmonary clearance but cannot prevent systemic dissemination of highly virulent Cryptococcus neoformans H99. Am J Pathol *175*, 2489-2500.

5.8 Figures

Figure 5.1: _QSP1 is a direct target of the Gat201-Gat204-Liv3 regulatory network._
a) Representative ChIP-Seq data from chromosome 8. The y-axis represents number of reads.
b) Network diagram showing promoters bound by each regulator.
c) ChIP-Seq traces for Gat201, Gat204, and Liv3 at the promoter of each regulator. The y-axis represents number of reads.
d) Binding motifs derived from Gat201, Gat204, and Liv3 ChIP-Seq datasets.
e) ChIP-Seq data for Gat201, Gat204, and Liv3 at the promoter of _QSP1._
f) Plot of the rank and RPKM for each gene in a wild-type strain grown in tissue cultures conditions.
g) Schematic of predicted precursor encoded by _QSP1_. Signal sequence predicted by SignalP (Petersen et al., 2011).
h) Qsp1 accumulation in wild-type culture supernatants as measured by ELISA. Inset: Qsp1 accumulation at lower cell densities.

Figure 5.2: Qsp1 is required for virulence and accumulation within macrophages.
a) Survival analysis. Ten A/J mice were infected per genotype. Significance was determined using a log rank test. ** indicates $p < 10^{-4}$.
b) Lung burden analysis. Nine A/J mice were infected per genotype. At specified times, lung CFUs of three mice per genotype were measured and CFUs for each _qsp1Δ_ mutant were compared to wild type by student t-test. ** indicates $p<10^{-3}$ and *** indicates $p<10^{-4}$.
c) Leukocyte recruitment. Four C57BL/6 mice were infected with wild type and eight C57BL/6 mice per _qsp1Δ_ strains. 18 days after infection, number and class of immune cells present in the lungs were analyzed by flow cytometry.
d) Cytokine analysis. Ten C57BL/6 mice were infected per genotype. Full lungs were analyzed for cytokines at specified timepoints. Error bars are S.D.
e) Phagocytic index of unopsonized and f) opsonized _C. neoformans_. >1200 BMDMs quantified per genotype. Error bars are S.D.
g) Intracellular accumulation. CFUs isolated from BMDM cell lysates at specified times after infection, normalized to starting inoculum. Error bars represent 95% confidence intervals constructed by bootstrapping. Bootstrap analysis was used to test the difference between wild type and each _qsp1Δ_ mutant. * indicates a result with p value < 0.05.

Figure 5.3: Impact of Qsp1 on virulence-related phenotypes.
a) Colony morphology of wild-type and _qsp1Δ_ strains grown at specified temperatures.
b) Complementation assays. Patches of donor cells were grown at room temperature next to single colonies of recipient cells.
c) Capsule quantification. Representative pictures of cryptococcal capsule stained with india ink. >100 cells per genotype used for capsule size determination. Mutants were compared to wild type using student t-tests. *** indicates $p < 10^{-5}$. Error bars are S.E.
d) Melanin assay. All assays performed on L-DOPA plates at 37°C. Left: Cells were grown in liquid culture with or without synthetic peptides before spotting. Right: Patches were grown on sterile filters and transferred to L-DOPA plates.

Figure 5.4: Qsp1 impacts cell wall integrity and secreted protease activities
a) Stress assays. Saturated and log phase cultures were incubated with indicated levels of stressors and plated for viability.
b) Transmission electron microscopy. Representative micrographs of wild-type and _qsp1Δ_ cells from tissue culture conditions.

c) Quantification of cell wall width and d) the percentage of cells with visible cell wall layers. 50 cells per genotype. Mutants were compared to wild type by student t-tests. ** indicates p < 0.001 and * indicates p < 0.01. Error bars are S.D.

d) Secreted protease activities. *qsp1Δ* mutants were compared to wild type by student t-tests. ** indicates p < 0.001 and * indicates p < 0.01. Error bars are S.D.

Figure 5.5: Pqp1 is required for Qsp1 precursor processing.
   a) Predicted structure of Pqp1 protein.
   b) Spotting assays. Cultures were spotted with specified synthetic peptide stock and grown at room temperature.
   c) Qsp1 secretion assays. Isoelectric focusing gel electrophoresis and immunoblotting against Qsp1/Qsp24 in culture supernatants. Liquid chromatography of culture supernatants.
   d) Predicted Qsp1 precursor cleavage assays. Synthetic Qsp24 was incubated with cultures of specified strains. Supernatants were isolated and then analyzed for cleavage by isoelectric focusing gel electrophoresis and immunoblotting against Qsp1/Qsp24.
   e) Liquid chromatography of culture supernatants.
   f) Qsp1 MS2 spectra. Mass spectrometry of Qsp1 showing the fragment ions identified. Each fragment detected is shown with red (y ions) or blue (b ions) lines on the peptide sequence.
   g) Pqp1 activity assay. Shown are rates of cleavage of Qsp1-like substrate by cells or cell-free supernatants in wild-type and *pqp1Δ* cultures. Error bars are S.D.

Figure 5.6: A predicted oligopeptide transporter is required for cells to respond to Qsp1.
   a) Transcription factor binding sites in intergenic regions surrounding *QSP1* and *OPT1*.
   b) Colony morphology of the *opt1Δ* mutant.
   c) Capsule sizes. Representative pictures of cryptococcal capsule stained with india ink. >100 cells per genotype were used in quantification. Mutants were compared to wild type using student t-tests. *** indicates $p < 10^{-5}$. Wild-type data from Figure 3 included for comparison. Error bars are S.E.
   d) Stress assays. Saturated and log phase cultures were incubated with indicated levels of stressors and plated for viability. Wild-type data from Figure 4 included for comparison.
   e) Phagocytic index of unopsonized and f) opsonized *C. neoformans.* >1200 BMDMs quantified per genotype. Wild-type data from Figure 2 is included for comparison. Error bars are S.D.
   g) Macrophage intracellular accumulation assay. CFUs isolated from BMDM cell at specified times after infection, normalized to starting inoculum. Wild-type data from Figure 2 included for comparison. Error bars represent 95% confidence intervals constructed by bootstrapping. Bootstrap analysis was used to test the difference between wild type and mutants. * indicates p value < 0.05.
   h) Complementation assays. Patches of donor cells are grown at room temperature next to single colonies of recipient cells.
   i) Stress assays: Saturated wild-type and *opt1Δ* cultures were grown in the absence or presence of synthetic peptides and incubated with SDS or caffeine and plated for viability.
   j) Internal Qsp1 Strain Constructs. Schematic of the wild-type *QSP1* locus and the design of both internal Qsp1 expression constructs.
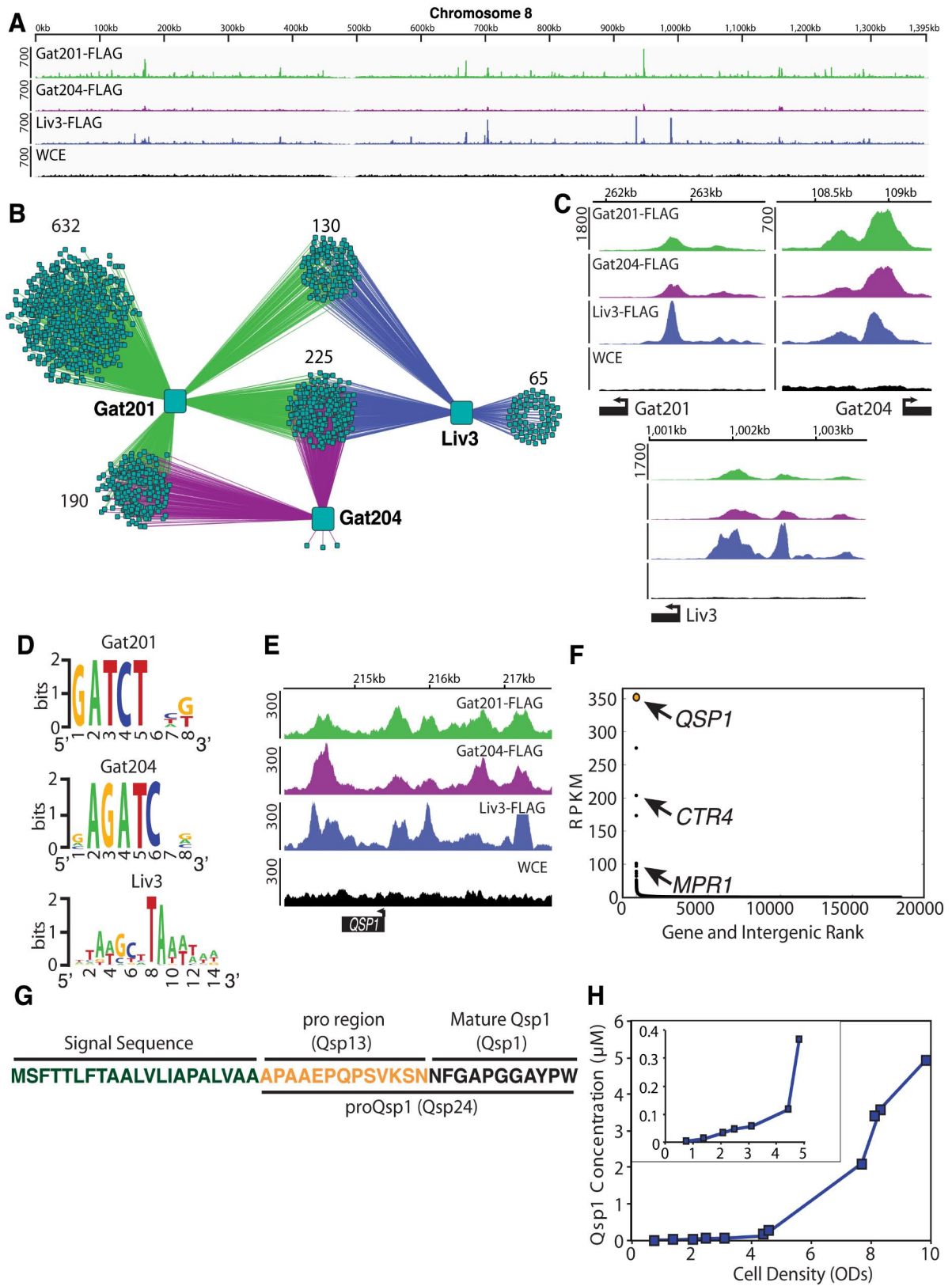   k) Colony morphology of internal Qsp1 strains.
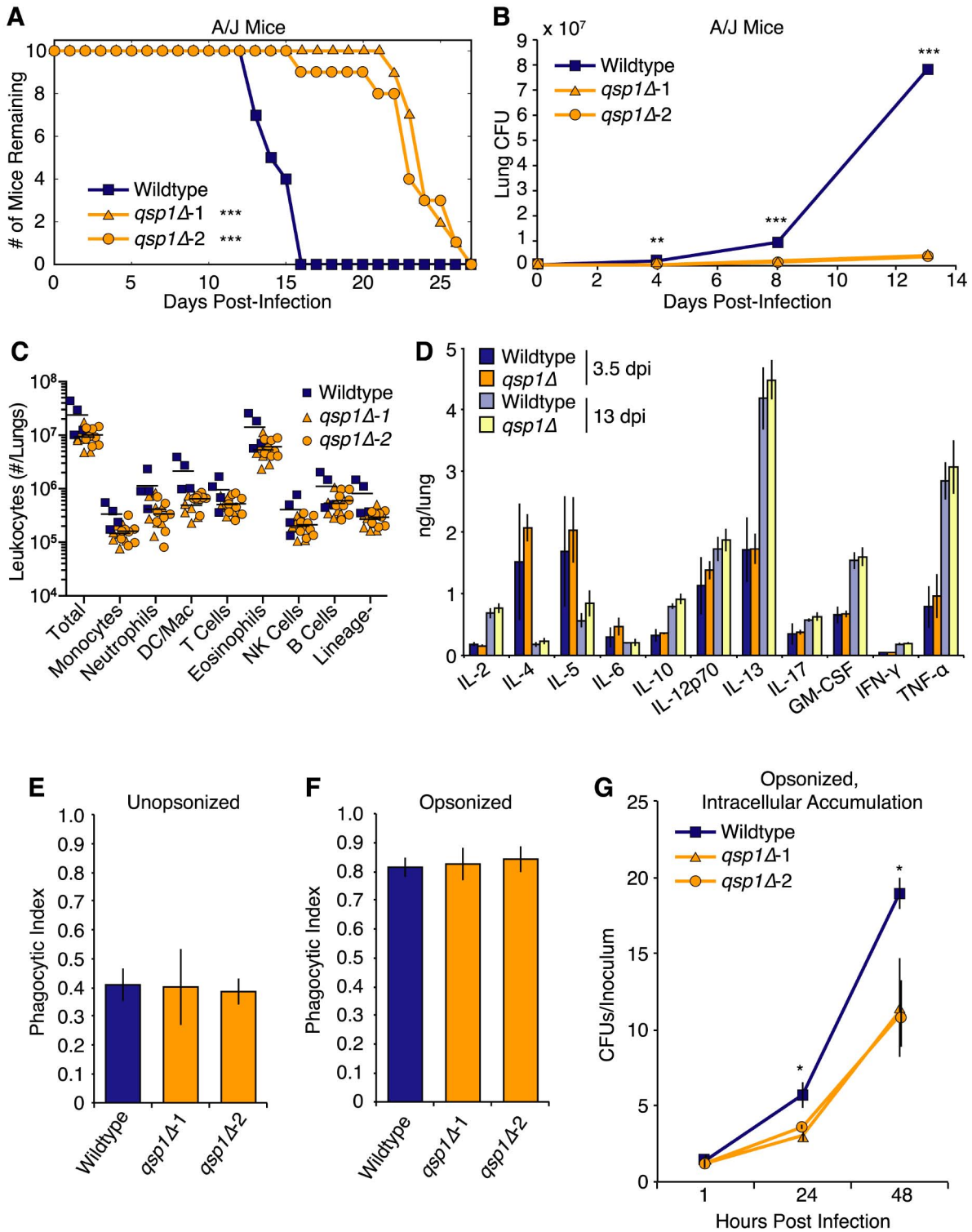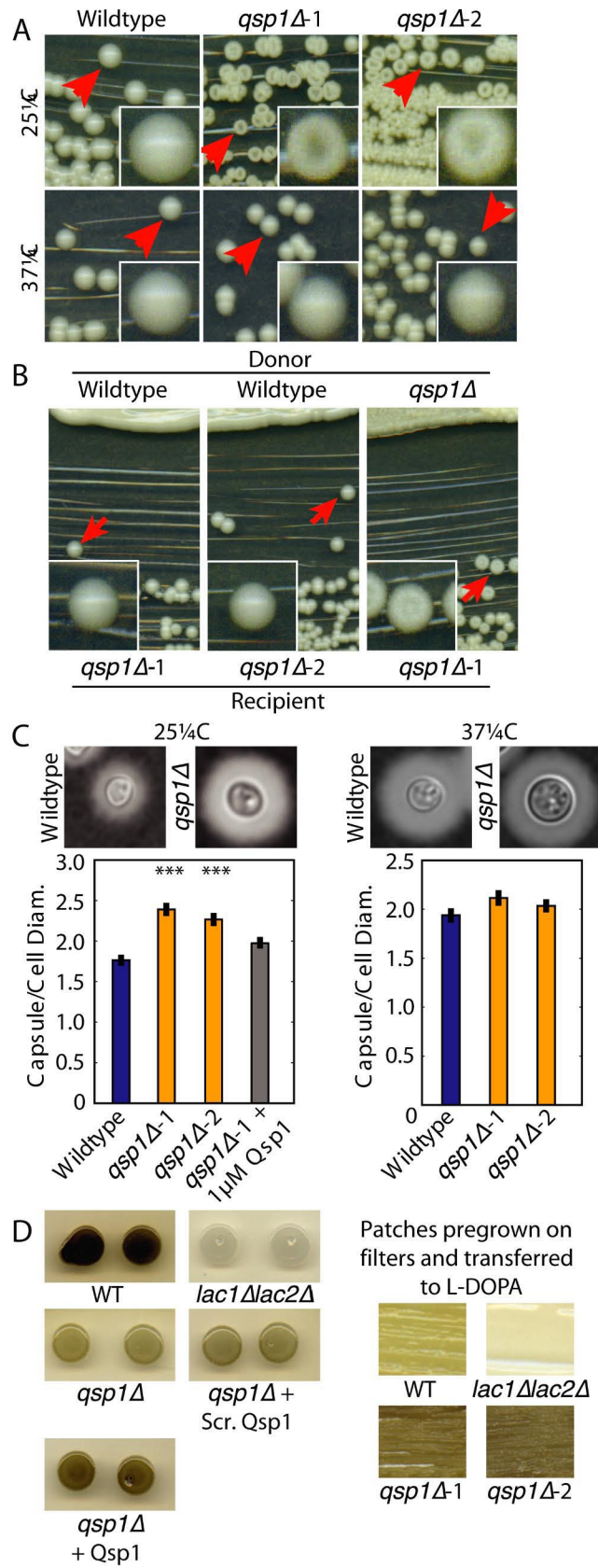
Figure 5.1

Figure 5.2

Figure 5.3

A

% SDS   Caffeine (mg/mL)   Nitrite (M)   H₂O₂ (%)

0.06 0.12 0.23 0.47 0.94 1.9 3.75 7.50   0.59 1.17 2.34 4.69 9.38 18.8 37.5 75.0   0.02 0.04 0.07 0.14 0.28 0.56 1.12 2.25   0.20 0.40 0.70 1.40 2.80 5.60 11.2 22.5

qsp1Δ-1
qsp1Δ-2
qsp1Δ-1+1μM Qsp1
Wildtype

Saturated Log Phase Cultures

qsp1Δ-1
qsp1Δ-2
qsp1Δ-1+1μM Qsp1
Wildtype

0.02 0.04 0.07 0.14 0.28 0.56 1.12 2.25
H₂O₂ (%)

B

Wildtype

qsp1Δ-1

qsp1Δ-2

C

Cell Wall Width (A.U.)

0.8
0.6
0.4
0.2
0

**
**

D

% Walls with Layers

1.0
0.8
0.6
0.4
0.2
0

**
*

■ Wildtype
■ qsp1Δ-1
▨ qsp1Δ-2

E

Type of Secreted Endoprotease Activity

Metallo   Aspartyl   Serine

RFU/Sec

0.3   0.3   0.9
0.2   0.2   0.6
0.1   0.1   0.3
0   0   0

■ WT
■ qsp1Δ-1
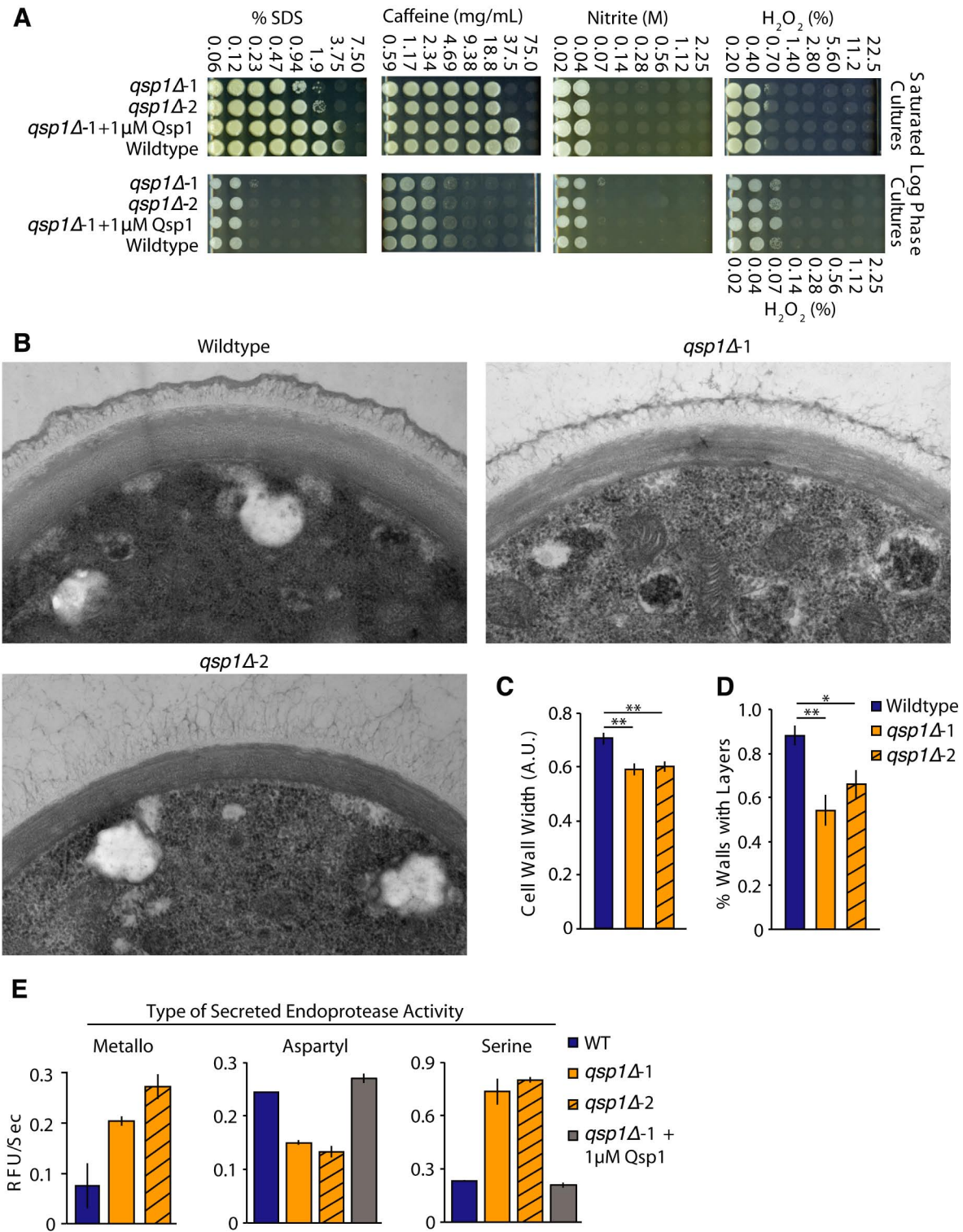▨ qsp1Δ-2
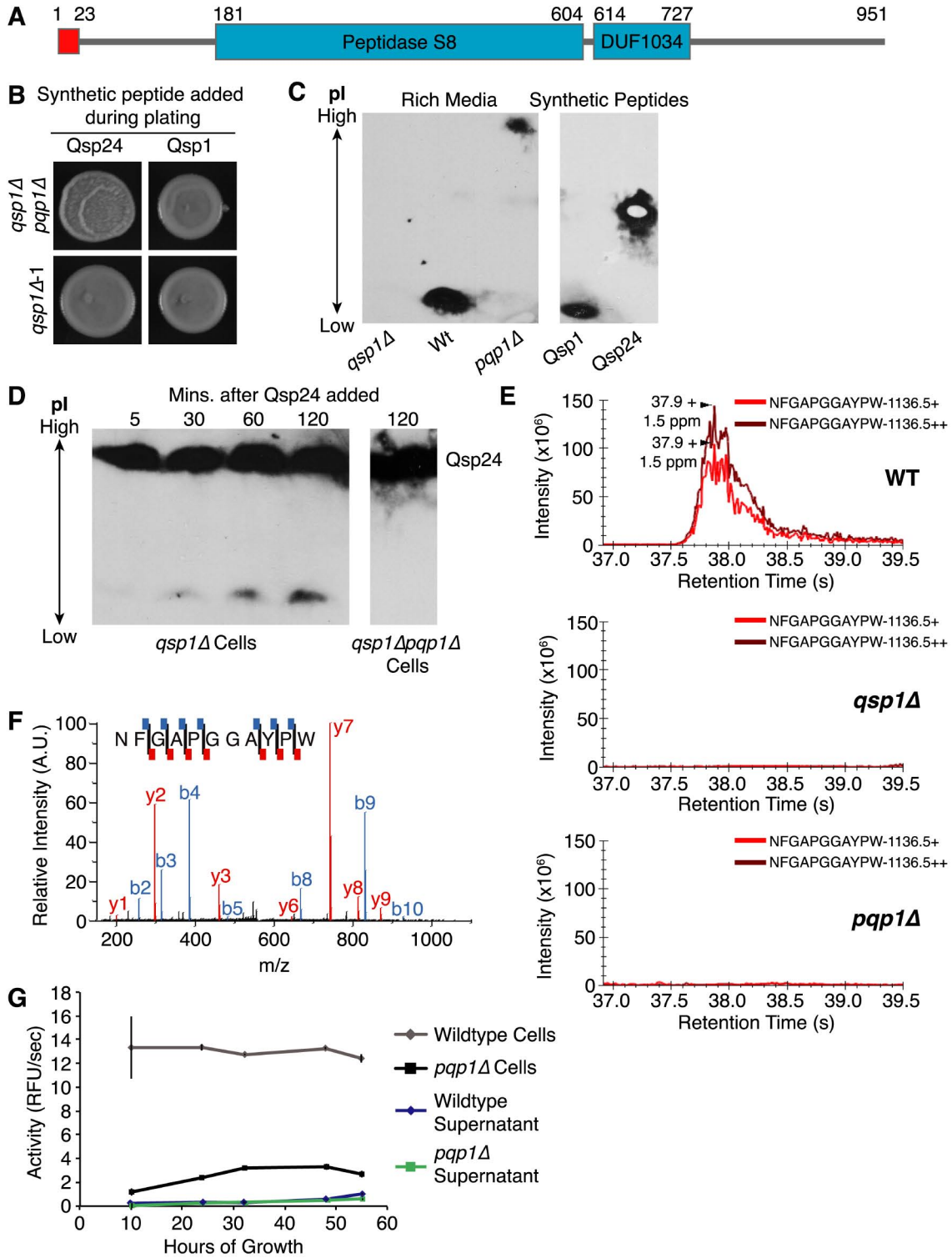■ qsp1Δ-1 + 1μM Qsp1
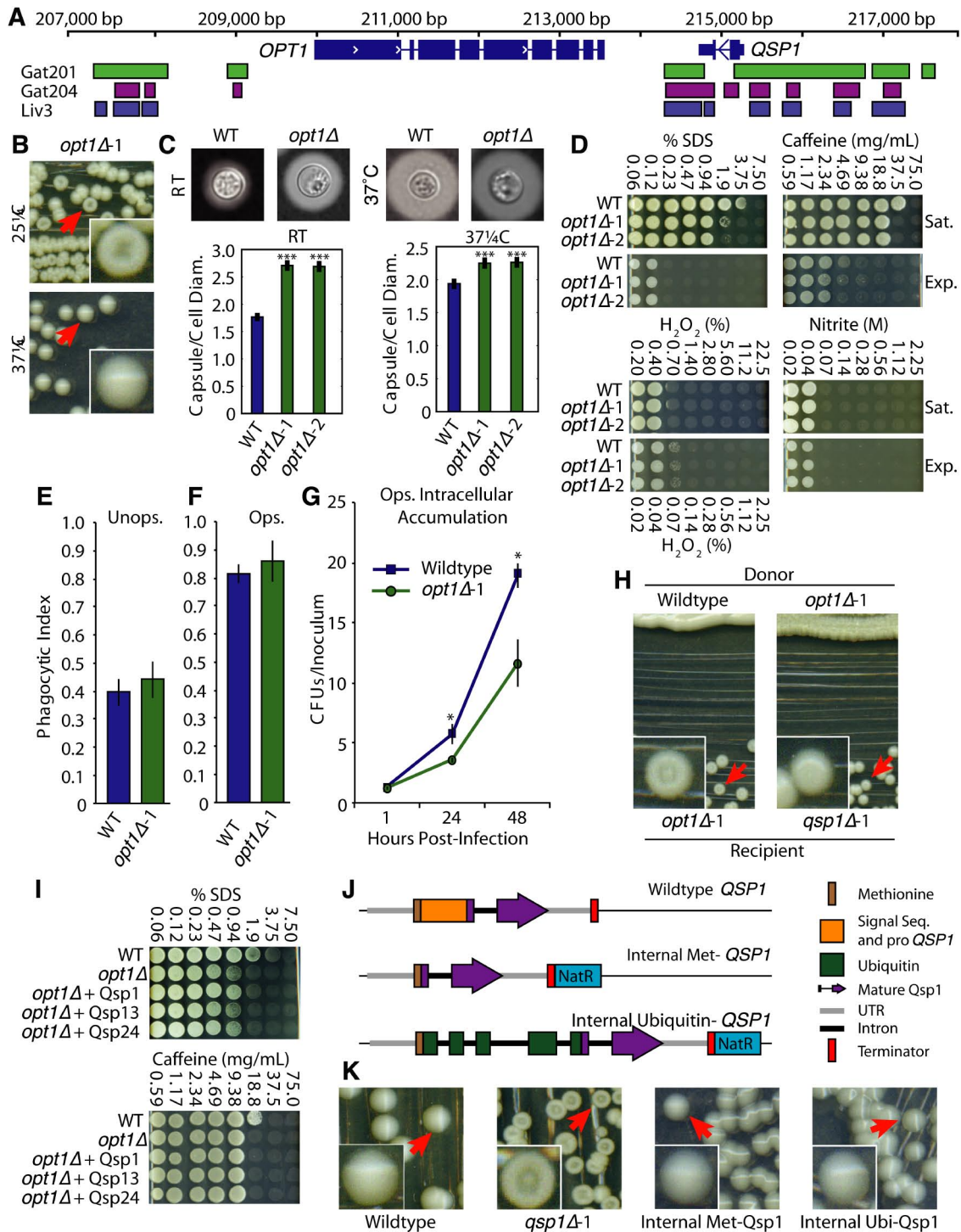
Figure 5.4

166

Figure 5.5

Figure 5.6

Chapter 6. Evolutionary selection on barrier activity: The aspartyl protease Bar1 is an enzyme with novel substrate specificity

6.1 Forward

Mating in the fungal pathogen *Candida albicans* occurs between **a** cells and α cells and is coordinated by the secretion of **a** pheromone and α pheromone by the respective cell type as well as expression of the opposite pheromone receptor. In other words **a** cells secrete **a** pheromone and express the α pheromone receptor whereas α cells secrete α pheromone and express the **a** pheromone receptor. Bar1 is an aspartyl peptidase that is secreted by **a** cells to cleave and inactivate α pheromone. Bar1 activity promotes growth by preventing **a** cells exposed to low levels of α pheromone from entering pheromone induced growth arrest and is also important for preventing mating among **a** cells. Although its target substrate is known, prior to this report the biochemical characteristics of Bar1 had not been thoroughly investigated and it was unknown whether or not this enzyme activity helps to prevent mating between different *Candida* species.

The results presented here indicate that Bar1 has a role in restricting inter-species mating and also a number of unusual characteristics for an aspartyl peptidase. One of the most surprising findings was that Bar1 substrate specificity is determined by residues in the P5'-P8' position of its substrates. In most cases, peptidase substrate specificity is driven by residues in the P4-P4' position, while the residues that form the scissile bond often have the greatest impact on selectivity. The observation that Bar1 cleavage is unaffected by substitution of the residues forming the scissile bond suggests that the specificity of this unique enzyme evolved due to the requirement that its substrate (α pheromone) be recognized by both Bar1 and the α pheromone receptor. Our discovery of the ability of Bar1 to cleave upstream from its recognition site has applications in the biotechnology field, particularly in the removal of affinity tags following protein purification.

My contribution to this report was to interrogate Bar1 substrate specificity by MSP-MS. Consistent with the hypothesis that the majority of substrate specificity of this enzyme is determined by residues in the distal carboxyl region, I found that the consensus cleavage sequence of the P4-P3' residues generated by MSP-MS was not similar to the native substrate, α pheromone. However, the P4' specificity did match the sequence of α pheromone and although not included in the published manuscript, residues in the P5'-P8' positions of MSP-MS substrates were also somewhat similar to these sites within α pheromone, in keeping with our expectations.

6.2 Abstract

Peptide-based pheromones are used throughout the fungal kingdom for coordinating sexual responses between mating partners. Here, we address the properties and function of Bar1, an aspartyl protease that acts as a "barrier" and antagonist to pheromone signaling in multiple species. *Candida albicans* Bar1 was purified and shown to exhibit preferential cleavage of native α pheromone over pheromones from related fungal species. This result establishes that protease substrate specificity co-evolved along with changes in its pheromone target. Pheromone cleavage by Bar1 occurred between residues Thr-5 and Asn-6 in the middle of the tridecapeptide sequence. Surprisingly, however, proteolytic activity was independent of the amino acid residues present at the scissile bond, and instead relied on residues at the C-terminus of α pheromone. Unlike most aspartyl proteases, Bar1 also exhibited a near-neutral pH optimum and was resistant to the class-wide inhibitor, pepstatin A. In addition, a genetic analysis was performed on *C. albicans BAR1,* and demonstrates that the protease not only regulates endogenous pheromone signaling but also can limit inter-species pheromone signaling. We discuss these findings and propose that the unusual substrate specificity of Bar1 is a consequence of its co-evolution with the α pheromone receptor, Ste2, for their shared peptide target.

6.3 Introduction

Pheromone signaling involves the secretion of species-specific chemicals to coordinate cell behavior. Fungi have been intensively studied for their use of sexual pheromones to regulate intercellular signaling and conjugation. In ascomycetes and basidiomycetes, pheromones are peptides or lipopeptides that are secreted into the extracellular milieu and induce morphological and transcriptional responses in target cells. In both of these fungal lineages, mating specificity is determined by the sets of pheromones and pheromone receptors expressed by different cell types (1-6).

In the model ascomycete *Saccharomyces cerevisiae,* **a** cells secrete **a** pheromone and express the α pheromone receptor, Ste2, while α cells secrete α pheromone and express the **a** pheromone receptor, Ste3 (7). Pheromone signaling between the two cell types activates a conserved MAPK signaling cascade, leading to the induction of mating genes and the formation of mating projections or 'shmoos' (8-10). The pheromones produced by *S. cerevisiae* **a** and α cells show distinct physical properties and are secreted by different mechanisms; α pheromone is an unmodified peptide that is secreted by the traditional secretory pathway, whereas **a** factor is farnesylated and carboxymethylated, and transport requires a specific transmembrane translocator (11, 12). The use of both modified and unmodified pheromones is conserved across the ascomycetes, whereas basidiomycetes utilize only lipid-modified pheromones (5, 13).

Candida albicans is a human fungal pathogen related to *S. cerevisiae*, although these species are as divergent as humans and fish (14). Similar to *S. cerevisiae*, mating in *C. albicans* involves pheromone signaling between **a** and α cells (15-18). However, sexual competency in *C. albicans* is dependent on cells undergoing a phenotypic switch from the conventional 'white' state to the alternative 'opaque' state (19, 20). These two states show marked differences in morphology, metabolism, and signaling, including distinct responses to

pheromone (21-23). Pheromone signaling between opaque **a** and α cells leads to the formation of mating projections and conjugation, producing tetraploid **a**/α cells. Although mating incompetent, *C. albicans* white cells can respond to pheromones secreted by opaque cells of the opposite mating type. However, rather than forming mating projections, pheromone-treated white cells adhere to inert surfaces and undergo biofilm formation (24-26).

In addition to the secretion of pheromones, yeast cells also produce degradative enzymes that target mating pheromones for destruction. In both *S. cerevisiae* and *C. albicans*, **a** cells secrete an aspartyl protease, Bar1, that acts as a "barrier" to α pheromone signaling by inactivating it (27-30). Bar1 is produced by **a** cells and sharpens the gradient of α pheromone produced by α cells, thereby facilitating chemotropism and mating between **a** and α partners (31-33). In addition to increasing mating efficiency, Bar1 promotes higher growth rates in subpopulations of **a** cells exposed to α pheromone by overcoming pheromone-induced cell cycle arrest (34). Interestingly, loss of Bar1 results in *C. albicans* transitioning from a heterothallic to a homothallic (selfing) mode of sexual reproduction. *C. albicans* **a** cells secrete low levels of α pheromone (in addition to **a** pheromone), and loss of Bar1 results in positive feedback of α pheromone on these cell types, triggering same-sex **a-a** mating (35). The presence of Bar1 therefore prevents autocrine signaling by α pheromone in *C. albicans* **a** cells, whereas the high level of α pheromone produced by α cells is sufficient to override Bar1 activity during **a**-α mating (30, 35). Interestingly, the distantly related archiascomycete *Schizosaccharomyces pombe* produces a secreted carboxypeptidase, Sxa2, that acts to degrade its own α pheromone-like peptide (36, 37). This indicates that distinct pheromone-degrading proteases have evolved in different ascomycete lineages, and supports the assertion that these proteases play key roles in pheromone signaling.

C. albicans Bar1 contains the Asp-(Thr/Ser)-Gly motif found in other aspartyl proteases and shares significant homology with well-characterized members of this class (38). Aspartyl proteases have been referred to as acidic proteases, given that they are often optimally active

at pH 3-5, and most are susceptible to inhibition by the aspartyl protease inhibitor, pepstatin A (39). However, purified *S. cerevisiae* Bar1 has been shown to have several unusual characteristics for an aspartyl protease, including a pH optimum of 6.5 and resistance to pepstatin A (28, 40). In contrast, the biochemical properties of *C. albicans* Bar1 have not been examined, nor has its specificity for endogenous α pheromone.

In this work, we investigate the biochemical properties of *C. albicans* Bar1 and reveal that the protease is highly selective in cleaving the endogenous α pheromone. Interestingly, substrate cleavage occurs in the middle of the pheromone, but specificity is largely determined by amino acids present at the C-terminus of the sequence. As such, this substrate specificity is highly unusual for an endoprotease, and may have arisen due to constraints on Bar1 activity imposed by the pheromone receptor Ste2, which must also recognize α pheromone. We also examine the role of *BAR1* in intra- and inter-species pheromone signaling in *C. albicans*. These experiments reveal that in addition to modulating endogenous α pheromone activity, *C. albicans* Bar1 also restricts signaling in response to pheromones secreted by related *Candida* species. Together, our studies reveal that *C. albicans* Bar1 is an aspartyl protease with novel substrate specificity that acts to regulate both intra- and inter-species signaling events, and is a potentially invaluable tool with applications in biotechnology.

6.4 Methods

Media

All media was prepared using previously described methods (58, 59). Spider media contained 1% nutrient broth, 0.4% potassium phosphate, and 2% mannitol (pH 7.2). Solid media were made with 1.35% agar.

Strain and plasmid construction

Strains used here are listed in Table 1. To overexpress *C. albicans* Bar1 protein, polymerase chain reaction (PCR) was used to generate a DNA product containing the entire open reading frame (ORF) of *BAR1* from *C. albicans* genomic DNA. The primers used added a 6x His tag to the end of the ORF and restriction endonuclease cut sites. The PCR product was then cloned into the pPinkα-HC plasmid under the *AOX1* promoter (Life Technologies) to yield the plasmid pBar1-His. Site-directed mutagenesis via the QuikChange II kit (Agilent Technologies) was used to mutate position +695 from adenine to cytosine to yield plasmid pBar1(D232A)-His, which contains amino acid substitution D232A in the predicted active site of the Bar1 protein. Both plasmids were transformed into *PichiaPink* strains to generate strains CAY794 and CAY800.

To construct *C. albicans* deletion strains of *MFα* and *BAR1*, plasmids pRB13 and pRB35, respectively, were used with the previously described SAT1-flipper method (60). Briefly, regions flanking the ORF of the gene of interest were PCR amplified together with oligonucleotides containing ApaI/XhoI (5' flank) and SacII/SacI (3' flank) restriction sites. The products were cloned into plasmid pSFS2a. The resulting plasmids were digested with ApaI and SacI and transformed into *C. albicans*. Correct integration was confirmed by PCR and the *SAT1* marker was excised by growth on maltose medium to induce the FLP recombinase (60). Transformations were then repeated to replace the second copy of the gene of interest. To fix cells in the opaque state strains were transformed with plasmid pRB99 (p*ACT1-WOR1*) to constitutively express *WOR1* (41).

C. albicans Bar1 protein expression and purification

P. pastoris strains CAY794 and CAY800 were grown in BMGY media (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 0.00004% biotin, 1% glycerol) for two days at 28°C, then cells were concentrated and switched to BMMY media (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 0.00004% biotin,

0.5% methanol) to induce expression of Bar1 for two days at 28°C. Methanol concentration was

maintained by further addition of methanol after 24 h. Cells were centrifuged and the

supernatant was collected, followed by 40-fold concentration using an Amicon cell (EMD

Millipore, Darmstadt, Germany). The cell was also used to perform buffer exchange into Ni-

NTA binding buffer (300 mM NaCl, 50 mM NaH_2PO_4, 10 mM imidazole, pH 8.0). The resulting

solution was batch-bound to HisPur Ni-NTA resin (Thermo Fisher Scientific, Waltham, MA), then

washed twice (300 mM NaCl, 50 mM NaH_2PO_4, 20 mM imidazole, pH 8.0), and the product

eluted (300 mM NaCl, 50 mM NaH_2PO_4, 500 mM imidazole, pH 8.0). The product was dialyzed

into storage buffer (PBS, 20% glycerol). Protein concentration was determined using a Bradford

protein assay (Thermo Scientific).

Analysis of peptide cleavage by liquid chromatography – mass spectrometry

Lyophilized peptides (Lifetein LLC) were dissolved in 10% DMSO, mixed with 127 nM of

either *C. albicans* Bar1 or Bar1(D232A) protein in 50 µM potassium phosphate solution (pH 6.5)

for one hour at 37°C, then passed through a centrifugal filtration unit to separate peptide

fragments from the peptidase. Samples were analyzed using a Shimadzu HPLC and Thermo

LCQ Deca XP Max ion trap mass spectrometer system.

Multiplex Substrate Profiling by Mass Spectrometry (MSP-MS).

MSP-MS assays were carried out as described previously (45). Briefly, ~60 nM Bar1,

Bar1(D232A) or matched no-enzyme control were assayed against a diverse library of 228

tetradecapeptides pooled at 500 nM in 10 mM KH_2PO_4 (pH 6.5) and incubated at 37°C. After

60, 240 and 1200 min, 30 µL of assay mixture was removed, quenched with 7.5 µL 2% formic

acid, and frozen. Prior to mass spectrometry acquisition, peptide samples were desalted using

Millipore C_{18} ZipTips and rehydrated in 0.2% formic acid. LC-MS/MS data were acquired using

a Thermo Scientific LTQ-Orbitrap mass spectrometer, which was equipped with a Thermo

Scientific EASY-Spray Ion Source, EASY-Spray PepMap C_{18} Column (3 μM, 100 Å), and Waters nanoACQUITY UPLC System. Mass spectrometry peak lists were generated using PAVA, and data were searched against the 228-member peptide library using Protein Prospector software v.5.12.4 (UCSF). Protein Prospector score thresholds were selected with a minimum protein score of 15 and minimum peptide score of 10. Maximum expectation values of 0.01 and 0.05 were used for protein and peptide matches, respectively. Peptides corresponding to cleavage products were imported into iceLogo to generate substrate specificity profiles as described (45). Octapeptides corresponding to P4-P4' were used as the positive dataset, and octapeptides corresponding to all possible cleavages in the MSP-MS library (N = 2,964) were used as the negative data set.

Protein characterization

To determine Bar1 levels in culture, *C. albicans* strains were grown at 22°C overnight in synthetic complete dextrose (SCD) medium and cells washed. Next, 2*10^8 cells were grown for 5 hours in 10 ml SCD medium with or without 0.3 μM synthetic α pheromone. The supernatant and cells were separated by centrifugation, and 40 μl of each was prepared for SDS-PAGE in Laemmli sample buffer. Samples were analyzed by SDS-PAGE and western blotting. For anti-His blots, a 1:5000 dilution of Sigma's HRP-conjugated monoclonal anti-hexahistidine antibody (A7058) was used according to manufacturer's directions. For anti-Bar1 blots, a 1:1000 dilution of sera containing a poly-clonal anti-Bar1 antibody designed by New England Peptides (Project 2095, raised against epitopes Ac-YFINETIRSNDWKC-amide and Ac-CSDDKITSVTSNPQ-amide) was used, followed by a 1:10,000 dilution of GAR-HRP (Jackson Immuno Research). Proteins were visualized using a chemiluminescent substrate (Pierce) and a Chemidoc XRS+ system (Bio-Rad).

For assessing glycosylation of *C. albicans* Bar1, a protein deglycosylation assay was performed on the purified Bar1 protein according to the manufacturer's directions (New England

Biolabs P6039S), and followed by SDS-PAGE and western blotting as described above.

Assessment of the signal peptide sequence of Bar1 was completed using SignalP-4.1 prediction

(http://www.cbs.dtu.dk/services/SignalP/).

Internally quenched (IQ) peptide assays

IQ peptides were produced by LifeTein, LLC (Hillsborough, New Jersey) with the

DABCYL at the N terminus and EDANS at the C terminus. Unless otherwise noted, all

experiments were done at 37°C in 50 µM potassium phosphate buffer, pH 6.5, with a substrate

concentration of 50 µM and 60 nM recombinant Bar1 protease or Bar1(D232A) mutant

protease. The protease and substrate were combined in a 96-well plate immediately before

beginning the experiment, and then fluorescence emission was measured over 90 min using a

BioTek Synergy HT plate reader (Winooski, Vermont). For experiments containing inhibitors,

the inhibitor was added to the protease before combining with the substrate. Each experimental

replicate was an average of three technical replicates. Values were normalized to the

Bar1(D232A) mutant control when noted. Rates were determined over the time period where

substrate was in excess of the enzyme.

Pheromone-induced cell death

C. albicans opaque **a** cells derived from strain AM2003 (41) were grown in SCD medium

overnight at 22°C in a rotating drum. Cultures were washed with water, then resuspended in

Spider media at 2X10^7 cells/ml, and α pheromone was added at the indicated concentration.

After 5 h incubation at 22°C in a rotating drum, 500 µl of culture was removed and washed with

water. Then, the cells were resuspended in SCD media containing 1 µg/ml propidium iodide for

15 min. Cells were washed again, resuspended in SCD media and analyzed by flow cytometry

on a FACSCalibur cell sorter (BD Biosciences). At least 10,000 events were analyzed for

fluorescence in channel FL3. For halo assays, 10^5 opaque cells were plated onto Spider

medium. 20 μg of α pheromone was spotted onto the center of the plate, and the plates incubated at 22°C for three days, after which images were acquired.

Pheromone-induced adherence assay

C. albicans white **a** cells derived from strain P37005 (61) were grown overnight in Spider medium at 22°C on a rotating drum. Cells were washed with water, then administered to each well of a 12-well plate containing 1 ml of Lee's + Glucose media at a concentration of 4*10^7 cells/ml. Pheromones were added at the indicated concentration, then plates incubated at 22°C for 2 days with no shaking. Each well was washed gently 3 times with water to remove non-adherent cells. Adherent cells were scraped and resuspended in 1 ml of water, and the optical density of each sample determined at 600 nm on a Nanodrop 2000c (Thermo Scientific).

Statistical Analysis

Statistics were performed using the PAST software package (folk.uio.no/ohammer/past/). Using the Anderson-Darling method, datasets were tested for normality, and parametric tests were used when appropriate and possible. Tests between samples were performed using the Student's T-test, Tukey's test or Mann-Whitney test, as indicated in figure legends. Asterisks indicate a p-value of 0.05 or less, unless otherwise noted. In order correlate Bar1 cleavage of the tested peptides and activation of Ste2 signaling by the same peptides, the efficiency of Bar1 cleavage was first binned into three groups (products detected by HPLC-MS after 1 h incubation with Bar1, after 24 h incubation, or products not detected even after 24 h co-incubation). Next, activation of Ste2 signaling was compared using data from published pheromone-induced biofilm assays (49) and binned into three groups (similar or better response than with wildtype α pheromone, reduced but detectable response, or no response). The binned data was then correlated using Spearman's rho.

6.5 Results

Purification and characterization of *C. albicans* Bar1 protease

To define the biochemical activity of *C. albicans* Bar1, a recombinant form of the protein was overexpressed and purified from *Pichia pastoris*. The *C. albicans BAR1* gene was engineered with a C-terminal hexahistidine tag and overexpressed using the PichiaPink expression system (see *Materials and Methods*). *BAR1* was expressed under the control of the *P. pastoris AOX1* promoter and contained the *S. cerevisiae* α pheromone signal peptide sequence to ensure efficient secretion into the supernatant. Secreted Bar1 protein was collected and purified by nickel affinity chromatography (Fig. 1A) to a final yield of 6 mg/Liter. The protein migrated as a broad range of high molecular weight bands in western blots, suggestive of protein glycosylation. To test this, *C. albicans* Bar1 was treated with a commercial mixture of glycosylases that increased protein migration by SDS-PAGE (Fig. 1B), indicating that the recombinant protein was highly glycosylated. A mutant Bar1 protein was also purified that contained an amino acid substitution (D232A) at one of the predicted catalytic aspartyl residues (Supp. Fig. 1).

To assess the functionality of recombinant Bar1, we tested its ability to complement a *C. albicans* mutant lacking the *BAR1* gene. A quantitative readout of the *C. albicans* response to pheromone is the pheromone-induced cell death (PID) assay; a significant proportion of opaque **a** cells responding to α pheromone undergo cell death that can be measured by staining with the vitality dye, propidium iodide (41). The pheromone response is dependent on Bar1 activity, as **a** cells lacking *BAR1* undergo PID at lower pheromone concentrations than wildtype Bar1[+] cells (41). Thus, wild-type opaque **a** cells exposed to 10 µM α pheromone showed low levels (<10%) of cell death, whereas *bar1Δ /bar1Δ* mutants showed ~70% cell death (Fig. 1C). Addition of recombinant Bar1 (12.7 nM) to *bar1Δ /bar1Δ* mutants reduced cell death to <10% of the population, whereas addition of the Bar1(D232A) mutant failed to restore wild type levels of PID (Fig. 1C).

A halo assay was also used to verify the functionality of recombinant Bar1. In this assay, α pheromone is spotted onto a lawn of **a** cells, which are then allowed to grow to confluence. Growth is inhibited in *C. albicans* cells that respond to the exogenous pheromone generating 'halos' that indicate the strength of the pheromone response (30). Unlike the weak halos produced by wildtype opaque **a** cells, *bar1Δ/bar1Δ* mutants produced a more distinct halo in response to α pheromone. Addition of recombinant Bar1 reduced the clarity of the halo formed by *bar1Δ/bar1Δ* cells (Fig. 1D), whereas the Bar1(D232A) mutant did not show complementation. Together, these results establish that recombinant Bar1 inhibits its physiological target, and that mutation of one of the catalytic aspartyl residues abolishes activity consistent with its classification as an aspartyl protease.

Biochemical characterization of *C. albicans* Bar1

To determine the kinetics of *C. albicans* Bar1 cleavage, an internally quenched (IQ) version of *C. albicans* α pheromone (GFRLTNFGYEPG) was utilized. Quenching of the C-terminal EDANS fluorophore occurs by resonance energy transfer to the N-terminal dabcyl quencher, and upon cleavage of the peptide the quencher is released so that fluorescence is detected. Under defined conditions, Bar1 processed α pheromone with a K_{cat}/K_m of ~7.7*10^6 M^{-1}s^{-1} (Fig. 2A), whereas the Bar1(D232A) mutant showed no detectable cleavage activity on the IQ substrate. This rate is comparable to that of other secreted *Candida* aspartyl proteases for their respective substrates (42). Bar1 exhibited maximal activity at a pH of 6.5 and at 37°C, with greatest activity in low molarity solutions (Fig. 2B-D). Interestingly, the potent aspartyl protease antagonist pepstatin A did not inhibit Bar1 activity (Fig. 2E). *S. cerevisiae* Bar1 is likewise insensitive to this compound (43). Pepstatin A acts as transition state analog of aspartyl protease substrates (44), and may be a poor mimic of *C. albicans* α pheromone, which contains polar, rather than hydrophobic, residues at its cleavage site (see below and (39)).

Cleavage of α pheromone peptides by *C. albicans* Bar1

To determine the position at which Bar1 cleaves *C. albicans* α pheromone we co-incubated recombinant Bar1 with synthetic α pheromone, and analyzed the digestion products by liquid chromatography and mass-spectroscopy (LC-MS, see *Materials and Methods*). A single cleavage event was detected at the bond between threonine 5 and asparagine 6 (Fig. 3A). In contrast, no cleavage products were generated by the catalytically inactive Bar1(D232A) mutant enzyme (data not shown), confirming that we were monitoring Bar1 activity and not that of a contaminating protease.

C. albicans cells were previously shown to respond to several pheromones from related *Candida* species (41), leading us to test Bar1 activity on each of these peptides using LC-MS (Fig. 3B and C). Bar1 cleaved both *C. dubliniensis* and *C. tropicalis* α pheromones, and the former, which differs by only two residues from *C. albicans* α pheromone, was cleaved with similar efficiency to the native substrate. As with *C. albicans* α pheromone, both *C. dubliniensis* and *C. tropicalis* α pheromones were cleaved between residues 5 and 6 (Thr-Asn and Thr-Arg, respectively).

To provide a more sensitive and quantitative analysis of Bar1 activity, IQ peptides were compared representing *C. albicans*, *C. tropicalis* and *C. parapsilosis* α pheromones. Bar1 cleaved each of the three IQ peptides, although the cleavage rates of *C. tropicalis* and *C. parapsilosis* α pheromones were significantly lower (by ~10-fold) than that of the *C. albicans* α pheromone (Fig. 3C). Together, these results indicate that Bar1 cleaves pheromones besides *C. albicans* α pheromone, although the endogenous pheromone is, by far, the preferred substrate.

Analysis of the cleavage specificity of *C. albicans* Bar1

To further determine the substrate specificity of *C. albicans* Bar1, a series of pheromone-like peptides were incubated with recombinant Bar1 and the products analyzed by

LC-MS. The peptides included scanning di-alanine substitutions across the α pheromone sequence (Fig. 4A, B), as well as N- and C-terminal truncations or extensions of the native sequence (Fig. 4C). Surprisingly, Bar1 cleaved most alanine-substituted peptides, even those with substitutions immediately flanking the scissile bond. However, substitution of several C-terminal residues (residues 10-13) effectively blocked cleavage (Fig. 4B). In each case, proteolysis occurred at the same position as in the native peptide (i.e., between residues Thr 5 and Asn 6). We found that Bar1 did not efficiently degrade peptides in which either the two most N-terminal or C-terminal residues were absent (see peptides N-2 and C-2, Fig. 4C). In comparison, *S. cerevisiae* Bar1 requires four amino acids or more on either side of the scissile bond for activity (40).

To generate an unbiased profile of protease substrate specificity, Bar1 was tested using a Multiplex Substrate Profiling by Mass Spectrometry (MSP-MS) approach (45). For this experiment, a physiochemically diverse library of 228 tetradecapeptides was incubated with Bar1 and the products analyzed at multiple time points by LC-MS/MS (Fig. 4D and 4E). Recognition and cleavage by most proteases is heavily dependent on no more than two amino acids positioned close to the scissile bond (usually within positions P4-P4′), so the library of peptides was designed to contain all combinations of neighbor and near-neighbor amino acid pairs (45). Based on the set of cleavage sites identified in the peptide library, a consensus cleavage motif for Bar1 was generated using iceLogo (46), highlighting substrate specificity on either side of the scissile bond. Notably, the iceLogo motif for Bar1 (Fig. 4E) showed little similarity to the sequence of *C. albicans* α pheromone. The exception was at the P4 position, where a phenylalanine is present in α pheromone and this amino acid was significantly favored in the MSP-MS assay (Fig. 4E). Taken together, these results suggest that Bar1 recognition of α pheromone is largely independent of the amino acids present near the scissile bond. This contrasts with most other endoproteases, where the majority of substrate selectivity originates from residues at, or very close to, the scissile bond (39, 45, 47, 48).

We further explored the specificity of *C. albicans* Bar1 using IQ versions of peptides 5 and 11 (Fig. 4B). These peptides have substituted the amino acid residues immediately flanking the scissile bond (P1 and P1', peptide IQ-5) or within the C-terminal region (P6' and P7', peptide IQ-11) with alanine residues. Cleavage of peptide IQ-5 was found to be more efficient than that of native α pheromone, whereas cleavage of peptide IQ-11 was significantly lower than that of the natural substrate (Fig. 4F). These experiments further establish that amino acid residues at the C-terminus of α pheromone are important for efficient cleavage, whereas those on either side of the scissile bond can both be substituted without negatively affecting Bar1 activity.

Pheromone signaling in *C. albicans* is mediated by α pheromone binding to the Ste2 receptor on the surface of **a** cells. We therefore compared the specificity of *C. albicans* Bar1 with that of Ste2 by examining the ability of different pheromones to activate a cellular response in *C. albicans* **a** cells (see Fig. 3B and (49)). Interestingly, Bar1 substrate specificity mirrored that of Ste2; both *C. albicans* activities were negatively affected by substitution of amino acid residues at the C-terminus of α pheromone, but not by residues in the middle of the peptide (see Fig. 4B and 4C). Furthermore, a significant correlation was found between the pheromone peptides that activate Ste2 and those that are efficiently cleaved by Bar1 (p = 0.00061, Pearson's rho = 0.51, n = 27), suggesting that Bar1 and Ste2 are dependent on the same residues of α pheromone for activity.

Contribution of Bar1 to pheromone signaling in *C. albicans*

C. albicans **a** and α cells can exist in two alternative phenotypic states, white and opaque. Mating involves pheromone signaling between cells in the opaque (mating-competent) state, but white cells also respond to pheromone by becoming adhesive and forming biofilms (Fig. 5A)(24-26). In *C. albicans* **a** cells, signaling often occurs in response to α pheromone

produced by opaque α cells, but it can also be activated by an autocrine loop resulting from α pheromone secreted by **a** cells (35).

To determine the function of *C. albicans* Bar1 in both white and opaque cells, we first examined Bar1 levels secreted by cells in the two phenotypic states. A custom antibody was raised against *C. albicans* Bar1 and western blots performed on conditioned media from cultures of white and opaque **a** cells. Unstimulated cells of both types did not produce Bar1, whereas opaque cells secreted detectable levels of Bar1 when stimulated with α pheromone for 5 hours (Fig. 5B). This data is consistent with RNA profiling experiments, as *BAR1* is highly induced (>200-fold) upon pheromone challenge in opaque cells (50).

Next, to test the contributions of Bar1 and autocrine pheromone signaling to cellular phenotypes, we constructed *bar1Δ/bar1Δ* and *bar1Δ/bar1Δ mfαΔ/mfαΔ* mutants in *C. albicans* **a** cells. Pheromone signaling in opaque cells was detected using the PID assay described previously, whereas white cell responses were quantified by measuring pheromone-induced cell adherence to a plastic substrate (24, 25). Deletion of the *BAR1* gene resulted in heightened sensitivity of opaque cells to exogenous α pheromone, in agreement with previous studies (Schaefer *et al.,* 2007). Thus, the majority of *bar1Δ/bar1Δ* cells underwent cell death at pheromone concentrations as low as 30 nM, while wildtype cells required > 1 µM of pheromone to elicit a similar level of cell death (Fig. 6B). Deletion of the *MFα* gene prevents autocrine signaling, and yet *bar1Δ/bar1Δ mfαΔ/mfαΔ* cells showed similar levels of PID to *bar1Δ/bar1Δ* mutants (Fig. 5C). This result indicates that endogenous α pheromone production and autocrine signaling has a negligible effect on the pheromone response under these conditions. Similar results were obtained with white cells; loss of *BAR1* increased the sensitivity of white cells to exogenous α pheromone, and deletion of the *MFα* gene, either singly or in combination with *BAR1*, did not show a significant effect on pheromone-induced adherence (Fig. 6C).

Bar1 regulation of inter-species signaling in *C. albicans*

C. albicans **a** cells have been shown to respond to α pheromones from related *Candida* species, with opaque cells induced to undergo same-sex mating whereas white cells exhibit increased biofilm formation (49). These results led us to ask two questions: First, can Bar1 regulate inter-species signaling by degradation of α pheromones secreted by other species? Second, does secretion of endogenous α pheromone from *C. albicans* **a** cells enhance inter-species signaling by activating autocrine signaling? There are three scenarios for how pheromone signaling could be influenced by *BAR1* and *MFα* in *C. albicans* **a** cells: (1) Signaling between species is unaffected by Bar1 or endogenous α pheromone produced by **a** cells. (2) Bar1 degrades non-*albicans* pheromones thereby directly restricting inter-species pheromone signaling. (3) Bar1 does not influence inter-species signaling directly, but restricts it by degrading endogenous α pheromone produced by **a** cells, thereby reducing autocrine signaling (Fig. 6A).

Deletion of the *BAR1* gene in *C. albicans* white **a** cells led to a heightened response to *C. dubliniensis, C. tropicalis,* and *C. parapsilosis* α pheromones (Fig. 6B-6C). This is consistent with the biochemical activity of Bar1, which was capable of degrading each of these three pheromones *in vitro* (Fig. 3C). Sensitivity to *C. dubliniensis* pheromone was considerably higher than that to *C. tropicalis* and *C. parapsilosis* pheromones, consistent with the close sequence homology of *C. albicans* and *C. dubliniensis* pheromones (Fig. 3B). In contrast to *BAR1*, deletion of the *MFα* gene did not affect the response to the three non-*albicans* pheromones (Fig. 6B-6C). This indicates that Bar1 directly affects the response to these pheromones similar to the result with *C. albicans* α pheromone (compare Fig. 5C and Fig. 6B-6C).

Analogous experiments were performed using *C. albicans* opaque **a** cells to determine the contributions of *BAR1* and *MFα* to inter-species signaling (Fig. 6D-6E). *C. dubliniensis* α pheromone produced a similar response to the native *C. albicans* α pheromone in opaque cells and deletion of *MFα* did not affect the outcome (Fig. 6D). In contrast, deletion of both *MFα* and *BAR1* genes depressed the response of *C. albicans* opaque **a** cells to *C. parapsilosis* α

pheromone, indicating that autocrine signaling significantly affects the response to this pheromone (Fig. 6E). Taken together, these results recapitulate our biochemical findings that Bar1 acts most effectively on *C. albicans* α pheromone, but also has the capacity to degrade pheromones from other *Candida* species thereby limiting inter-species signaling.

6.6 Discussion

Mating in fungi is choreographed by pheromone signaling between cell partners. In this work, we addressed the biochemical properties and function of *C. albicans* Bar1, an aspartyl protease secreted by **a** cells that degrades α pheromone. We show that Bar1 cleaves endogenous α pheromone between threonine-5 and asparagine-6, thereby inactivating the 13 amino acid peptide. This protease exhibits several features common to secreted aspartyl proteases, including extensive glycosylation and an Asp-(Thr/Ser)-Gly motif. However, Bar1 also displays novel properties for an aspartyl protease, including a near-neutral pH optimum and resistance to pepstatin A, a potent aspartyl protease antagonist that inhibits members of the *C. albicans* SAP protease family (Naglik et al., 2003). Unlike Bar1, most aspartyl proteases have optimal activity at acidic pH (28, 39, 51), although *S. cerevisiae* Bar1 is a notable exception, as this enzyme also shows a near-neutral pH optimum and is not inhibited by pepstatin A (28, 29, 38, 40). It therefore appears that *Sc*Bar1 and *Ca*Bar1 share several unusual properties, despite the extensive evolutionary divergence between *S. cerevisiae* and *C. albicans* (14).

A comparative analysis of *C. albicans* Bar1 cleavage of α pheromones from multiple *Candida* species revealed that Bar1 preferentially cleaved *C. albicans* α pheromone, whereas α pheromones from closely related species (including *C. tropicalis* and *C. parapsilosis*) were also cleaved, albeit at a much-reduced rate. Bar1 did not cleave pheromones from more evolutionarily distant fungal species. Thus, *C. albicans* Bar1 has evolved high specificity towards the endogenous α pheromone. Studies on *Sc*Bar1 also suggest a species-specific

function, as *S. cerevisiae* cells were unable to cleave *C. albicans* α pheromone (52). Together, these experiments establish that Bar1 has evolved selectivity towards its native pheromone in diverse fungal lineages.

An unexpected feature of *C. albicans* Bar1 cleavage activity was its dependence on amino acid residues located distal to the scissile bond. The specificity of most endoproteases is guided by amino acids located at, or very close to, the scissile bond (45, 48). In contrast, Bar1 cleaved α pheromone by recognition of amino acid residues at the C-terminus of the peptide. Thus, alanine substitutions at positions 10-13 of α pheromone (P4′-P7′) inhibited cleavage, whereas substitutions at the scissile bond did not (Fig. 4B and 4F). The unusual specificity of Bar1 was further supported by an unbiased analysis of activity using MSP-MS, which identifies amino acids close to the scissile bond that are enriched or de-enriched in substrates (45). The amino acids favored on either side of the scissile bond by MSP-MS analysis had little similarity to those present in *C. albicans* α pheromone. Together, these studies demonstrate that the cleavage specificity of Bar1 for α pheromone is mediated by residues distal to the scissile bond rather than amino acids close to the site of cleavage.

Why might *C. albicans* Bar1 have evolved specificity for residues away from the scissile bond? We show that the specificity of Bar1 exhibits a striking similarity to that of the *C. albicans* Ste2 receptor for α pheromone. A previous study revealed that substitution of residues at the C-terminus of α pheromone (particularly residues 11 and 12) resulted in the largest reduction in pheromone signaling by *C. albicans* **a** cells (see "Ste2" column, Fig. 4B and C) (49). Thus, both *C. albicans* Bar1 and Ste2 are critically dependent on the same C-terminal residues of α pheromone for activity. We therefore propose that there has been co-evolution of these factors due to their dependence for the same peptide target. In particular, as receptor-pheromone interactions diverged between species, Bar1 specificity for α pheromone would have had to evolve in parallel to retain activity. By having Bar1 and Ste2 recognize the same residues in α pheromone this would have facilitated co-evolution of these factors together with their peptide

target during speciation (see Fig. 7). Studies have also addressed the interaction between *S. cerevisiae* Ste2 and α pheromone, and shown that *Sc*Ste2 makes multiple contacts with α pheromone, although residues near the C-terminus of the pheromone (residues 10-13) play the most important role in receptor binding (53-56). Thus, in both *C. albicans* and *S. cerevisiae,* the carboxyl terminal residues of α pheromone are critical for receptor function. It will now be of interest to examine pheromone-receptor and pheromone-Bar1 interactions across multiple species to further address how these factors co-evolved, and if the C-terminus of α pheromone has a conserved role in mediating interactions with both of its protein targets across diverse species.

To complement the biochemical analysis of *C. albicans* Bar1, we performed genetic experiments to define the role of this protease in regulating intra- and inter-species pheromone signaling. Loss of *BAR1* sensitized *C. albicans* white and opaque **a** cells to native α pheromone, indicating functional secretion of the Bar1 enzyme by both cell types. Deletion of the *BAR1* gene also sensitized *C. albicans* **a** cells to α pheromones from *C. dubliniensis*, *C. tropicalis* and *C. parapsilosis,* indicating that Bar1 effectively inhibits signaling to pheromones produced by closely related species. This could be important in nature, as *C. albicans* is often found co-colonizing the oral cavity with multiple other *Candida* clade species (57).

We also addressed whether autocrine pheromone production contributes to inter-species signaling, as *C. albicans* **a** cells have been shown to secrete α pheromone in addition to the conventional **a** pheromone (35). To test this, the *MF*α gene (encoding α pheromone) was deleted in the presence or absence of *BAR1*, and **a** cells challenged with α pheromones. In most cases, loss of *MF*α did not affect the response of *C. albicans* **a** cells to other species' pheromones. However, the response to *C. parapsilosis* α pheromone was reduced upon deletion of both *BAR1* and *MF*α genes, indicating that autocrine signaling enhanced inter-species signaling to this pheromone (Fig. 6E). These results are the first to demonstrate that

Bar1 acts to restrict inter-species pheromone signaling, and also show a role for endogenously produced pheromone in enhancing the response to pheromones from related species.

Finally, we note that the unusual specificity of *C. albicans* Bar1 makes it an attractive candidate for bioengineering a protease with specialized properties. There are a number of potential technical applications for a protease whose specificity is dependent on residues distal to the scissile bond, including the ability to completely remove affinity tags from purified proteins without leaving behind the cleavage site residues. In principle, this is analogous to restriction enzymes (e.g., type III restriction enzymes) that cleave the DNA at a short distance from their recognition sequence. In fact, enteropeptidase is a serine protease whose specificity determinants are located on the C-terminal side of the scissile bond, so that N-terminal tags can be removed leaving the product with a native N-terminus (48). *C. albicans* Bar1 could be engineered to perform a similar function for the complete removal of C-terminal tags from recombinant proteins. Bar1 can be highly overexpressed, shows strong substrate selectivity, and is active at neutral pH, all properties that would favor its use in many applications. Further analysis of Bar1 homologs from *Saccharomycotina* species could also help identify the residues in Bar1 that are responsible for determining its substrate specificity, as Bar1 orthologs appear to be under strong selection to evolve in parallel with their pheromone substrates. Such experiments could shed light on the precise mechanism by which α pheromone is recognized, and also allow further refining of the proteolytic activity for important commercial applications.

6.7 References

1.   **Bender A, Sprague GF, Jr.** 1986. Yeast peptide pheromones, a-factor and alpha-factor, activate a common response mechanism in their target cells. Cell **47:**929-937.

2.   **Stanton BC, Giles SS, Staudt MW, Kruzel EK, Hull CM.** 2010. Allelic exchange of pheromones and their receptors reprograms sexual identity in *Cryptococcus neoformans*. PLoS Genet **6:**e1000860.

3.   **Bender A, Sprague GF, Jr.** 1989. Pheromones and pheromone receptors are the primary determinants of mating specificity in the yeast *Saccharomyces cerevisiae*. Genetics **121:**463-476.

4.   **Bolker M, Urban M, Kahmann R.** 1992. The a mating type locus of *U. maydis* specifies cell signaling components. Cell **68:**441-450.

5.   **Goncalves-Sa J, Murray A.** 2011. Asymmetry in sexual pheromones is not required for ascomycete mating. Curr Biol **21:**1337-1346.

6.   **Seike T, Nakamura T, Shimoda C.** 2015. Molecular coevolution of a sex pheromone and its receptor triggers reproductive isolation in *Schizosaccharomyces pombe*. Proc Natl Acad Sci U S A **112:**4405-4410.

7.   **Johnson AD.** 1995. Molecular mechanisms of cell-type determination in budding yeast. Curr Opin Genet Dev **5:**552-558.

8.   **Jones SK, Jr., Bennett RJ.** 2011. Fungal mating pheromones: choreographing the dating game. Fungal Genet Biol **48:**668-676.

9.   **Merlini L, Dudin O, Martin SG.** 2013. Mate and fuse: how yeast cells do it. Open biology **3:**130008.

10.   **Bardwell L.** 2005. A walk-through of the yeast mating pheromone response pathway. Peptides **26:**339-350.

11.   **Chen P, Sapperstein SK, Choi JD, Michaelis S.** 1997. Biogenesis of the *Saccharomyces cerevisiae* mating pheromone a-factor. J Cell Biol **136:**251-269.

12.   **Kuchler K, Sterne RE, Thorner J.** 1989. *Saccharomyces cerevisiae STE6* gene product: a novel pathway for protein export in eukaryotic cells. EMBO J **8:**3973-3984.

13.   **Casselton LA, Olesnicky NS.** 1998. Molecular genetics of mating recognition in basidiomycete fungi. Microbiol Mol Biol Rev **62:**55-70.

14.   **Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, De Montigny J, Marck C, Neuveglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich JM, Beyne E, Bleykasten C, Boisrame A, Boyer J, Cattolico L, Confanioleri F, De Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud JM, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard GF, Straub ML, Suleau A, Swennen D, Tekaia F, Wesolowski-Louvel M, Westhof E, Wirth B, Zeniou-Meyer M, Zivanovic I, Bolotin-Fukuhara M, Thierry A, Bouchier C, Caudron B, Scarpelli C, Gaillardin C, Weissenbach J, Wincker P, Souciet JL.** 2004. Genome evolution in yeasts. Nature **430:**35-44.

15.   **Bennett RJ, Miller MG, Chua PR, Maxon ME, Johnson AD.** 2005. Nuclear fusion occurs during mating in *Candida albicans* and is dependent on the *KAR3* gene. Mol Microbiol **55:**1046-1059.

16.   **Bennett RJ, Uhl MA, Miller MG, Johnson AD.** 2003. Identification and characterization of a *Candida albicans* mating pheromone. Mol Cell Biol **23:**8189-8201.

17. **Lockhart SR, Zhao R, Daniels KJ, Soll DR.** 2003. Alpha-pheromone-induced "shmooing" and gene regulation require white-opaque switching during *Candida albicans* mating. Eukaryot Cell **2:**847-855.

18. **Panwar SL, Legrand M, Dignard D, Whiteway M, Magee PT.** 2003. *MFalpha1*, the gene encoding the alpha mating pheromone of *Candida albicans*. Eukaryot Cell **2:**1350-1360.

19. **Miller MG, Johnson AD.** 2002. White-opaque switching in *Candida albicans* is controlled by mating-type locus homeodomain proteins and allows efficient mating. Cell **110:**293-302.

20. **Slutsky B, Staebell M, Anderson J, Risen L, Pfaller M, Soll DR.** 1987. "White-opaque transition": a second high-frequency switching system in *Candida albicans*. J Bacteriol **169:**189-197.

21. **Geiger J, Wessels D, Lockhart SR, Soll DR.** 2004. Release of a potent polymorphonuclear leukocyte chemoattractant is regulated by white-opaque switching in *Candida albicans*. Infect Immun **72:**667-677.

22. **Si H, Hernday AD, Hirakawa MP, Johnson AD, Bennett RJ.** 2013. *Candida albicans* white and opaque cells undergo distinct programs of filamentous growth. PLoS Pathog **9:**e1003210.

23. **Sudbery PE.** 2011. Growth of *Candida albicans* hyphae. Nat Rev Microbiol **9:**737-748.

24. **Daniels KJ, Srikantha T, Lockhart SR, Pujol C, Soll DR.** 2006. Opaque cells signal white cells to form biofilms in *Candida albicans*. Embo J **25:**2240-2252.

25. **Lin CH, Kabrawala S, Fox EP, Nobile CJ, Johnson AD, Bennett RJ.** 2013. Genetic control of conventional and pheromone-stimulated biofilm formation in *Candida albicans*. PLoS Pathog **9:**e1003305.

26. **Yi S, Sahni N, Daniels KJ, Pujol C, Srikantha T, Soll DR.** 2008. The same receptor, G protein, and mitogen-activated protein kinase pathway activate different downstream regulators in the alternative white and opaque pheromone responses of *Candida albicans.* Mol Biol Cell **19:**957-970.

27. **Sprague GF, Jr., Herskowitz I.** 1981. Control of yeast cell type by the mating type locus. I. Identification and control of expression of the a-specific gene *BAR1*. J Mol Biol **153:**305-321.

28. **Manney TR.** 1983. Expression of the *BAR1* gene in *Saccharomyces cerevisiae*: induction by the alpha mating pheromone of an activity associated with a secreted protein. J Bacteriol **155:**291-301.

29. **MacKay VL, Welch SK, Insley MY, Manney TR, Holly J, Saari GC, Parker ML.** 1988. The *Saccharomyces cerevisiae BAR1* gene encodes an exported protein with homology to pepsin. Proc Natl Acad Sci U S A **85:**55-59.

30. **Schaefer D, Cote P, Whiteway M, Bennett RJ.** 2007. Barrier activity in *Candida albicans* mediates pheromone degradation and promotes mating. Eukaryot Cell **6:**907-918.

31. **Andrews SS, Addy NJ, Brent R, Arkin AP.** 2010. Detailed simulations of cell biology with Smoldyn 2.1. PLoS computational biology **6:**e1000705.

32. **Barkai N, Rose MD, Wingreen NS.** 1998. Protease helps yeast find mating partners. Nature **396:**422-423.

33. **Jin M, Errede B, Behar M, Mather W, Nayak S, Hasty J, Dohlman HG, Elston TC.** 2011. Yeast dynamically modify their environment to achieve better mating efficiency. Science signaling **4:**ra54.

34. **Diener C, Schreiber G, Giese W, del Rio G, Schroder A, Klipp E.** 2014. Yeast mating and image-based quantification of spatial pattern formation. PLoS computational biology **10:**e1003690.

35. **Alby K, Schaefer D, Bennett RJ.** 2009. Homothallic and heterothallic mating in the opportunistic pathogen *Candida albicans*. Nature **460:**890-893.

36. **Ladds G, Davey J.** 1996. Characterisation of Sxa2, a carboxypeptidase involved in pheromone recovery in fission yeast. Biochemical Society transactions **24:**210S.

37. **Ladds G, Rasmussen EM, Young T, Nielsen O, Davey J.** 1996. The sxa2-dependent inactivation of the P-factor mating pheromone in the fission yeast *Schizosaccharomyces pombe*. Mol Microbiol **20:**35-42.

38. **Naglik JR, Challacombe SJ, Hube B.** 2003. *Candida albicans* secreted aspartyl proteinases in virulence and pathogenesis. Microbiol Mol Biol Rev **67:**400-428.

39. **Dunn BM.** 2001. Overview of pepsin-like aspartic peptidases. Current protocols in protein science / editorial board, John E. Coligan. **Chapter 21:**Unit 21.23.

40. **MacKay VL, Armstrong J, Yip C, Welch S, Walker K, Osborn S, Sheppard P, Forstrum J.** 1991. Characterization of the Bar proteinase, an extracellular enzyme from the yeast *Saccharomyces cerevisiae*., p. 161-172. *In* B.M.Dunn (ed.), Structure and Function of the Aspartic Proteinases. Plenum Press, New York.

41. **Alby K, Schaefer D, Sherwood RK, Jones SK, Jr., Bennett RJ.** 2010. Identification of a cell death pathway in *Candida albicans* during the response to pheromone. Eukaryot Cell **9:**1690-1701.

42. **Fusek M, Smith EA, Monod M, Dunn BM, Foundling SI.** 1994. Extracellular aspartic proteinases from *Candida albicans, Candida tropicalis,* and *Candida parapsilosis* yeasts differ substantially in their specificities. Biochemistry **33:**9791-9799.

43. **Nath R.** 1993. Properties of Barrier, a novel *Saccharomyces cerevisiae* acid protease. Biochimie **75:**467-472.

44. **Marciniszyn J, Jr., Hartsuck JA, Tang J.** 1976. Mode of inhibition of acid proteases by pepstatin. J Biol Chem **251:**7088-7094.

45. **O'Donoghue AJ, Eroy-Reveles AA, Knudsen GM, Ingram J, Zhou M, Statnekov JB, Greninger AL, Hostetter DR, Qu G, Maltby DA, Anderson MO, Derisi JL, McKerrow JH, Burlingame AL, Craik CS.** 2012. Global identification of peptidase specificity by multiplex substrate profiling. Nature methods **9:**1095-1100.

46. **Colaert N, Helsens K, Martens L, Vandekerckhove J, Gevaert K.** 2009. Improved visualization of protein consensus sequences by iceLogo. Nature methods **6:**786-787.

47. **Dunn BM, Hung S.** 2000. The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases. Biochim Biophys Acta **1477:**231-240.

48. **Waugh DS.** 2011. An overview of enzymatic reagents for the removal of affinity tags. Protein expression and purification **80:**283-293.

49. **Alby K, Bennett RJ.** 2011. Interspecies pheromone signaling promotes biofilm formation and same-sex mating in *Candida albicans*. Proc Natl Acad Sci U S A **108:**2510-2515.

50.   **Bennett RJ, Johnson AD.** 2006. The role of nutrient regulation and the Gpa2 protein in the mating pheromone response of *C. albicans*. Mol Microbiol **62:**100-119.

51.   **Ghosh AK.** 2011. Aspartic Acid Proteases as Therapeutic Targets. John Wiley and Sons.

52.   **Janiak AM, Sargsyan H, Russo J, Naider F, Hauser M, Becker JM.** 2005. Functional expression of the *Candida albicans* alpha-factor receptor in *Saccharomyces cerevisiae.* Fungal Genet Biol **42:**328-338.

53.   **Abel MG, Zhang YL, Lu HF, Naider F, Becker JM.** 1998. Structure-function analysis of the *Saccharomyces cerevisiae* tridecapeptide pheromone using alanine-scanned analogs. The journal of peptide research : official journal of the American Peptide Society **52:**95-106.

54.   **Lee BK, Khare S, Naider F, Becker JM.** 2001. Identification of residues of the *Saccharomyces cerevisiae* G protein-coupled receptor contributing to alpha-factor pheromone binding. J Biol Chem **276:**37950-37961.

55.   **Naider F, Becker JM.** 2004. The alpha-factor mating pheromone of *Saccharomyces cerevisiae*: a model for studying the interaction of peptide hormones and G protein-coupled receptors. Peptides **25:**1441-1463.

56.   **Son CD, Sargsyan H, Naider F, Becker JM.** 2004. Identification of ligand binding regions of the *Saccharomyces cerevisiae* alpha-factor pheromone receptor by photoaffinity cross-linking. Biochemistry **43:**13193-13203.

57.   **Ghannoum MA, Jurevic RJ, Mukherjee PK, Cui F, Sikaroodi M, Naqvi A, Gillevet PM.** 2010. Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. PLoS Pathog **6:**e1000713.

58.   **Guthrie C, Fink GR.** 1991. Guide to Yeast Genetics and Molecular Biology. Academic Press, San Diego.

59.   **Liu H, Kohler J, Fink GR.** 1994. Suppression of hyphal formation in *Candida albicans* by mutation of a *STE12* homolog. Science **266:**1723-1726.

60.   **Reuss O, Vik A, Kolter R, Morschhauser J.** 2004. The *SAT1* flipper, an optimized tool for gene disruption in *Candida albicans*. Gene **341:**119-127.

61.   **Lockhart SR, Pujol C, Daniels KJ, Miller MG, Johnson AD, Pfaller MA, Soll DR.** 2002. In *Candida albicans*, white-opaque switchers are homozygous for mating type. Genetics **162:**737-745.
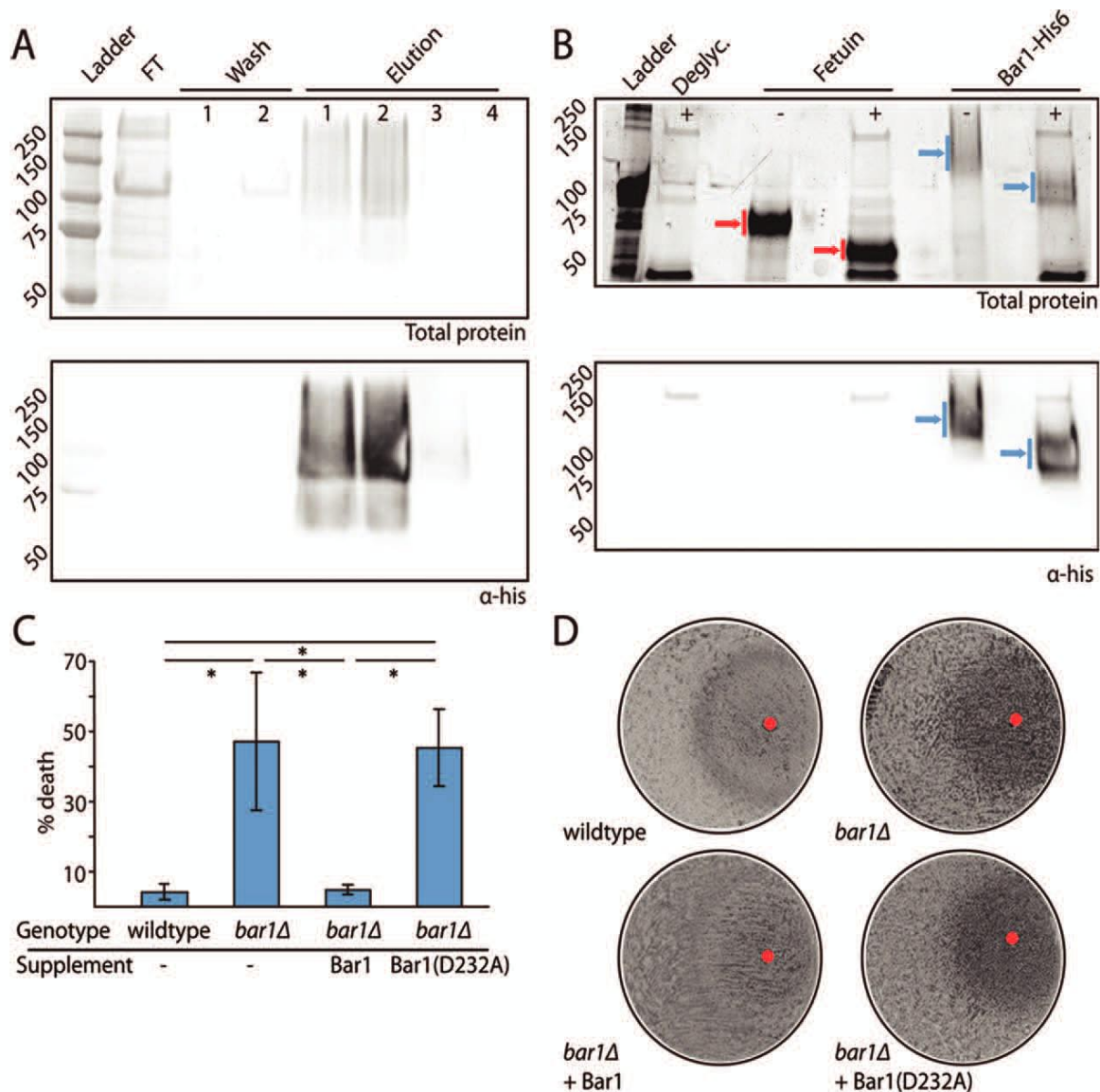
Figure 6.1 Purification and analysis of a recombinant CaBar1 protease.
(A) Purification of his-tagged C. albicans Bar1 (CaBar1) from Pichia pastoris using nickel affinity chromatography. Nickel beads were washed twice (20 mM imidazole) and protein eluted with 500 mM imidazole. FT (flow through). Total protein was assayed by coomassie staining (top panel), and CaBar1 detected by western blotting with an anti-his antibody (lower panel). (B) Protein glycosylation was assessed by treatment of CaBar1 or fetuin (positive control) with a cocktail of deglycosylation enzymes. Blue arrows = CaBar1, Red arrows = fetuin. Presence or absence of deglycosylase enzymes is noted with +/-.   Total protein is shown (top panel) using Biorad Stainfree indicator, and CaBar1 was detected by western blotting with an anti-his antibody (lower panel). (C) To evaluate the activity of recombinant CaBar1, wildtype and bar1Δ/bar1Δ opaque a strains were treated with or without 300 nM α pheromone for 5 hours, together with recombinant CaBar1 or CaBar1(D232A) at 127 nM. Percent death was analyzed by staining with propidium iodide and flow cytometry. Mean ± SD. N=3. * = p<0.05 by T-test. (D) Cell cycle arrest was analyzed by plating a lawn of opaque a cells on solid medium, and spotting 20 μg α pheromone (red circle) together with CaBar1 or CaBar1(D232A), then imaging of plates after 2 days culture at 22°C.
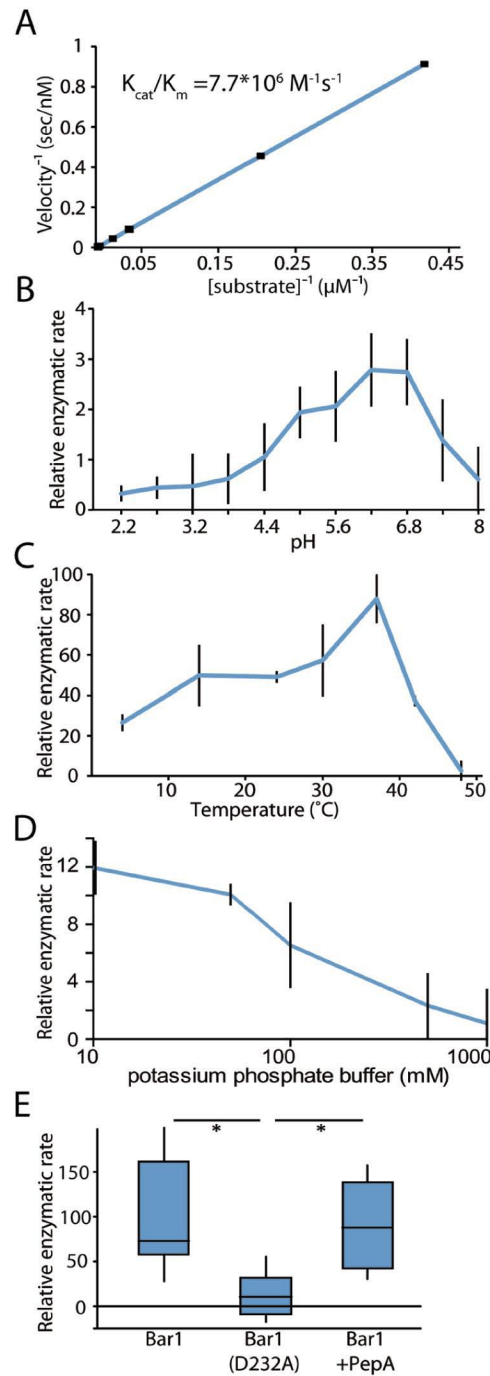
Figure 6.2. Characterization of CaBar1 protease activity. CaBar1 activity was determined using an internally quenched peptide corresponding to the native α pheromone sequence of C. albicans, GFRLTNFGYFEPG . (A) Activity was plotted using a Lineweaver-Burk plot to determine enzyme kinetics. (B) pH dependence, (C) temperature dependence, and (D) osmolarity dependence of CaBar1 activity. Relative enzymatic rate is the amount of fluorescence per unit time. Default conditions were 37°C at pH 6.5. Data shows mean ± SD. (E) Enzymatic activities of C. albicans α pheromone with CaBar1, CaBar1(D232A), or CaBar1 co-incubated with pepstatin A were determined by co-incubation with the internally quenched peptide substrate for 90 min at 37°C, pH 6.5. Tukey boxplot. N≥5. * = p<0.05 by Mann-Whitney U test.
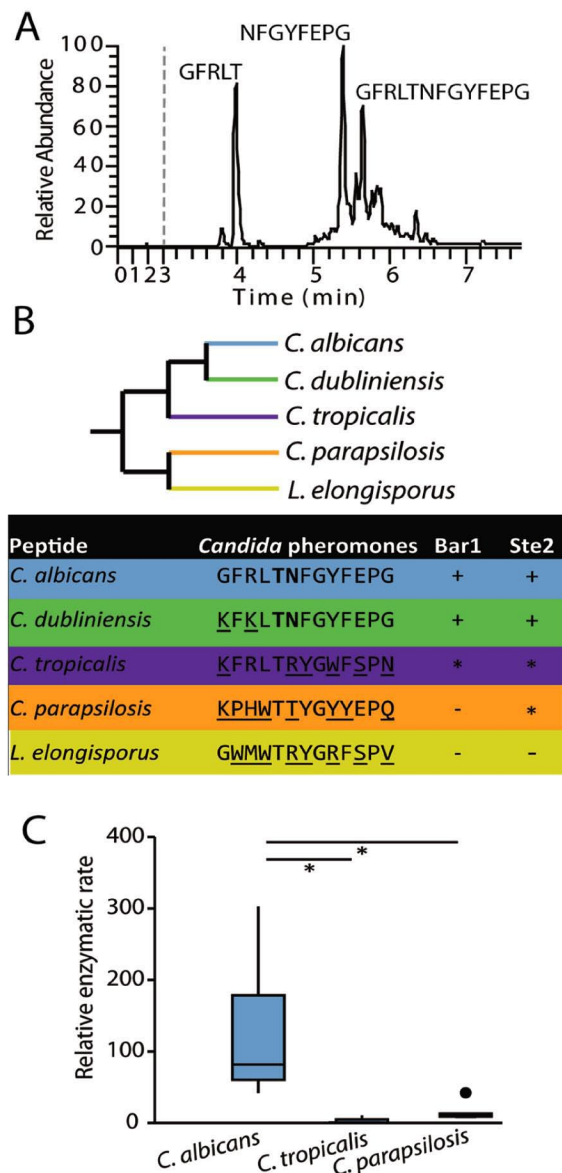
Figure 6.3. Analysis of CaBar1 protease activity on ? pheromones from multiple Candida species. (A) Cleavage of C. albicans α pheromone by CaBar1 protease. CaBar1 and ? pheromone were co-incubated for 1 h and the products analyzed by LC-MS. Mass spectrometry identified two major products (GFRLT and NFGYFEPG), as well as uncleaved pheromone (GFRLTNFGYFEPG). (B) An unrooted, phylogenetic tree of multiple Candida clade species is shown, as well as the activity of CaBar1 on α pheromone peptides from each species. The detection of specific degradation products is indicated by "+" in the Bar1 column; "-" indicates no products were detected while "*" indicates products were formed only after an extended (1 day) incubation. Bolded residues indicate amino acids adjacent to the cleavage site. The Ste2 column indicates if the pheromone induced a robust response (+), a weak response (*), or no response (-) in C. albicans MTLa cells. Ste2 data adapted from Alby and Bennett, 2011. (C) CaBar1 activity was compared on C. albicans, C. tropicalis, and C. parapsilosis ? pheromones using internally quenched peptide substrates. Purified CaBar1 was incubated with ? pheromones for 90 min at 37°C. Relative enzymatic rate is the amount of fluorescence per unit time due to cleavage of the fluorophore-conjugated peptide. Tukey boxplot with outliers noted as ●. N≥5. * =p<0.05 by Mann-Whitney U test.
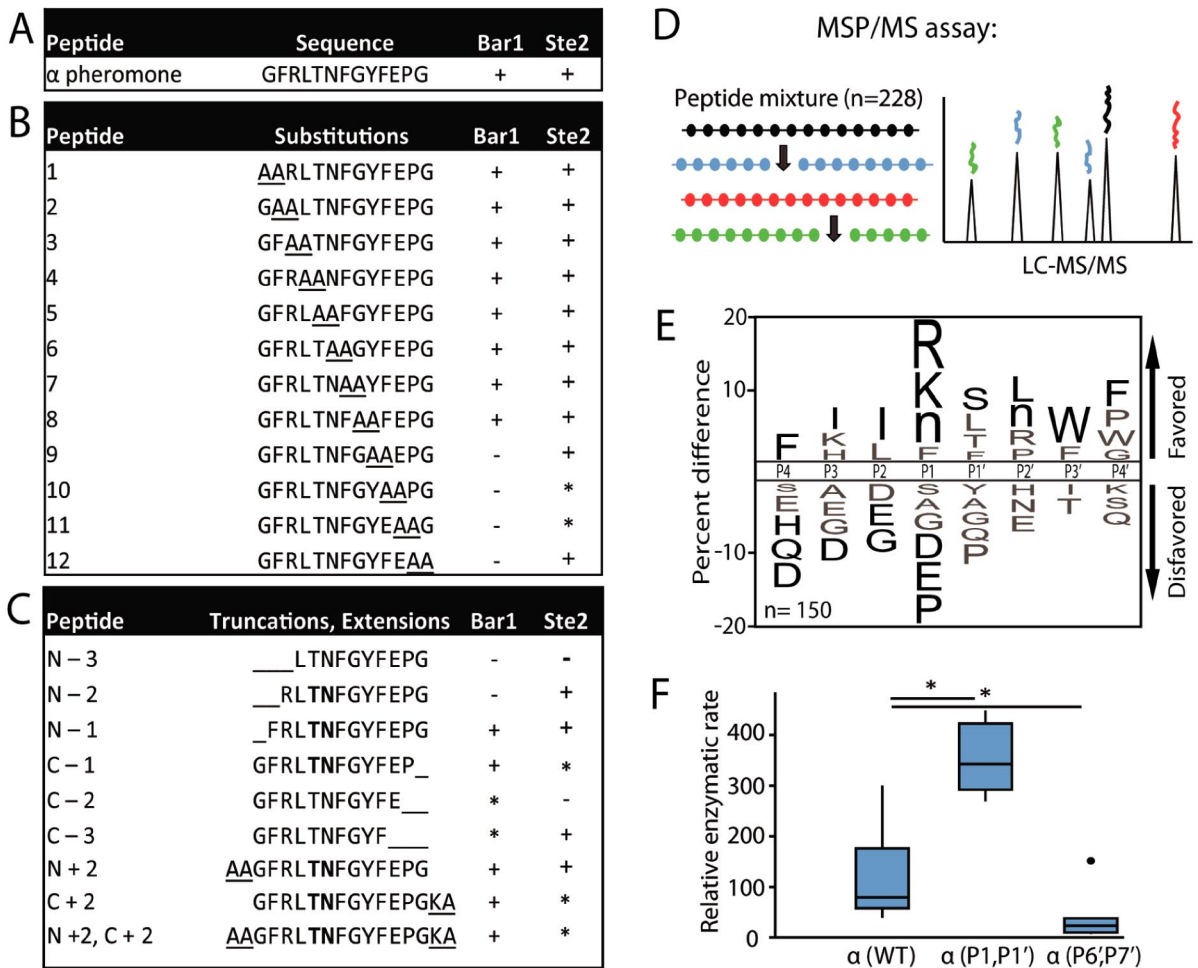
A

Peptide	Sequence	Bar1	Ste2
α pheromone	GFRLTNFGYFEPG	+	+

B

Peptide	Substitutions	Bar1	Ste2
1	AARLTNFGYFEPG	+	+
2	GAALTNFGYFEPG	+	+
3	GFAATNFGYFEPG	+	+
4	GFRAANFGYFEPG	+	+
5	GFRLAAFGYFEPG	+	+
6	GFRLTAAGYFEPG	+	+
7	GFRLTNAAYFEPG	+	+
8	GFRLTNFAAFEPG	+	+
9	GFRLTNFGAAEPG	-	+
10	GFRLTNFGYAAPG	-	*
11	GFRLTNFGYEAAG	-	*
12	GFRLTNFGYFEAA	-	+

C

Peptide	Truncations, Extensions	Bar1	Ste2
N − 3	LTNFGYFEPG	-	-
N − 2	RLTNFGYFEPG	-	+
N − 1	FRLTNFGYFEPG	+	+
C − 1	GFRLTNFGYFEP	+	*
C − 2	GFRLTNFGYFE	*	-
C − 3	GFRLTNFGYF	*	+
N + 2	AAGFRLTNFGYFEPG	+	+
C + 2	GFRLTNFGYFEPGKA	+	*
N +2, C + 2	AAGFRLTNFGYFEPGKA	+	*

D

MSP/MS assay:

E

F

Figure 6.4. CaBar1 degradation of α pheromone analogs. Recombinant CaBar1 was incubated for 1 h at 37°C with the indicated peptides and the products analyzed by mass spectrometry. (A) C. albicans α pheromone, (B) di-alanine substitutions within the C. albicans α pheromone sequence, (C) peptides corresponding to truncated or extended α pheromone. The detection of specific degradation products is indicated by "+" in the Bar1 column; "-" indicates no products were detected while "*" indicates products were formed only after extended (1 day) incubation. Bolded residues indicate amino acids flanking the cleavage site. The Ste2 column indicates if the pheromone induced a robust response (+), a weak response (*), or no response (-) in C. albicans MTLa cells using published data (Alby & Bennett, 2011). (D) Overview of the Multiplex Substrate Profiling by Mass Spectrometry (MSP-MS) assay. CaBar1 was co-incubated with 228 unique dodecapaptides for 1 h, and cleavage products identified using mass spectrometry. (E) MSP-MS data was used to generate an iceLogo identifying amino acids that were enriched or selected against in CaBar1 cleavage sites. Percent difference is the difference in amino acid frequency surrounding the cleavage sites relative to the frequency of amino acids surrounding all peptide bonds in the library (n = 2,964). Residues above the midline are favored, while those below the midline are disfavored. Residues colored black significantly influence CaBar1 activity (P < 0.05), whereas residues colored grey did not reach significance. Methionines were replaced with norleucines in the peptide library and are represented as "n". (F) Enzymatic activity of CaBar1 on substituted forms of α pheromone. Internally quenched peptides included ? pheromone with alanine substitutions at positions P1 and P1', or positions P6' and P7', corresponding to peptides 5 and 11 in panel B. Relative enzymatic rate is the amount of fluorescence per unit time due to cleavage of the fluorophore-conjugated peptide. Rates were normalized to the CaBar1(D232A) mutant control. Tukey boxplot with outliers noted as ●. N≥5. * =p<0.05 by Mann-Whitney U test.
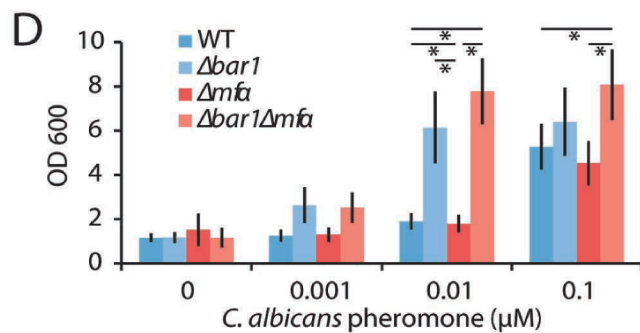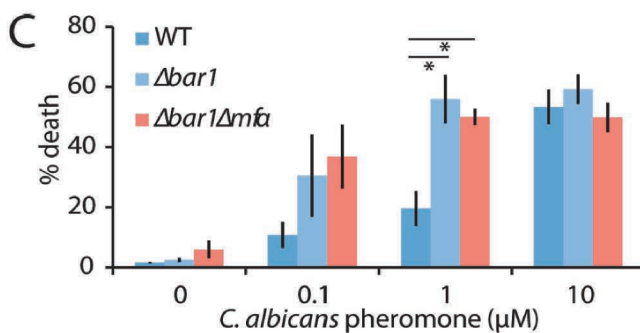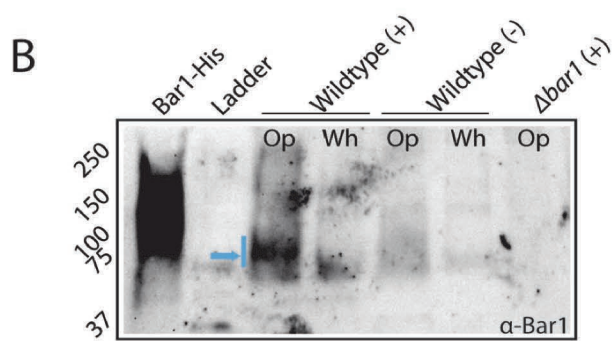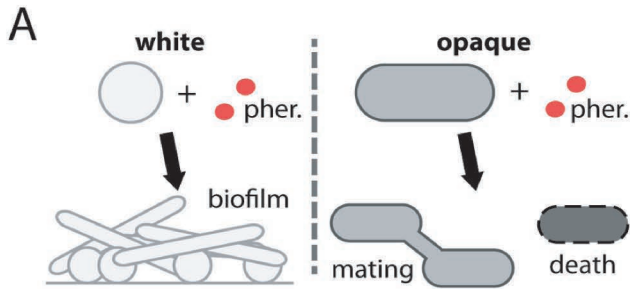
Figure 6.5. Analysis of the role of CaBAR1 on the pheromone response in white and opaque MTLa cells.
C. albicans wildtype, bar1Δ/bar1Δ, and bar1Δ/bar1Δ mfαΔ/mfαΔ MTLa strains were compared for their response to α pheromone. (A) Schematic indicating pheromone responses of *C. albicans* white and opaque cells. White cells become adherent and form biofilms, whereas a fraction of opaque cells experience cell death. (B) Secretion of CaBar1 by white or opaque MTLa cells. Cells were exposed to a vehicle control (-) or 0.3 μM α pheromone (+) for 5 hours, and CaBar1 protein detected from the supernatant via western blot using an anti-Bar1 antibody. Recombinant CaBar1-His protein is used as a positive control, and bar1Δ/bar1Δ opaque cells as a negative control. Blue arrow = dominant CaBar1 band. (C) Evaluation of gene function in opaque MTLa cells. A pheromone-induced death (PID) assay was used to assess the response of opaque a cells to α pheromone. Cells were treated with the indicated amount of α pheromone for 5 h, stained with propidium iodide, and percent death determined by flow cytometry. Error bars indicate SE. (D) Evaluation of gene function in white MTLa cells. A biofilm assay was used to assess the response of white cells to α pheromone. Cells were incubated with α pheromone for 2 days, and adherent cells quantified using absorbance at 600 nm. Error bars indicate SD. * = p<0.05 by T-test.
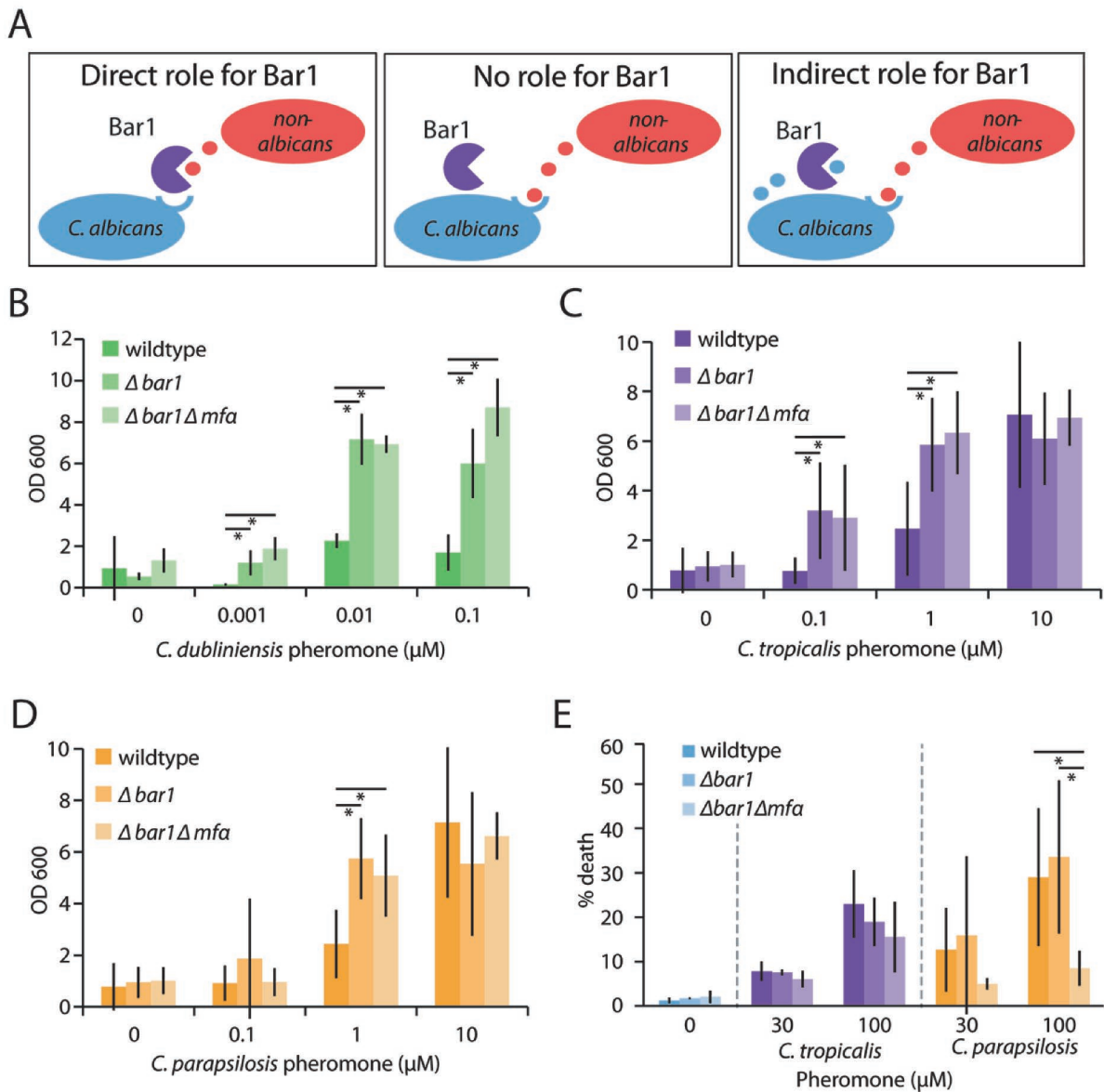
Figure 6.6. Role of CaBAR1 in inter-species pheromone signaling.
(A) Three alternative models for how CaBar1 may affect inter-species signaling between Candida species. (i) CaBar1 may affect signaling through direct recognition and degradation of non-native pheromones, (ii) it may have no effect on interspecies signaling events, or (iii) may play an indirect role in signaling due to degradation of C. albicans α pheromone (and blocking of autocrine signaling) following receptor activation with a non-native pheromone. *C. albicans* wildtype, bar1Δ/bar1Δ, and bar1Δ/bar1Δ mfαΔ/mfαΔ MTLa white cells were incubated with *C. dubliniensis* (B), *C. tropicalis* (C), or *C. parapsilosis* (D) α pheromones in a biofilm assay. After two days, adherent cells were quantified using spectrometry at 600 nm. Mean ± SD. * = p<0.05 by T-test. (E) C. albicans wildtype, bar1Δ/bar1Δ, and bar1Δ/bar1Δ mfαΔ/mfαΔ MTLa opaque cells were treated with C. tropicalis or C. parapsilosis α pheromones. The pheromone response was quantified by determining the percent cell death after 5 h of pheromone exposure using flow cytometry. Mean ± SD. * = p<0.05.
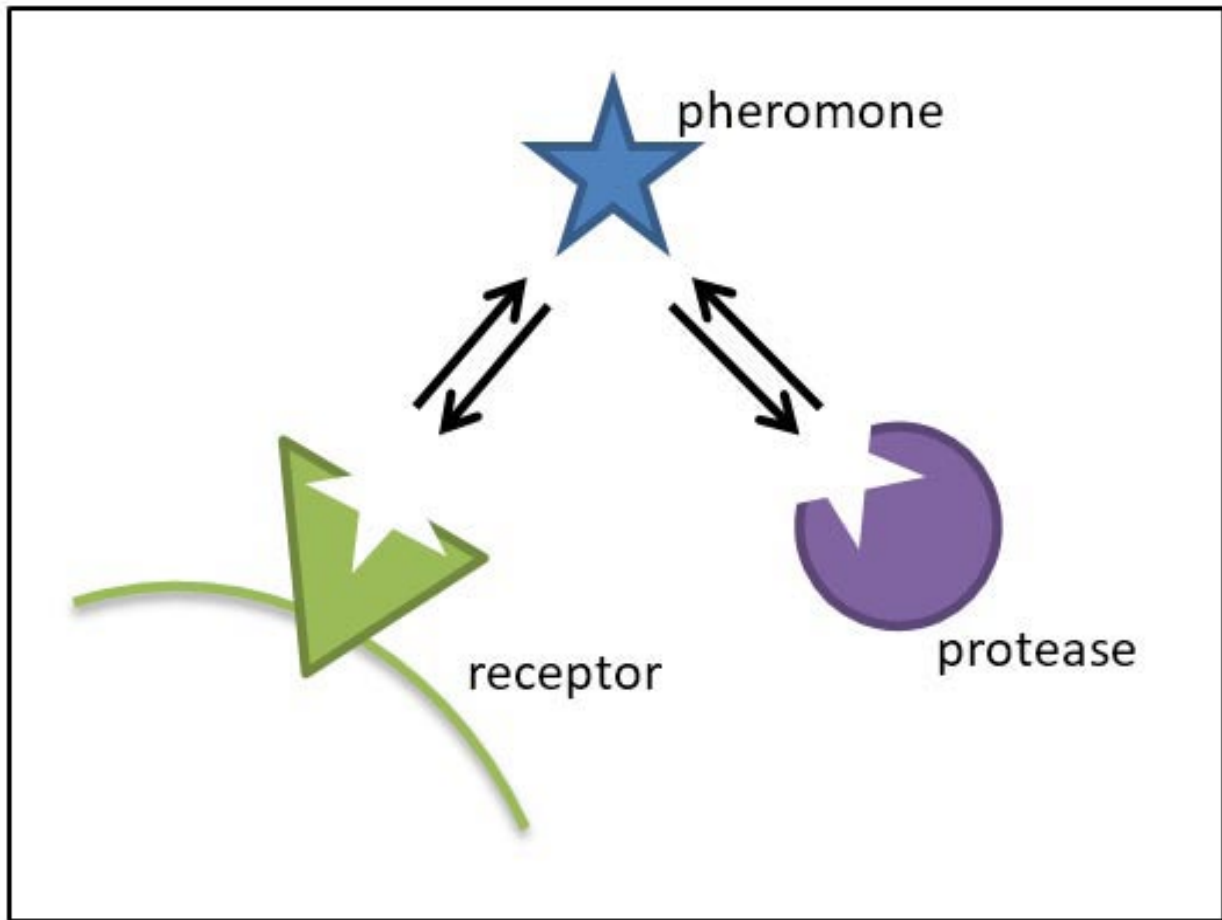
Figure 6.7. Model of interactions between Bar1, Ste2, and α pheromone.
Both Bar1 and Ste2 recognize the same, or overlapping, regions of α pheromone, thereby
facilitating co-evolution of their substrate specificities together with pheromone divergence
during speciation.

Chapter 7: Conclusions and Future Directions

7.1 Investigating the host-protein interactions of HIV protease

Much remains to be done before the role of HIV protease host-protein interactions will be fully understood. While many host substrates have been identified in recent years, a long list of candidate substrates have been identified through AP-MS assays but not yet studied and it is likely that others are still undiscovered (see Chapters 2 and 3 and manuscript under preparation Mahon et al). Significant advancements have been made in identification of candidate host-protein substrates using a mass spectrometry-based approach. This is an important improvement because studies validating HIV protease host-protein interactions are not easily amenable to high throughput approaches. Therefore, it is essential to implement the best initial screen possible as this focuses efforts on relevant interactions and avoid wasting time and resources pursuing artifacts. In work conducted primarily by Eckhardt and Mahon (manuscript under preparation), a screen for host protein substrates of HIV protease was conducted by infecting human T cells with wild type HIV virus with and without treatment by an HIV protease inhibitor. A method for labeling proteins called stable isotope labeling by amino acids in cell culture (SILAC) was used to label cells under one of the treatment conditions. Thus, when the lysates from cells treated with and without the inhibitor were combined, the proteins originating from each lysate could be distinguished and compared by mass spectrometry. Proteins found to decrease in molecular weight in cells not treated with the HIV protease inhibitor were determined to be potential host substrates (Fig 7.1). The design of this screen has several advantages over the initial AP-MS approach described in Chapter 2. The primary advantage is that it enriches for substrates of HIV protease rather than identifying all interaction partners. In addition, it reveals substrates which are cleaved in a context relevant to HIV infection, a factor that has been a serious roadblock in previous studies of HIV protease host-protein interactions (see Chapters 2 and 3).

Another challenge inherent to the study of HIV protease host-protein substrates is the identification of the sites of cleavage within the proteins. HIV protease is an unusual enzyme in that its substrate specificity is significantly influenced by the overall shape of the amino acids surrounding the scissile bond, rather than the sequence itself [1]. This presents a challenge when attempting to predict likely cleavage sites *de novo*. Studies to improve the understanding of the HIV protease substrate recognition mechanism and use this information to locate new cleavage sites are currently underway by Bohn, Mahon and Khuri. Their studies will likely result in an improved ability to screen new candidate host-proteins for potential sites of HIV proteolysis, a contribution that will greatly speed analysis of the role of these substrates in the viral life cycle.

Finally, the greatest challenge in the study of HIV protease host-protein interactions is determining the significance of these interactions and what role they play in viral replication. From the number of substrates identified thus far (Table 1.1) it appears that this enzyme impacts several host pathways. From previous studies of HIV protease host-protein substrates, it is plausible that each interaction makes a small contribution to viral replication; therefore interrogating the individual contribution of each substrate will require extremely sensitive approaches. One technique that we have begun to investigate is the use of nanosensors to detect HIV protease cleavage (see Chapter 3). Nanosensors can report single cleavage events potentially allowing us to quantify even very low levels of HIV protease activity within virally infected cells and determine if the activity is sufficient to cleave non-viral sequences [2]. In addition to this approach, the role of a particular substrate on viral replication can be examined by altering the protein such that it is no longer cleaved. After identifying the proteolysis site, a non-cleavable version of the substrate protein can be generated by mutating the P1 or P1' residues to an amino acid that is disfavored for cleavage (see Chapter 3). Replacement of the endogenous protein sequence with the non-cleavable version using the CRISPR/Cas system offers a method for dissecting the specific role of an individual substrate in the viral lifecycle [3].

This approach is currently being studied in the Krogan laboratory, however given the low throughput nature of this technique it is being reserved for host-protein substrates where the relevance of the interaction has already been verified.

7.2 The secreted peptidases of *Cryptococcus neoformans*

In the data presented in Chapters 4 and 5, we detail our novel findings regarding peptidases secreted by the fungal pathogen *C. neoformans.* Our evidence supports a role for these enzymes in pathogenesis and we provide a basis for the development of potent small-molecule inhibitors of a virulence-associated peptidase. There are several follow-up studies that are suggested by our results. We uncovered the substrate specificity and function of a principle enzyme contributing to extracellular peptidase activity (May1) however, in this study we do not address which substrates it cleaves to influence tolerance to acidic environments and virulence. Our results indicate that May1 cleaves fungal proteins and potentially host proteins as well. Determining the fungal substrates will likely be more straightforward than identifying host protein substrates given the genetic tractability of this organism. While an approach similar to the AP-MS technique used with HIV protease is possible for May1, the multiple post-translational modifications of fungal proteins pose a barrier to mass spectrometry in *C. neoformans* [4]. Because May1 deletion strains have an easily detectable phenotype, the substrate(s) can be identified by screening deletion collections of secreted proteins and identifying mutants with a similar phenotype. This approach has proven successful for elucidation of the components of other pathways in *C. neoformans*, as we detail in Chapter 5.

Our results provide a basis for the development of potent small molecule inhibitors of May1. However, when we tested these compounds in a fungal culture model we found that they were less efficacious than anticipated based on their *in vitro* characteristics. This indicates a need for further medicinal chemistry to improve solubility and stability of these compounds. We must also investigate the specificity of these compounds for aspartyl peptidases found within the

host organism before considering testing their effectiveness in an *in vivo* infection model. Further design may be required to improve the selectivity of these inhibitors as informed by our MSP-MS analysis of May1 specificity.

Another area of investigation stemming from our results is the determination of the functions of the additional secreted peptidases that we have identified. Our findings provide clues as to the pathways that some of these enzymes may be involved in, for example we determine that Prb1 has an impact on melanin production, while Cxd1is potentially involved in nutrition. However further studies investigating the functions of these enzymes are necessary as well as experiments in a mouse model of cryptococcus infection to address whether or not these proteolytic activities affect virulence. An abundant serine peptidase activity detected in our experiments after growth in host-like conditions also remains to be assigned to a gene (or genes). We propose a candidate enzyme responsible for this activity, Kex2, but this deletion strain will need to be generated to test our hypothesis.

7.3 Bar1: an unconventional peptidase

In Chapter 6 we describe biochemical and functional characterization of a unique aspartyl peptidase secreted by the fungal pathogen *C. albicans*, called Bar1. This peptidase has an important role in regulating mating by degrading α pheromone, an initiator of the mating response in **a** cells. We detail numerous surprising characteristics of Bar1 including its resistance to the aspartyl peptidase inhibitor pepstatin A, its pH optimum of 6.5 and most unusual of all, its mechanism of substrate selectivity. Whereas peptidase substrate selectivity is typically driven by the residues forming the scissile bond and those closest to the scissile bond, Bar1 appears to recognize its substrates by binding to residues in the distal carboxyl terminus of the protein, positions P5'-P8' of its substrates. We put forth an argument that this specificity evolved to match the specificity of the receptor for α pheromone, which also appears to bind the

C-terminus of this peptide. However, to conclusively test this we need to investigate the binding of Bar1 and α pheromone in more detail. A crystal structure will reveal the contacts that are made between Bar1 and its substrate, while further peptide mutagenesis experiments similar to those described in Chapter 6, can reveal the precise tolerance of Bar1 to amino acids in each position of α pheromone.

Another direction suggested by our studies of Bar1 is the commercial use of this enzyme to remove C-terminal tags from proteins. Tags are frequently added to proteins for purification purposes, however they can interfere with downstream applications and thus they are usually removed [5]. An enzyme such as Bar1 which cleaves upstream from its recognition site has the potential to allow complete removal of a protein tag without leaving a several amino acid overhang, as with most other peptidases. Our studies thus far indicate that Bar1 is a highly specific enzyme, an advantage in protein-tag removal applications; however we will need to further understand the factors contributing to its substrate specificity before it can be optimized for commercial use. Structural studies of Bar1 and α pheromone will likely be informative. We will also need to investigate the ability of Bar1 to cleave a substrate within the context of a protein. Because the native substrate of this enzyme is a short peptide it is possible that engineering of Bar1 will be required to optimize it for cleaving a protein substrate.

Identification and characterization of secreted peptidases is a topic of great interest to those studying pathogenic organisms. Not only are these enzymes viable drug targets, they can also be useful for diagnostic purposes through the development of specific substrates [6,7]. As detailed in the data presented here, their study has the potential to reveal important host-pathogen interactions, the elucidation of which may provide new opportunities for alternative therapeutic intervention and reveal critical differences between benign and disease-causing organisms.

7.4 References

1. Nalam MN, Ali A, Altman MD, Reddy GS, Chellappan S, et al. (2010) Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. J Virol 84: 5368-5378.
2. Tajon CA, Seo D, Asmussen J, Shah N, Jun YW, et al. (2014) Sensitive and selective plasmon ruler nanosensors for monitoring the apoptotic drug response in leukemia. ACS Nano 8: 9199-9208.
3. Lee J, Chung JH, Kim HM, Kim DW, Kim H (2015) Designed nucleases for targeted genome editing. Plant Biotechnol J.
4. Campbell LT, Simonin AR, Chen C, Ferdous J, Padula MP, et al. (2015) Cryptococcus strains with different pathogenic potentials have diverse protein secretomes. Eukaryot Cell 14: 554-563.
5. Waugh DS (2011) An overview of enzymatic reagents for the removal of affinity tags. Protein Expr Purif 80: 283-293.
6. Kaman WE, Hays JP, Endtz HP, Bikker FJ (2014) Bacterial proteases: targets for diagnostics and therapy. Eur J Clin Microbiol Infect Dis 33: 1081-1087.
7. Aoki W, Kitahara N, Fujita A, Shibasaki S, Morisaka H, et al. (2013) Detection of Candida albicans by using a designed fluorescence-quenched peptide. J Biosci Bioeng 116: 573-575.
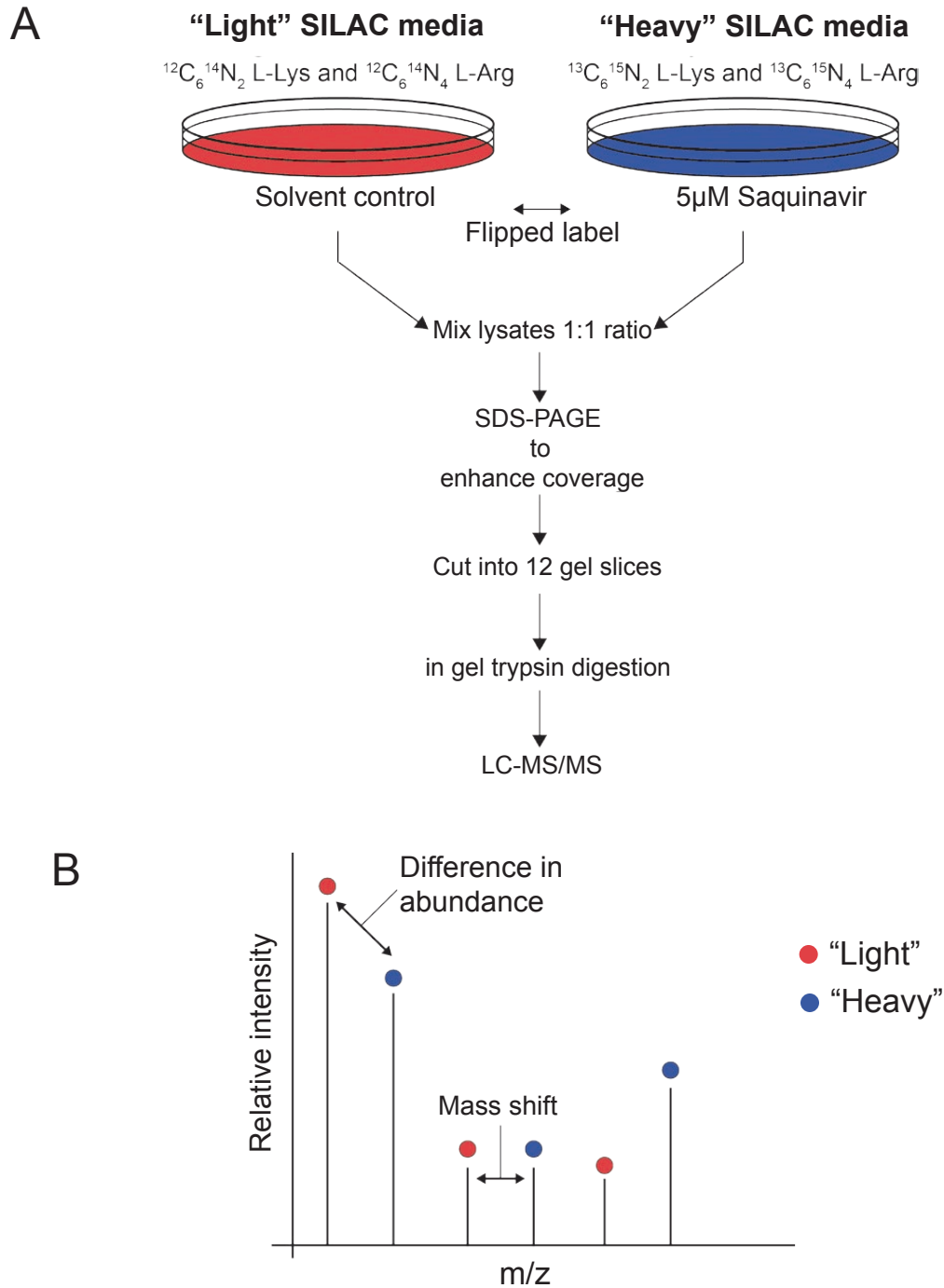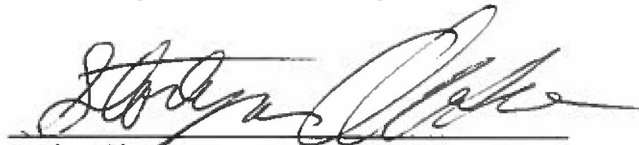
Figure 7.1. Schematic of screen for HIV protease host-protein substrates using SILAC.
(A) Cells cultured in "heavy" or "light' SILAC media are infected by wild type HIV and one sample is treated with the HIV protease inhibitor Saquinavir. To control for changes due to the media alone, two experiments are performed where the inhibited sample is "flipped". Lysates are then made from the infected cells, combined and run on a protein gel. The gel is cut into slices which are then analyzed by LC-MS/MS. (B) Because proteins coming from each condition can be distinguished by their label, it is possibile to compare the relative abundance and mass shift of proteins. Schematic curtesy of Dr. Manon Eckhardt (UCSF).

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:
I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

_____   9-23-15
Author Signature   Date