UCSF
UC San Francisco Electronic Theses and Dissertations

Title
The characterization of a family of secreted aspartyl proteinases in Candida albicans

Permalink
https://escholarship.org/uc/item/2k02x0sw

Author
Miyasaki, Shelley H.

Publication Date
1997

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

The Characterization of a Family of Secreted Aspartyl Proteinases in *Candida albicans*

by

Shelley H. Miyasaki

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Oral Biology

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

UCSF LIBRARY

Copyright (1998)

by

Shelley Hiroko Miyasaki

Acknowledgements

Portions of Chapter 2 are reprinted with permission from:

The Identification and Tracking of *Candida albicans* Isolates from Oral Lesions in HIV-Seropositive Individuals.

Journal of Acquired Immune Deficiency Syndromes. 1992. 5:1039-1046.
©1992 Raven Press, Ltd., New York

Portions of Chapter 3 are reprinted with permission from:

Three Distinct Secreted Aspartyl Proteinases in *Candida albicans*.

Journal of Bacteriology. 1993. 175:6126-6133.
© 1993, American Society for Microbiology

Portions of Chapter 5 are reprinted with permission from:

A Fourth Secreted Aspartyl Proteinase Gene (*SAP4*) and a *CARE2* Repetitive Element are Located Upstream of the SAP1 Gene in *Candida albicans*.

Journal of Bacteriology. 1994. 176:1702-1710.
© 1994, American Society of Microbiology

This thesis dissertation has been completed by Shelley Miyasaki in fulfillment of the requirements of the Oral Biology Graduate Group in the School of Dentistry at the University of California, San Francisco, California.

In completing this research, Shelley has performed both independent research and has also collaborated with her laboratory colleagues who appear as co-authors on the three manuscripts included in this dissertation. In each of these cases, Shelley has made important, substantiative and unique contributions to the research for these manuscripts.

In the first manuscript, The identification and tracking of *Candida albicans* isolates from oral lesions in HIV-seropositive individuals - S.H. Miyasaki, J.B. Hicks, D. Greenspan, I. Polacheck, L.A. Macphail, T. C. White, N. Agabian and J.S. Greenspan (authors), Shelley was responsible for the majority of the work presented in the study, bringing her clinical training and expertise to bear on the development and testing of a DNA fingerprinting methodology for the identification of *C. albicans* in clinical isolates. Her studies led to the realization

that *C. albicans* strains are vast in number and each appears to have the ability to express both commensal and pathogenic phenotypes depending upon the host immune status. Her studies also provided some of the first evidence for the potential transfer of *C. albicans* between sexual partners.

In the second manuscript, Three distinct secreted aspartyl proteinases in *Candida albicans* - T.C. White, S.H. Miyasaki, N. Agabian (authors), Shelley again performed a significant portion of these studies. She designed and prepared the PCR product which she then used for the identification and isolation of several of the then unknown secreted aspartyl proteinases (Saps). She focused on the characterization of the complex *SAP* gene family in *C. albicans* strains which exhibited differential gene expression of *SAP*1 under various switch phenotype and growth conditions. She applied her findings to the characterization of the clinical isolates and was able to test the importance of specific restriction fragment length polymorphisms (RFLPs) in the expression of members of this gene family. Her studies revealed that the differences in *SAP*1 alleles could be attributed to the insertion of a repetitive sequence element upstream of one of the two alleles of *SAP*1 in an expressor strain. She further demonstrated that the insertion of this element was benign in that it did not prevent or otherwise modulate the expression of the *SAP*1 allele downstream. Her analysis of *SAP*1 in expressor and non-expressor strains did not reveal any potential *cis*-acting regulatory regions which might account for the expression of *SAP*1 in certain phenotypic switching, suggesting that the expression of switch phenotype specific trans-acting elements may be important in regulating *SAP*1 activity.

The third manuscript, A fourth secreted aspartyl proteinase gene (*SAP*4) and a *CARE*2 repetitive element are located upstream of the *SAP*1 gene in *Candida albicans* - S.H. Miyasaki, T.C. White, N. Agabian (authors), describes the identification and sequencing of a fourth secreted aspartyl proteinase gene. These studies were the outcome of characterization of the upstream regulatory regions of *SAP*1. Shelley was an important contributor to the identification and cloning of this fourth *SAP*.

These studies are comparable in scope and complexity to those required by the Oral Biology Graduate Group for the completion of dissertation requirements towards a Ph.D. in Oral Biology. Shelley, as one of the first students in this program has carried out these studies while also completing her residency in Oral and Maxillofacial surgery. Her work has been of the highest professional calibre and she has set an exemplary standard for the combined degree program throughout her studies and productivity.

Nina Agabian Ph.D.
Professor Molecular Pathogenesis
Department of Stomatology.

DEDICATION

This dissertation is dedicated to my family, friends and mentors, who have always stood behind me in all of my endeavors.

LIBRARY

The Characterization of a Family of Secreted Aspartyl Proteinases in *Candida albicans*

Shelley Hiroko Miyasaki

ABSTRACT

The opportunistic fungus *Candida albicans* continues to be a major cause of morbidity and mortality in immunocompromised hosts. Little is known about the adaptive capabilities that enable this previously commensal organism to become pathogenic. This thesis is actually series of investigations designed to characterize at a molecular level first, the individual types of *Candida* isolates which infected a population of human immunodeficiency virus (HIV)-infected men; and second, a family of secreted aspartyl proteinases (Saps) which may function as putative virulence factors in *C. albicans*. The first study showed that although a large number of different *C. albicans* isolates exist in a population of immunocompromised patients, for the most part, each person maintains a unique strain throughout the course of their HIV-infection. Each isolate of *C. albicans* must therefore be capable of causing disease in the right environment. What fungal characteristics might enable this commensal to become a pathogen? The following studies focused on the identification and characterization of a family of putative virulence factors, the Saps. Secreted aspartyl proteinases have been identified in a number of *Candida* species which are pathogenic in humans. The *C. albicans SAP* 3 and *SAP* 4 genes were isolated and sequenced. Additionally, the patterns of expression of *SAP*s 1,2 and 3 in the *C. albicans* strain WO-1 were correlated with different phenotypic switch states. This correlation is perhaps suggestive of differing functional roles for the various Saps. Sequencing of the flanking regions from the four alleles of *SAP*1 in two different strains of *C. albicans* identified some potential *cis*-acting transcriptional regulatory regions. Future studies will test the functionality of these regions using a mutant *Aequorea victoria* green fluorescent protein reporter system in an auxotrophic isolate of the phenotypic switching strain WO-1.

CONTENTS

List of Tables vii
List of Figures viii

Chapter 1. Introduction and Review of the Literature 1

References 7

Chapter 2. The Identification and Tracking of *Candida albicans* Isolates from Oral Lesions in HIV-Seropositive Individuals 11

Abstract 12
Introduction 13
Materials and Methods 14
Results 19
Discussion 22
References 29

Chapter 3. Candidal Virulence Factors and Pathogenesis 40

References 44

Chapter 4. Three Distinct Secreted Aspartyl Proteinases in *Candida albicans.* 47

Abstract 48
Introduction 49
Materials and Methods 52
Results 57
Discussion 65
References 71

Chapter 5. A Fourth Secreted Aspartyl Proteinase Gene (*SAP*4) and a *CARE*2 Repetitive Element are Located Upstream of the *SAP*1 Gene in *Candida albicans.* 86

Abstract 88
Introduction 89
Materials and Methods 93
Results 98
Discussion 106
References 112

Chapter 6. Future Directions 129

References 135

LIST OF TABLES

Chapter 2

Table 1.	Recurrent Oral Candidiasis Despite Antifungal Treatment in One HIV-Seropositive Subject	36
Table 2.	Presumed Candidal Transmission Between Sexual Partners	39

Chapter 3.

Table 1.	Putative Virulence Factors of *Candida albicans*	46
Table 2.	Differences Between White and Opaque Cells	46

Chapter 4.

Table 1.	Characteristics of the *SAP* Gene Products	83

Chapter 5.

Table 1.	Differences Between Coding Regions in Alleles from Three Strains of *C. albicans*	120
Table 2.	Sequence Differences Between Flanking Regions of Alleles from Three Strains of *C. albicans*	121

UCSF LIBRARY

LIST OF FIGURES

Chapter 1.

Figure 1.	Pseudomembranous Candidiasis	8
Figure 2.	Erythematous Candidiasis	9
Figure 3.	Angular Cheilitis	10

Chapter 2.

Figure 1.	*C albicans* Isolates Show a High Degree of Polymorphism	32
Figure 2.	Patients Maintain A Single Isolate of *C. albicans* Over Time	34
Figure 3.	Sexual Partners May Share the Same *C. albicans* Isolate	37
Figure 4.	*C. albicans* Strains May be Transmitted Between Sexual Partners	38

Chapter 4.

Figure 1.	Detection of Three Proteinases	75
Figure 2.	Expression of Three Proteinases During Growth	76
Figure 3.	Isoelectric Focusing of the Three Isoenzymes	78
Figure 4.	N-Terminal Sequences of Secreted Aspartyl Proteinases	79
Figure 5.	Sequence of the *SAP*3 Gene	81
Figure 6.	Genomic Blot Showing Genomic Organization of Three Proteinases	84

Chapter 5.

Figure 1.	RFLP Mapping Distinguishes Three Alleles in Laboratory Strains	117
Figure 2.	Restriction Maps and Summary of Major Allelic Differences	118
Figure 3.	Sequence of the 5' -Flanking Region of the WO-1 2.6 Allele	122
Figure 4.	Both the 2.4 and 2.6 *SAP*1 Alleles of Strain WO-1 are Transcribed in RNAs from O Cells Grown in YCB-BSA.	123
Figure 5.	Sequence of the *SAP*4 Gene and Deduced Amino Acid Sequence of the Protein	125
Figure 6.	Comparisons of the Deduced Protein Sequences of the Four *SAP* Genes	127

Chapter 6.

UCSF LIBRARY

CHAPTER 1

INTRODUCTION AND REVIEW OF THE LITERATURE

The clinical condition known as "thrush" has been a scourge since ancient times. Hippocrates and Galen (Odds 1988) are thought to have provided the first descriptions of the affliction. The microorganism responsible for these infections was not identified until 1842, when Gruby correctly associated the disease with the fungus, *Candida albicans* (Ref Gruby 1842). The disease in its many forms has continued to plague humankind. With the onset of both novel immunosuppressive disease states and therapies, candidiasis has become one of the most frequent and well-recognized opportunistic infections of the twentieth century.

The species most often associated with oral fungal infections in humans is *Candida albicans*. *C. albicans* is medically the most important of the yeasts and thought to be responsible for the majority of infections caused by this genus. However, with the advent of drug resistance to azoles, other species such as *C. tropicalis, C. glabrata, C. parapsilosis, C. stellatoidea, C. krusei and C. kyfer* have taken on added importance. The genus *Candida* is a member of the class *Deuteromycetes* which includes the "imperfect yeasts," so called because they do not have a sexual or haploid stage. This is an important distinction, because the absence of sexual recombination is likely to limit the ways in which *C. albicans* is able to generate genetic diversity. *C. albicans* may have evolved other means of generating genetic variation, such as the phenomenon of

USF LIBRARY

spontaneous phenotypic switching, in order to compete and survive in a constantly changing environment.

C. albicans has been termed a "dimorphic yeast" because it is capable of switching between elongated hyphal forms and budding yeast forms as well as pseudohyphae. We also now know that this species is capable of rapidly switching between many more phenotypes. A description of these spontaneously generated "switch phenotypes" as well as their interrelationships with other factors involved in pathogenesis, will be discussed in a later chapter with respect to their putative roles as virulence factors in *C. albicans*.

A variety of epidemiological surveys have found *C. albicans* as a normal inhabitant of the oral cavity in upwards of 10 to 75% of the population studied (Reviewed by Odds 1988). The onset of morbidity is correlated with compromise of the host's immune system and/or changes in the oral flora or oral environment. Both cellular and humoral immunodeficiencies can predispose one to overgrowth and subsequent suprainfection with *C. albicans*. Although this pathogen is capable of causing systemic (including both deep organ and bloodborne) infections (Bodey 1985), it is best known as the cause of superficial or mucocutaneous infections, in immunocompromised patients. In this patient population, *C. albicans* infections can be a significant source of morbidity and mortality. Depending upon the cause and severity of the immunodeficiency, the microenvironment supplied by the host, and as yet undefined factors present in

the parasite, the infection can be localized or systemic and exhibit different manifestations.

Oral mucocutaneous candidal infections can be classified into four different types based upon their clinical appearance as described by Greenspan et al. (1990). The first and most familiar type is pseudomembranous candidiasis, commonly known as "thrush." This lesion is shown in Figure 1 and is characterized by the presence of removable white plaques or "pseudomembranes," which when removed, reveal a bleeding ulcerated base. These plaques are composed of epithelial cells, materia alba and numerous *Candida* hyphae. The second type, illustrated in Figure 2, is erythematous candidiasis. This lesion is much more subtle and is often overlooked. It appears as red patches which on a microscopic level are seen to be atrophic epithelium. When erythematous candidiasis appears on the tongue, it can cause atrophy of the filiform papillae. Interestingly, these lesions are commonly associated with a paucity of fungal organisms, which are often only detectable by culture.

The third type is hyperplastic candidiasis (not shown). The lesion is most commonly associated with smoking and consists of white plaques which do not rub off. Microscopically, these lesions consist of hypertrophic and hyperplastic epithelium with candidal hyphae invading the most superficial aspects of the lesion. The fourth type of lesion is known as angular cheilitis. It is seen in Figure 3 and is characterized by erythema, and cracking of the corners of the mouth and lips. All of the lesions can cause pain, soreness and sometimes changes in taste.

Factors which can cause an immunocompromised state and therefore predispose one to an overgrowth of *C. albicans* include: systemic malignancies (Boggs 1961), defects in cell-mediated immunity including infection with the human immunodeficiency virus (HIV), and other rare specific T-cell deficiencies (Rothschild 1976), endocrinopathies such as diabetes mellitus (Odds et al.1987, Tapper-Jones et al. 1981), and iatrogenically induced states of altered microbial flora competition or immunocompromise such as those found in patients taking antibiotics and steroids, or undergoing radiation or chemotherapy.

Examples of states in which alteration of the host oral environment can lead to overgrowth of *Candida* include: Patients who have alterations in either the quality or quantity of saliva which is produced. Commonly encountered categories of patients with decreased salivary flow are those who have undergone radiation therapy for a head and neck malignancy and those with alterations in their salivary flow due to treatment with widely used drugs such as the phenothiazine antipsychotics, antihypertensives and antihistamines.

The last group of patients at risk of fungal overgrowth are those who wear dentures. *Candida* are highly adherent to plastic surfaces and the small porosities in the denture material provide excellent growth environments for the fungus. In addition, the wearing of dentures also causes changes in the host mucosal surfaces which can contribute to suprainfection by *Candida*. Suprainfection may be enhanced as saliva, with its inherent buffering capacity

and antimicrobial agents such as lactoferrin, histatins and secretory IgA , is not readily available underneath a denture.

My specific interests have centered on immunodeficiency states caused by infection with the human immunodeficiency virus or HIV. In this population, suprainfection with *C. albicans* is a major source of morbidity and in some cases mortality. Thrush was one of the primary disease manifestations to be recognized in association with this progressive immunosuppressive disease (Jaffe 1983). In conjunction with HIV infection, *C. albicans* is primarily associated with superficial mucocutaneous infections, but deep organ infections, such as esophageal candidiasis, are not unknown.

It is apparent that there are many ways to affect the local host environment and predispose one to the onset of a candidal infection. However, not all affected people will develop a candidal infection. Moreover, of those who do develop an infection, it is impossible to predict which type of infection they will develop. What factors make it possible for this previously commensal organism to suddenly become pathogenic? This very general question formed the basis for my focus on the identification and roles of some of the putative virulence factors in *C. albicans*.

Both the host and the fungus play a role in determining the balance of the delicate equilibrium which exists in the commensal state. For the host, as stated above, immune competence or incompetence plays a major role. The fungus, on the other hand, is likely to possess virulence factors which may potentiate

opportunism and/or a tendency toward infection. To investigate the relationship between host immune status and the pathogenicity of a given *C. albicans* strain, I first asked if perhaps more virulent strains of *Candida* are selected for in immunosuppressed patients. In these initial studies, using molecular fingerprinting methods, I identified and tracked over time, the candidal isolates in a population of HIV-infected men. The methods and findings of this study are detailed in Chapter 2.

REFERENCES - CHAPTER 1

Bodey, G. P., and V. Fainstein. 1985. Systemic candidiasis. In *Candidiasis*, ed. G. P. Bodey, V. Fainstein, 135-68. New York: Raven Press.

Boggs, D. R., A. F. Williams, and A. Howell. 1961. Thrush in malignant neoplastic disease. *Arch. Intern. Med.* 107:354-60.

Greenspan, D., M. Schiødt, J. S. Greenspan, J. J. Pindborg. 1990. *AIDS and the mouth*. Copenhagen: Munksgaard.

Gruby, M. 1842. Recherches anatomiques sur une plant cryptogame qui constitue le vrai muguet des enfants. *C. R. Acad. Sci. (Paris)* 14:634-6.

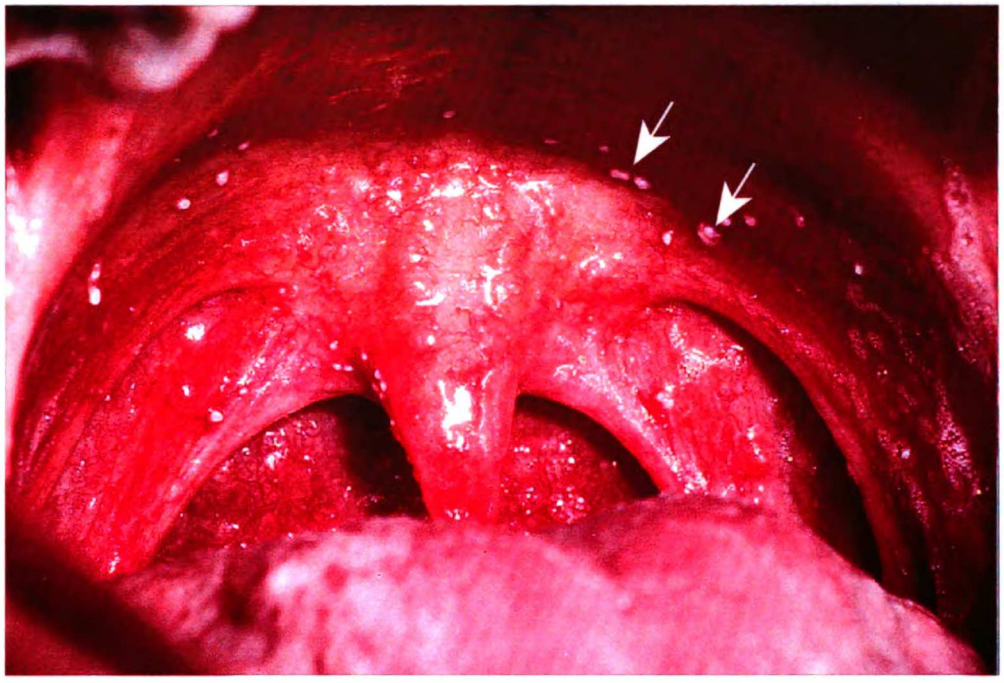
Jaffe, H. W., D. J. Bregman, and R. M. Selik. 1983. Acquired immune deficiency syndrome in the United States: the first 1000 cases. *J. Infect. Dis.* 148:339-45.

Odds, F. C. 1988. *Candida and Candidosis*. London: Baillière Tindall.

Odds F. C., E. G. V. Evans, M. A. R. Taylor, J. K. Wales. 1987. Prevalence of pathogenic yeasts and humoral antibodies to *Candida* in diabetic patients. *J. Clin. Pathol.* 31:840-44.

Rothschild H., M. Wilson, M. Lopez, J. Salvaggio, D. Bice. 1976. An immunological investigation of a family with chronic mucocutaneous candidiasis. *Int. Arch. Allergy Appl. Immunol.* 52:291-6.

Tapper-Jones L., J. M. Aldred, D. M. Walker, T. M. Hayes. 1981. Candidal infections and populations of *Candida albicans* in the mouths of diabetics. *J. Clin. Pathol.* 34:706-11.



NOSE LIBRARY

Figure 1

Pseudomembranous Candidiasis

The arrows point to the white removable plaques, which are composed of fungal hyphae, materia alba and epithelial cells.

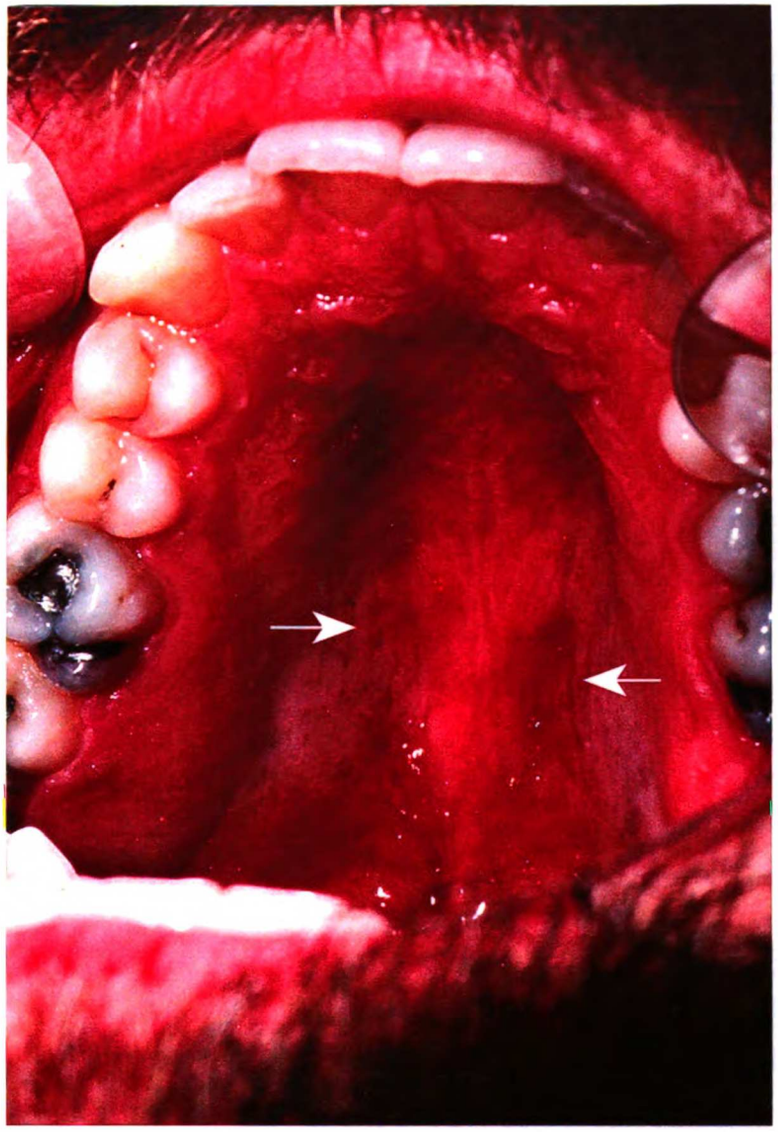


Figure 2

Erythematous Candidiasis of the palate

The arrows denote areas of macular erythematous candidiasis.

UCSF LIBRARY

Figure 3

Angular Cheilitis

USF LIBRARY

CHAPTER 2

The Identification and Tracking of *Candida albicans* Isolates from Oral Lesions in HIV-Seropositive Individuals

Portions of the following chapter have been previously published in:

Journal of Acquired Immune Deficiency Syndromes

1992

Vol. 5, No. 10

pp. 1039-1046

Shelley H. Miyasaki[1], James B. Hicks[3], Deborah Greenspan[1], Itzhack Polacheck[4], Laurie A. MacPhail[1], Theodore C. White[2], Nina Agabian[2] and John S. Greenspan[1]

From the [1]Department of Stomatology and Oral AIDS Center, University of California San Francisco, and [2]University of California San Francisco and Berkeley Intercampus Program in Molecular Parasitology, San Francisco, California; [3]Department of Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, California; and [4]Department of Clinical Microbiology and Infectious Disease, Hadassah Medical Center, Jerusalem, Israel

UCSF LIBRARY

Chapter 2

CANDIDA ISOLATES IN HIV INFECTION

ABSTRACT

Restriction fragment polymorphism analysis was used to investigate the identity and genotypic relatedness of *Candida albicans* strains isolated from human immunodeficiency virus (HIV)-infected patients with or without oral candidiasis and from some of their sexual partners. Use of the species-specific DNA probe Ca3 revealed that most subjects carried a single distinct *Candida albicans* strain throughout the course of the study, during both symptomatic and asymptomatic periods. Sexual partners were more likely to carry the same or similar *C. albicans* isolates than unrelated subjects, raising the possibility of transmission via intimate contact. One patient appeared to acquire his partner's isolate, which then became predominant in both partners in subsequent isolations. These findings indicate that recurrent oral candidiasis is usually caused by a single persistent strain unique to each patient, but that in some cases transmission via intimate contact may occur between sexual partners.

Keywords: *Candida albicans*; oral candidiasis, Fingerprinting; repetitive DNA.

USF LIBRARY

INTRODUCTION

Candida albicans is recognized as the most common cause of oral candidiasis in immunosuppressed persons. It has been estimated that one in two individuals with human immunodeficiency virus (HIV) infection will suffer from oral candidiasis during the course of their disease (Samaranayake and Holmstrup 1989, Feigal et al. 1991). The significance of oral candidiasis as a disease entity in HIV-related immunosuppression is its frequent recurrence and, in some cases, its spread to the esophagus. A study by Carpenter et al. (1989) of 62 HIV-infected women found esophageal candidiasis to be the most frequent AIDS-defining event, occurring even more frequently than *Pneumocystis carinii* pneumonia in that group. Efforts to explain why and how candidiasis develops in HIV infection have focused on the underlying host cellular immunodeficiency. Accordingly, little is known about the biotypes and genotypes of *C. albicans* present in these individuals. The mechanism behind the ability of this fungus to cause recurrent disease is unknown, although in several studies a carrier state was found in upwards of 30-60% of subjects (Martin and Wilkinson 1983, Wright, Clark and Hardie 1985). Whether or not the *Candida* strains harbored by these persons act as a pool from which recurrent oral candidiasis might develop during immunosuppression has not been well documented.

We used the techniques of restriction fragment length polymorphism analysis (RFLP) and nucleic acid hybridization to characterize the molecular epidemiology of *Candida* in a group of HIV-infected patients seen at the

University of California San Francisco Oral AIDS Center. Isolates were obtained from patients before, during and after the development of oral candidiasis and associated antifungal treatment. For comparison, oral cultures were also obtained from a second group of HIV-seronegative individuals who were free of oral candidiasis. In addition, we investigated the genotypic relatedness of strains isolated from sexual partners of HIV-seropositive patients to evaluate the possibility of transmission of *C. albicans* strains via intimate contact.

METHODS

Study Populations.

Study participants included 33 HIV-seropositive persons, 12 of whom had no history of oral candidiasis at the time of their entry into the study. This group was recruited from patients seen in the Oral AIDS Center Clinic at the University of California, San Francisco and consisted of 32 homosexual men and one heterosexual woman. Subjects were selected upon the basis of fulfillment of the following specific study criteria: (a) presence or absence of oral candidiasis; (b) documentation of HIV serostatus; (c) ability to attend frequent recall visits to the clinic; and (d) willingness to sign a consent form in order to give informed consent for participation in the study.

They were enrolled between September 1988 and April 1990 and followed up for periods ranging from 2 to 15 months. Oral samples were taken from asymptomatic subjects at 3-month intervals for *Candida* culture. This was done by rubbing a sterile cotton swab over the intraoral soft tissues including the

UCSF LIBRARY

buccal mucosa, hard and soft palates, gingiva and tongue. Those who had, or developed, signs or symptoms of oral candidiasis were diagnosed by the criteria described by Greenspan et al. (1990), and oral samples were obtained from those persons at the time of their diagnosis. Both nonlesional and lesional sites were sampled when present in the same individual. All samples were analyzed separately. Symptoms of oral candidiasis were often but not always present, and included alteration in taste or the development of a painful or burning sensations. The criteria for diagnosis of the four major types of oral candidiasis are briefly summarized as follows: Pseudomembranous candidiasis is characterized by the presence of white removable plaques, which may be scraped away to reveal a red base. Erythematous candidiasis presents as erythematous patches upon the oral mucosa that may be extremely subtle and therefore more difficult to diagnose. Hyperplastic candidiasis is uncommon among HIV-infected subjects and is characterized by white plaques that are not removable by scraping. Angular cheilitis appears as fissuring and ulceration of the oral commissures. The diagnosis of candidiasis relied both upon the clinical presentation and the presence of *Candida* hyphae on smears examined by potassium hydroxide, periodic acid-Schiff (PAS) or Gram stain. Patients with oral candidiasis were then treated with antifungal agents, and samples were obtained for fungal culture at bi-weekly intervals until all signs and symptoms resolved. The antifungal agents used (nystatin, clotrimazole, ketoconazole and fluconazole) varied from case to case and were sometimes used in combination.

UCSF LIBRARY

For comparison, a second group, consisting of 9 HIV-seronegative heterosexual men with no history of oral candidiasis, were sampled for culture at 6-month intervals. These persons had no signs or symptoms of oral candidiasis throughout the period of the study. For the partner portion of the study, a subset of the HIV-seropositive group, consisting of 3 pairs of sexual partners, were followed for 6 months and cultures were obtained at 1- to 3-month intervals.

Isolation of Oral Candidal Strains.

Samples were removed from both oral candidiasis lesions and nonlesional sites in the oral cavities of the subjects with a sterile Culturett (American Scientific Products, McGraw Park, IL, U.S.A.). This was then vigorously swirled in 0.5 ml sterile water, which was plated directly onto agar plates containing yeast extract, peptone, and dextrose (YEPD) (Sherman, Fink and Hicks 1986) with 30 μg/ml tetracycline and 30 μg/ml ampicillin to prevent bacterial growth. Yeast colonies were isolated by streaking for single colonies. Several colonies were then chosen at random from each plate, inoculated separately into liquid YEPD medium and grown at 30°C overnight for DNA isolation.

Species Identification

Yeasts were biotyped for species identification by conventional biochemical (API-20 system, Analytab Products, Plainview, NY, U.S.A.) and enzyme profile (Yeast Ident™, Analytab Products) methods, according to the manufacturer's instructions.

UCSF LIBRARY

DNA Analysis.

DNA was isolated by standard procedures used with *Saccharomyces cerevisiae* as described by Sherman, Fink and Hicks (1986) with the exception that in the yeast "mini-prep" procedure, three phenol extractions were required to purify the DNA for enzymatic restriction. For the DNA analysis, 1 μg of DNA sample from each of the isolates was incubated with 10 U of either *Eco*RI or *Cla*I restriction enzyme (New England BioLabs, Inc., Beverly, MA, U.S.A.) for 1 h at 37°C. The digestion products were separated by electrophoresis for 15 h in 1.0% agarose gels in TBE (0.089 *M* Tris-borate, 0.089 *M* boric acid, 0.002 *M* EDTA). The gels were stained with ethidium bromide (0.5 μg/ml in TBE) for 0.5 h, rinsed with TBE and photographed using 300-nm transillumination through an orange filter with Polaroid 667 film. These gels were then examined for comparison of restriction fragment length polymorphism (RFLP) banding patterns from the different isolates run on the same gel. After electrophoresis, the gels were soaked in 0.3 *N* NaOH for 30 min to denature the DNA, then in a solution containing 12 m*M* Tris, 6 m*M* sodium acetate and 0.3 m*M* EDTA (pH 7.5, transfer buffer) for 30 min to neutralize the gel. The DNA fragments were transferred to GeneScreen Plus hybridization membrane (New England Nuclear Corp., Boston, MA, U.S.A.). The transfer was performed in a Hoeffer electroblot transfer apparatus containing transfer buffer. Electrophoresis was performed for 1 h at 10 V and 5°C, then the voltage was then increased to 30 V for 2 h at 5°C. The DNA was then fixed by baking at 80°C for 2 h.

Strain Classification.

The *Candida* isolates obtained from oral cultures were classified by DNA "fingerprinting" with a dispersed middle repetitive DNA sequence probe from *C. albicans*. This method was first reported by Stevens and Scherer (1988) and has been used to track *Candida* isolates in other patient populations (Soll et al. 1987, Soll et al. 1988). Blots were probed with the nick-translated ^{32}P-labeled dispersed *Candida albicans* repeat sequence Ca3 (Sadhu et al. 1991). Differences in the hybridization patterns of the isolates were used to identify and track the isolates obtained from each subject. Hybridizations for Southern blots were performed as described previously (Sadhu et al. 1991). Hybridization patterns were demonstrated by autoradiography.

This method was used in preference to a previously described DNA genotyping method utilizing electrophoretic karyotyping (Merz, Connelly and Hieter 1988). Electrophoretic karyotyping of *Candida* strains can distinguish strains only upon the basis of rather gross chromosomal rearrangements. The method of fingerprinting through the use of moderately repetitive DNA probes provides the advantage of much finer resolution of differences in the genome, down to the nucleotide level. By using this method, it is possible to perform a longitudinal analysis of *C. albicans* isolates from patients during the course of symptomatic and asymptomatic infection through repeated samplings. It can also be used to compare the similarity of isolates obtained from different patients.

Criteria for Asessment of Strain Relatedness.

UCSF LIBRARY

Strain similarity was assessed on the basis of both the RFLP patterns in the ethidium bromide-stained gel and the hybridization patterns with the strain-specific Ca3 probe. In some cases, differences in strains were obvious on examination of the RFLP banding pattern alone. In other cases, in which only subtle differences in either mobility or position of minor bands were present, identification of differing strains was not as obvious. For this reason we also examined the autoradiographic patterns produced by hybridization with the Ca3 probe. The Ca3 patterns always confirmed the obvious RFLP differences in different strains and, in many cases, highlighted some less obvious differences. Therefore, in this study, two strains were judged to be identical when 1.) DNA samples run on the same gel produced the same DNA banding patterns with two restriction enzymes, *Cla* I and *Eco*RI, and 2.) Southern blots of these gels showed identical Ca3 hybridization patterns.

RESULTS

C. albicans Isolates Show a High Degree of Polymorphism.

Thirty of 33 HIV-seropositive subjects and 3 of 9 HIV-seronegative subjects were culture positive for fungi. Conventional biotyping and DNA genotyping analyses identified all of the isolates obtained as *C. albicans*, with two exceptions: one HIV-seropositive subject was found to harbor *Candida tropicalis* concurrently with *Candida albicans* (Figure 1) and one HIV-seropositive patient was found to harbor *Candida lipolytica* (see Figure 4). The strains that were isolated were genotypically distinct in almost all cases. In no case was

UCSF LIBRARY

more than one strain of *C. albicans* isolated simultaneously from the oral cavity of a single patient. Figure 1 demonstrates the individual variation seen in *C. albicans* strains isolated from different individuals. Although at first the banding patterns appear quite similar, subtle differences between patients can be discerned upon close inspection of both the *Eco*RI and *Cla* I patterns. A total of 33 different *C. albicans* strains were isolated from 33 subjects who harbored *Candida*. Of these 33 subjects, 30 were HIV-seropositive and 3 were HIV-seronegative. Twenty-one of the 30 HIV-seropositive patients, and none of the HIV-seronegative subjects, developed active oral candidiasis during the study period.

Eighteen of the 21 patients who developed oral candidiasis infections maintained a single strain of *C. albicans*, characteristic for each patient, throughout treatment, irrespective of the presence or absence of signs or symptoms of oral candidiasis. For these 18 patients, the same strain could be isolated during successive rounds of oral candidiasis, regardless of drug treatment regimen.

The remaining 3 patients exhibited more than one strain of *Candida* during the course of the study. Figure 2 shows the "fingerprinting" pattern obtained for one of these patients with recurrent oral candidiasis over a period of 15 months. During this time the patient was treated with different combinations of topical and systemic antifungal medications detailed in Table 1. As seen in the *Eco*RI pattern, the patient's infecting *C. albicans* strain at the time of entry into

the study was different from the strain obtained from the ensuing samples. However, the fact that this patient's *C. albicans* isolate remained unchanged over the course of a year attests to the stability of unique isolates in many individuals despite multiple drug treatment regimens.

Of the 12 HIV-seropositive patients who had no history of oral candidiasis on entry into the study, two subsequently developed oral candidiasis. On entry, these patients had been without lesions but were culture-positive for *Candida*. Cultures taken from the lesions at the time of diagnosis revealed the same strain of *C. albicans* as was present during the asymptomatic carrier state in each patient (data not shown).

UCSF LIBRARY

Analysis of Candida Isolates from Sexual Partners.

Three sets of HIV-seropositive sexual partners were followed in order to evaluate the potential for transmission of *C. albicans* strains through intimate contact in this group. The first set of partners maintained unique strains of *C. albicans* for 6 months and were either asymptomatic or presented with pseudomembranous lesions (data not shown). The second set of partners (2A and 2B) each harbored the same strain of *C. albicans* during the 6-month period for which they were followed. In this interval, one partner developed oral pseudomembranous and erythematous candidal lesions while his partner remained asymptomatic (Figure 3).

The third set of partners harbored different strains of *C. albicans* at the beginning of the study period. After one month, the strain of one partner (3A)

had become predominant in the oral cavity of the other partner (3B). At the same time, the oral sample from partner 3A showed no evidence of his initial *C. albicans* strain, the predominant oral fungal species being *Candida lipolytica*. Both partners then reverted to harboring the initial infecting *C. albicans* strain of partner 3A. It is interesting to note that the initial infecting strain of patient 3A was associated with pseudomembranous oral lesions in partner 3A and no lesions in 3B (Figure 4 and Table 2).

DISCUSSION

Our study demonstrates that in most cases, HIV-infected persons who harbor *Candida* species orally maintain a distinct strain, as assayed by the Ca3 probe; only rarely do two individuals have the same strain. This "signature" strain, to the exclusion of other species and strains of *Candida*, may be repeatedly isolated from the patients' oral cavity regardless of the presence or absence of oral candidal lesions or intervening antifungal treatment. This finding is in accordance with those of Soll et al. (1989) in their analysis of oral and vaginal *Candida* strains found in patients with recurrent fungal vulvovaginitis. Similar results in 60 of 63 immunocompromised patients have been reported by Fox, Mobley and Wade. (1989).

Although we did not undertake a precise quantitative assessment of the amount of *Candida* present in the oral cavity, data (not shown) from cultures taken from lesional and non-lesional sites in the same mouth suggest that lesions contained from 10 to 1000 times more organisms than the nonlesional

areas. In general, antifungal treatment that resolved the signs and symptoms of oral candidiasis correlated directly with a decreased amount of culturable *C. albicans* (data not shown).

Taken together, the data suggest that each patient carries a unique strain of *Candida* that is responsible for recurrent oral candidiasis in that patient. In addition, finding the same strain during both symptomatic and asymptomatic states bears on the longstanding question as to what differentiates a "non-pathogenic" from a "pathogenic" strain of *Candida*. In two subjects, the same strain was isolated both before and after the subject's initial episode of oral candidiasis. In addition, a study by Whelan et al. (1990) comparing drug resistance patterns, amino acid and nucleotide synthesis, sugar use and enzyme activities of *Candida* isolates from HIV-infected patients with those of isolates from healthy adults, found no significant differences in these variables between the two groups. It thus seems clear that the development of oral candidiasis in the immunocompromised host is not dependent upon replacement of an avirulent strain by a more virulent strain. Rather, whether or not a *Candida* strain is able to cause disease may depend upon the regulation of expression of specific genes within its gene repertoire. Although it was beyond the scope of this study, it would be interesting to know whether the HIV-infected patients harbored the same strain of *Candida* before and after the development of their immunosuppression.

USF LIBRARY

Since finding two individuals who carry the same strain of *Candida* as assessed by RFLP analysis is a relatively infrequent event, the finding of strain relatedness among sexual partners was unexpected. In the study by Fox and colleagues. (Fox, Mobley and Wade 1989), 60 of 63 patients were judged to harbor unique strains of *C. albicans*. Even in those instances in which identical *Candida* genotypic patterns were found in different individuals, no epidemiologic connection could be established. Similar findings have been reported by Powderly et al. (1992). Our finding of identical *C. albicans* strains in partners 2A and 2B could be ascribed to a chance event in which these two individuals had always harbored the same strain, and would be in accord with the findings described by Fox. (Fox, Mobley and Wade 1989). Another possibility is strain transmittal from one subject to another subject. In the case of partners 2A and 2B, the answer is unclear since we have no information on the genotypes of strains previously harbored by the two subjects. However, the situation of partners 3A and 3B is clearly different and excludes the first possibility (Figure 4). In this case, partner B's strain was replaced by that of partner A during the one-month period following collection of the first sample from both partners. During that same time period, partner A's original *C. albicans* strain became undetectable by our sampling methods and *C. lipolytica* became the predominant oral fungal species. The final oral sample was obtained from partner 3A four months after the first sample and showed that the subject's original *C. albicans* strain had again become predominant. These events took place over a 3- to 6-

month period and directly suggest the interesting possibility of transmission via intimate contact.

Indirect evidence for both nosocomial transmission and transmission via intimate contact has been gathered by a number of studies. Schmid and colleagues (Schmid, Voss and Soll 1990), using a computer-assisted method of genotyping, found genotypic similarity (but not identity) in isolates from seven immunocompromised patients hospitalized over a 2.5-month period in the same hospital, suggestive of a nosocomial route of transmission. In the same study a husband and wife who presented with simultaneous oral lesions were found to harbor identical *Candida* strains, providing evidence for oral transmission via intimate contact, although in this case, transmission through casual household exposure could not be ruled out. Stevens, Odds and Scherer, (1990) found two cases in which isolates in samples taken from the heterosexual partners of women with vaginal candidiasis matched those found in the vaginal samples. Ours is the first study to document, through longitudinal molecular analysis of strains in partners, the actual transmission of a strain from one patient to another.

Also of note is the observation that the clinical oral status of partners 3A and 3B varied directly with the identity of the infecting strain. Partner B presented with erythematous and pseudomembranous candidiasis while harboring his own strain, but was asymptomatic after it was replaced by partner A's original strain. Partner A, however, developed erythematous and

pseudomembranous candidiasis in the presence of the same strain. No conclusions can be drawn regarding the virulence of partner A's second candidal species, as the patient was taking clotrimazole at the time it was isolated. However, *C. lipolytica* is not usually considered a very virulent species. In fact, Holzschu et al. (1979) observed that this species was unable to cause disease even when given intraveneously to cortisone-treated mice. The variation in clinical presentation in different persons upon infection by the same strain highlights the importance of the host response in the pathogenesis of oral candidiasis.

These data also suggest another interesting possibility: i.e., that there is a hierarchy among *Candida* strains such that between otherwise virulent strains, only a single strain predominates in a given lesion or body site. DNA fingerprinting has been used to determine whether different strains predominate in different parts of the body. In most cases, the same strain is present in the oral and esophageal cavities (Fox, Mobley and Wade 1989, Whelan et al. 1990). However, isolates obtained from different parts of the body of the same person (anus, rectum, genitals, breasts, vagina etc.) can be quite different (Scherer and Stevens 1988, Soll et al., 1989 and Schmid, Voss and Soll 1990). It will important to determine whether different virulence factors make it possible for the yeast to survive and proliferate in different locations.

In conclusion, our findings indicate the continued presence of the same strain of *C. albicans* within an individual during asymptomatic infection and after

antifungal treatment in the majority of cases. However, our data also support strongly the possibility of transmission of *C. albicans* via either casual or intimate contact in the setting of the immunocompromised HIV-infected patient. The ability of the organism to transfer to and successfully replace the original host flora, and the variety of genotypes exhibited by organisms isolated from different individuals, may reflect the organism's ability to respond to selection pressures from the host environment. If this is true, then it is not surprising that *C. albicans* is such a successful opportunistic pathogen.

Finally, the use of the Ca3 probe is still in its earliest stages. Our data and those of others have been primarily obtained with clinical isolates from episodic candidiasis, as in oral or vaginal candidiasis, or in hospital settings with immunocompromised patients. Because these studies have been limited, we have no information regarding the spectrum or stability of *Candida* strains found in association with intravenous drug use, mother-child transmission, or heterosexual transmission in these important HIV-infected populations. In addition, information is lacking regarding the virulence factors and transmission of *Candida* in pharmaceutically induced immunosuppression. It is clear, however, that the use of probes such as Ca3 for molecular fingerprinting provides an indispensable and powerful tool for following the etiology of *Candida* infections in both epidemiological and clinical studies.

SUPPORT

Supported in part by a grant from the AIDS Clinical Research Center, by U.S. Public Health Service grant NIH PO1 DE07946, and by the California Universitywide AIDS Research Program.

ACKNOWLEDGEMENT

The authors thank Ms. E. Leash for editing this manuscript. This work was supported in part by a grant from the AIDS Clinical Research Center, by U.S. Public Health Service grant NIH PO1 DE07946 and by the California Universitywide AIDS Research Program.

UCSF LIBRARY

REFERENCES - CHAPTER 2

1. Carpenter, C. C. J., K. H. Mayer, A. Fisher, M. B. Desai, L. Durand. 1989. Natural history of acquired immunodeficiency syndrome in women in Rhode Island. *Am. J. Med.* 86:771-5.
2. Feigal, D. W., M. H. Katz, Greenspan D, J. Westenhouse, W. Winkelstein Jr. W. Lang, M. Samuel, S. P. Buchbinder, N. A. Hessol, A. R. Lifson, G. W. Rutherford, A. Moss, D. Osmond, S. Shiboski, J. S. Greenspan. 1991. The prevalence of oral lesions in HIV-infected homosexual and bisexual men: Three San Francisco epidemiologic cohorts. *AIDS* 5:519-25.
3. Fox, B. C., H. L. T. Mobley, and J. C. Wade. 1989. The use of a DNA probe for epidemiological studies of candidiasis in immunocompromised hosts. *J. Infect. Dis.* 159:488-94.
4. Greenspan, D., J. S. Greenspan, M. Schiødt, J. J. Pindborg. 1990. *AIDS and the mouth.* Copenhagen: Munksgaard.
5. Holzschu, D.L., F. W. Chandler, L. Ajello, D. G. Ahearn. 1979. Evaluation of industrial yeasts for pathogenicity. *Sabouraudia* 17:71-8.
6. Martin, M. V., Wilkinson, G. R. 1983. The oral yeast flora of 10-year-old schoolchildren. *Sabouraudia* 21:129-35.
7. Merz, W.G., C. Connelly, and P. Hieter. Variation of electrophoretic karyotypes among clinical isolates of *Candida albicans*. *J. Clin. Microbiol.* 26:842-5.

8. Powderly, W.G., K. Robinson, and E. J. Keath. 1992. Molecular typing of *Candida albicans* isolated from oral lesions of HIV-infected individuals. *AIDS* 6:81-4.

9. Sadhu, C., M. J. McEachern, E. P. Rustchenko-Bulgac, J. Schmid, D. R. Soll, J. B. Hicks. 1991. Telomeric and dispersed repeat sequences in *Candida* yeasts and their use in strain identification. *J. Bacteriol.* 173:842-50.

10. Samaranayake, L. P., and P. Holmstrup. 1989. Oral candidiasis and human immunodeficiency virus infection. *J. Oral Pathol. Med.* 18:554-64.

11. Scherer, S., and D. A. Stevens. 1988. A *Candida albicans* dispersed, repeated gene family and its epidemiologic applications. *Proc. Natl. Acad. Sci. U. S. A.* 85:1452-6.

12. Schmid, J., E. Voss, and D. R. Soll. 1990. Computer assisted methods for assessing strain relatedness in *Candida albicans* by fingerprinting with the moderately repetitive sequence Ca3. *J. Clin. Microbiol.* 28:1236-43.

13. Sherman, F., G. R. Fink, and J. B. Hicks. 1986. *Methods in yeast genetics: a laboratory course manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

14. Soll, D. R., C. J. Langtimm, J. McDowell, J. Hicks, R. Galask. 1987. High-frequency switching in *Candida* strains isolated from vaginitis patients. *J. Clin. Microbiol.* 25:1611-22.

UCSF LIBRARY

15. Soll, D. R., M. Staebell, C. Langtimm, M. Pfaller, J. Hicks, T. V. Gopala Rao. 1988. Multiple *Candida* strains in the course of a single systemic infection. *J. Clin. Microbiol.* 26:1448-59.

16. Soll, D. R., R. Galask, S. Isley, T. V. Gopala Rao, D. Stone, J. Hicks, J. Schmid, K. Mac, C. Hanna. 1989. Switching of *Candida albicans* during successive episodes of recurrent vaginitis. *J. Clin. Microbiol.* 27:681-90.

17. Stevens, D.A., F. C. Odds, and S. Scherer. 1990. Application of DNA typing methods to *Candida albicans* epidemiology and correlations with phenotype. *Rev. Infect. Dis.* 12:258-66.

18. Whelan, W.L., D. R. Kirsch, K. J. Kwon-Chung, S. M. Wahl, P. D. Smith. *Candida albicans* in patients with the acquired immunodeficiency syndrome: Absence of a novel or hypervirulent strain. *J. Infec.t Dis.* 162:513-18.

19. Wright, P. S., P. Clark, and J. M. Hardie. 1985. The prevalence and significance of yeasts in persons wearing complete dentures with soft-lining materials. *J. Dent. Res.* 64:122-5.

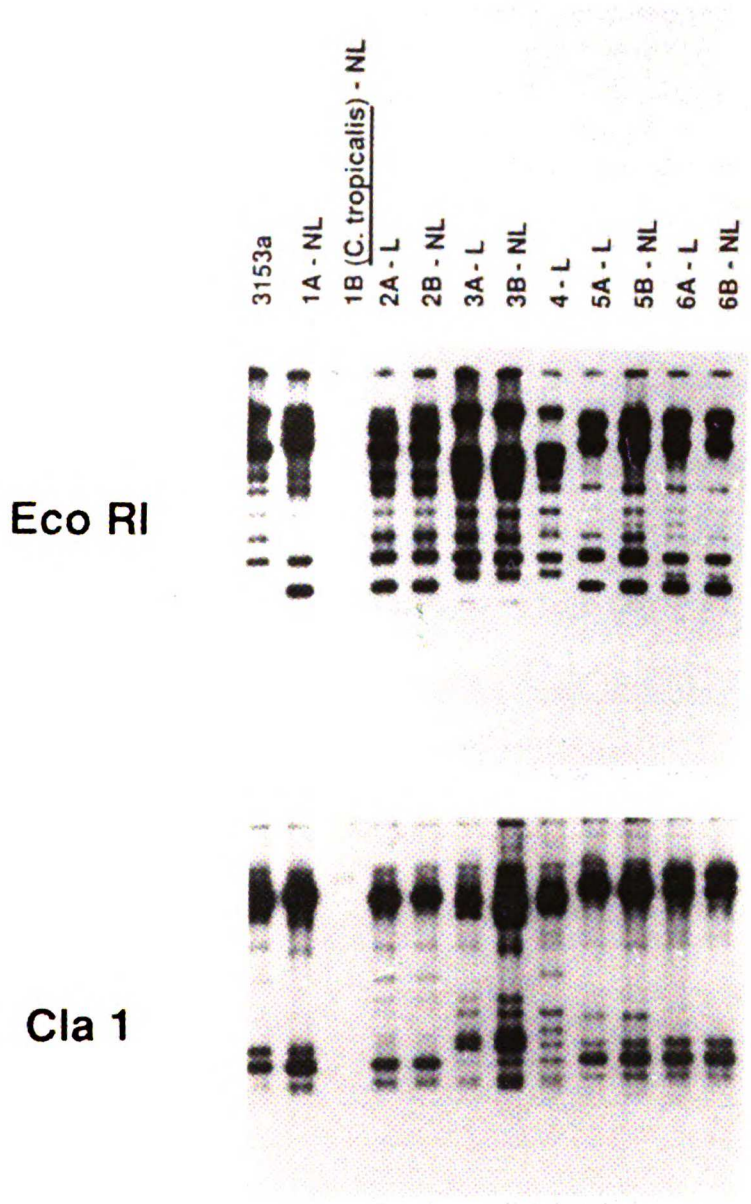


Figure 1

UCSF LIBRARY

Figure 1 Legend

Southern blot showing a representative sampling of genomic DNA from *C. albicans* strains cultured from six patients and hybridized with the ^{32}P-labeled species-specific Ca3 probe. DNA was isolated and digested with *Eco*RI and *Cla* I in separate experiments. Designations L and NL for strains of the same sample number indicate that the swabs were taken from lesions (L) or non-lesional (NL) areas of the mouth. Letters A and B, when they appear, indicate colonies of different morphology picked from the same primary culture. Lane marked 3153a is DNA from a standard laboratory control strain. Note that the patterns are more or less closely related. Strain 1B is a strain of the related yeast, *Candida tropicalis*. The Ca3 probe does not hybridize to *C. tropicalis* DNA.

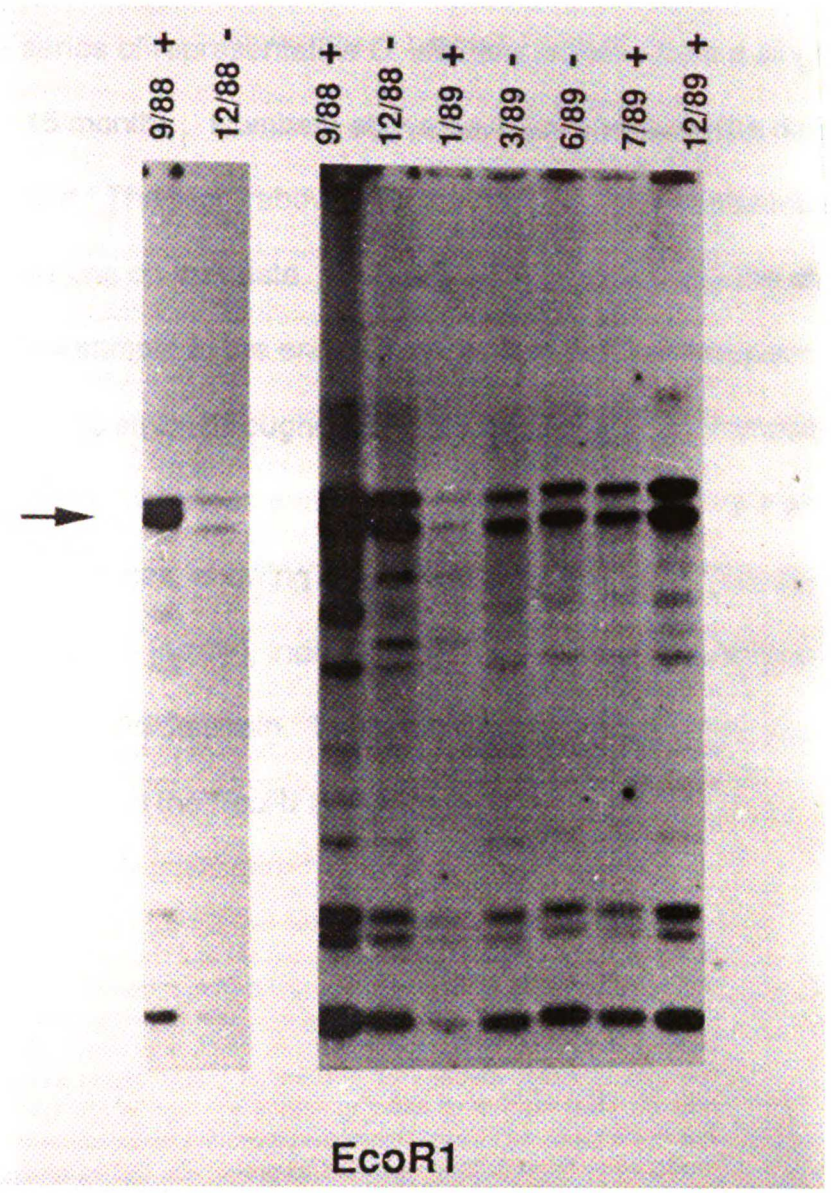


Figure 2

UCSF LIBRARY

Figure 2 Legend

Southern blot of genomic DNA hybridized with the ^{32}P-labeled Ca3 probe. DNA is from a series of representative *C. albicans* isolates from a single patient over a period of 15 months . Numbers above each lane indicate the date the sample was obtained. The + or - above indicate the presence or absence of oral candidal lesions on that date. This blot demonstrates both the change in strain from the first sample to the ensuing samples and the subsequent maintenance of a single unique strain throughout symptomatic and asymptomatic episodes and in the presence of various antifungal drugs. The left panel is a shorter exposure of the first two lanes, showing more clearly the change in mobility of the second band from the top (arrow), indicative of a change in the genotype of the predominant *Candida* strain. The patient's *Cla* I pattern (not shown) reflected the results seen in the *Eco*RI Southern blot.

Table 1. Recurrent oral candidiasis despite antifungal treatment in one HIV-seropositive subject

Date	Current antifungal Medication	Oral Condition	*Candida* culture*
9/88	Ketoconazole	Peudomembranous and erythematous candidiasis	++
12/88	Nystatin	No oral lesions	+
1/89	Ketoconazole, Nystatin	Pseudomembranous candidiasis	+++
3/89	Nystatin	No oral lesions	+++
6/89	Ketoconazole, Nystatin	No oral lesions	+++
7/89	None	Pseudomembranous candidiasis	+++
12/89	None	Pseudomembranous candidiasis	+++

*Cultures were scored as 0 (less than 2 colonies); + (2 to 10 colonies); ++ (11 to 100 colonies) and +++ (greater than 100 colonies) in the primary culture plate.

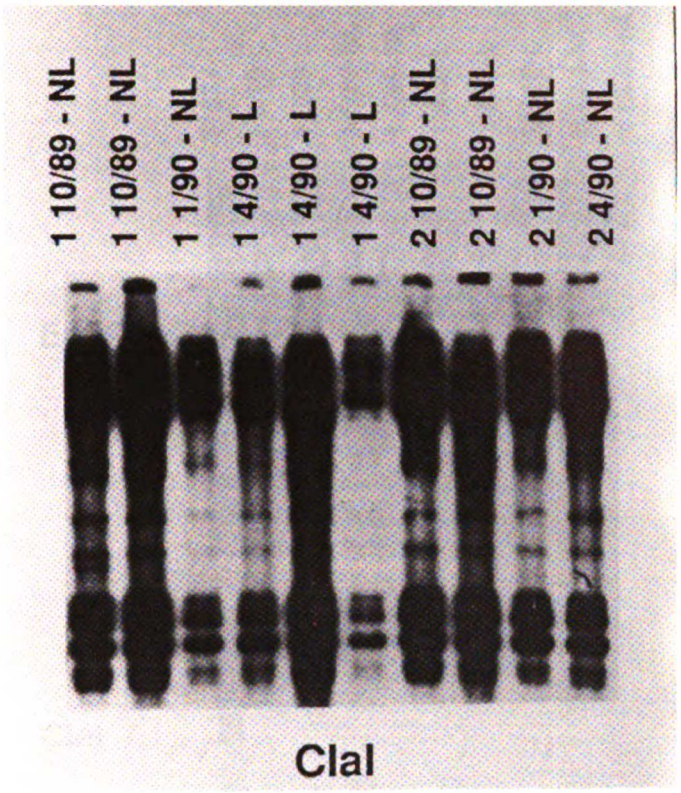
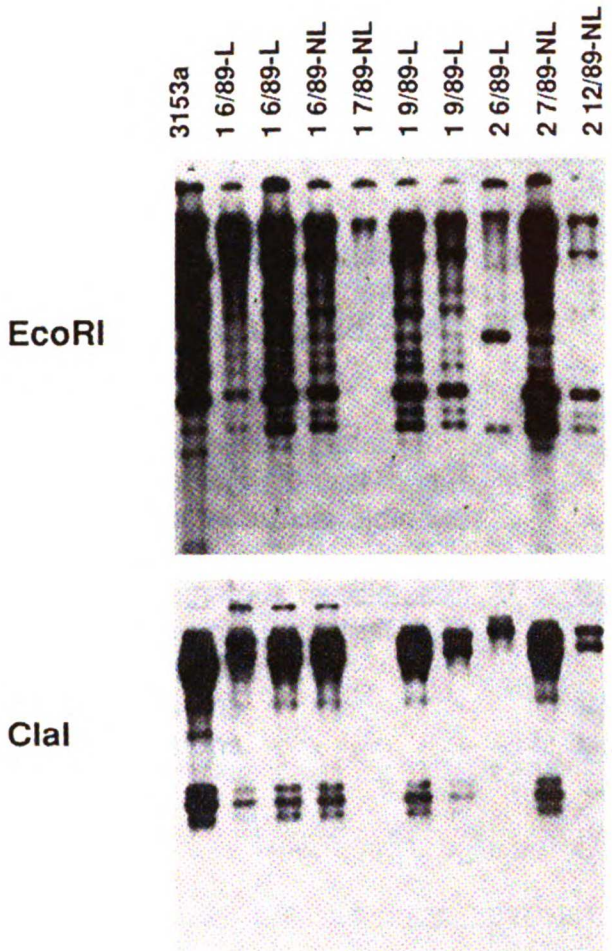


Figure 3

^{32}P-labeled Ca3 probe hybridization with a Southern blot of genomic DNA from isolates obtained from two sexual partners (partners 2A and 2B from the text) who maintained an identical strain throughout the 6 months of follow-up. In this *Cla* I blot, isolates from the two partners are differentiated by the numbers 1 and 2 above the lanes, along with the dates of the samples. In some instances, multiple samples representing different areas of the mouth were obtained. Letters L and NL stand for lesional and non-lesional areas of the mouth, respectively. The *Eco*RI blot (not shown) confirmed the results of the *Cla* I data.

UCSF LIBRARY



UCSF LIBRARY

Figure 4

Ca3 hybridization with DNA from isolates obtained from two sexual partners (partners 3A and 3B from the text) in which one partner's (3B) strain was replaced by a strain that gave a pattern identical to that of his partner's (3A) strain over 5 months of follow-up. In this *Eco*RI blot, the lane designated 3153a represents control DNA from a standard laboratory strain. Isolates from the two partners are differentiated by the numbers 1 for partner 3A and 2 for partner 3B above the lanes, along with the dates of the samples. In some instances, multiple samples representing different areas of the mouth were obtained. Letters L and NL stand for lesional and non-lesional areas of the mouth respectively. Isolate designated 1 7/89-NL was determined to be *Candida lipolytica*. The Ca3 probe does not hybridize to the *C. lipolytica* DNA

Table 2. Presumed candidal transmission between sexual partners.

Partner	Date	Clinical status	Current medication	Strain
3A	6/89	Pseudomembranous candidiasis, floor of mouth; erythematous candidiasis, palate	None	A-1
	7/89	No lesions	Clotrimazole	A-2*
	9/89	Pseudomembranous candidiasis, ventral tongue; erythematous candidiasis, palate	None	A-1
3B	6/89	Erythematous candidiasis, tongue; angular cheilitis, oral commissures	Clotrimazole	B-1
	7/89	No lesions	None	A-1
	12/89	No lesions	None	A-1

*A-2 = *Candida lipolytica*

UCSF LIBRARY

CHAPTER 3

Candidal Virulence Factors and Pathogenesis

In the study reported in Chapter 2, I had found that a patient most typically carried the same strain of *Candida* throughout the course of their HIV-infection, and that this same strain of *Candida was* found during both active disease states with oral lesions and inactive or carrier states. From this I concluded that it was not that new and more virulent strains of *Candida* were becoming more prevalent among this group of HIV-infected patients, but that most strains of *Candida* was likely capable of becoming pathogenic in the face of the developing host immunosuppression. As such, we can surmise that there are likely factors in *C. albicans* itself, which may be triggered to assist in the establishment of an active infection of the host, as opposed to just colonization.

This sparked my interest in the roles of various fungal virulence factors in pathogenesis. Potential virulence factors in *Candida* are many and are listed in Table 1; among these are thigmotropism, adhesins, tissue lytic enzymes, and the ability to express multiple switch phenotypes. Although any or all of these may assume a role in the development of pathogenesis in humans, I chose to investigate one of the most extensively studied putative virulence factors in *Candida,* a secreted aspartyl proteinase (Staib 1965; Hube 1991) and its relationship with another putative virulence factor, the ability of *C. albicans* to undergo rapid phenotypic switching.

UCSF LIBRARY

In addition to the familiar yeast/hyphae morphological transitions, each strain of *C. albicans* is capable of expressing a number of alternative switch phenotypes under appropriate growth conditions. These are manifest as alterations in cell morphology as well as colony shape and opacity. Each specific cell phenotype is correlated with a unique colony morphology on agar plates, and switching between alternative phenotypes occurs at low and predictable frequencies; from 10^{-2} to 10^{-4} as reported by Slutsky et al. (1985).

My interest in hydrolytic enzymes and phenotypic variation stems by extension from other parasites such as, *Trypanosoma brucei* (Cross 1996) and *Schistosoma mansonii* (Newport 1988). Studies of these parasites have revealed traits such as the ability to undergo antigenic variation and to secrete hydrolytic enzymes (respectively) to be potentiators of pathogenesis. These traits are thought to play a role in enabling the organism to avoid the host immune response or to invade and penetrate the host defenses, respectively. It is possible that morphologic variation of *C. albicans* at both the level of the blastospore/hyphae transition and other colony and cell variants could result in evasion or confusion of the host immune response. Of more general interest, phenotypic switching appears to be correlated with the selective transcription and expression of a number of gene products, some of which may be related to virulence. As an example:

The WO1 or White-Opaque *C. albicans* strain was originally isolated from the blood and lungs of a bone marrow transplant patient at the University of Iowa

Hospitals (Slutsky 1987.) This strain is known to switch spontaneously at a high frequency (between 10^{-2} and 10^{-4} per cell cycle) between white and opaque cell phenotypes. This strain has been investigated extensively in the laboratory of Dr. David Soll. It is from this work (Anderson, Mikhalik and Soll 1990) that we know that the white cells are oval to round in shape and bud with the same pattern as that of the standard laboratory strains. In contrast, the opaque cells are elongated and bean shaped, exhibit approximately twice the mass and volume of white cells and can sometimes bud in a bipolar fashion. Additional differences, as described by Kennedy et al. (1988) and Kolotila and Diamond (1990), between the two phenotypes are listed in Table 2, and include surface antigens, permeability, adhesiveness and drug resistance. Not listed are a number of differences in both the type and amounts of secreted proteinases which are produced, however this will be discussed at length in the following chapters.

In 1965, Staib (Staib 1965) first reported the identification of a pH-dependent proteolytic activity in *C. albicans* strains. Using an auxanographic method, he showed that some strains of *C. albicans* were capable of using human albumin as a nitrogen source. Following this report, Remold et al. (ref. Remold 1968) isolated a 40 kD secreted proteinase. A sequence of a *C. albicans* gene encoding a secretory aspartate proteinase was then published by Hube (Hube et al. 1991.) Since that time, a number of laboratories, including

ours, have identified a family of secreted aspartyl proteinases. The second secreted aspartyl proteinase or *SAP*2 was identified and cloned in 1992 by Wright (Wright et al. 1992). The discovery of a third Sap gene (*SAP*3) and elucidation of the pattern of its expression with relationship to that of *SAP*s 1 and 2 in the phenotypic switching strain WO1 are the subject of the remainder of the next chapter.

REFERENCES - CHAPTER 3

1. Anderson, J., R. Mikhalik, and D. R. Soll. 1990. Ultrastructure and antigenicity of the unique cell wall "pimple" of the *Candida* opaque phenotype. *J. Bacteriol.* 172:224-235.

2. Cross, G. 1996. Antigenic variation in trypanosomes: Secrets surface slowly. *Bioessays* 18:283-91.

3. Ghannoum, M. A., I. Swairjo, and D. R. Soll. 1990. Variation in lipid and sterol contents in *Candida albicans* white and opaque phenotypes. *J. Med. Vet. Mycol.* 28:103-15.

4. Hube, B., C. J. Turver, F. C. Odds, H. Eiffert, G. J. Boulnois, H. Kochel, and R. Rüchel. 1991. Sequence of the *Candida albicans* gene encoding the secretory aspartate proteinase. *J. Med. Vet. Mycol.* 29:129-32.

5. Kennedy, M. J., A. L. Rogers, L. R. Hanselman, D. R. Soll, R. J. Yancye. 1988. Variation in adhesion and cell surface hydrophobicity in *Candida albicans* white and opaque phenotypes. *Mycopathologia.* 102:149-56.

6. Kolotila, M. P., and R. D. Diamond. 1990. Effects of neutrophils and in vitro oxidants on survival and phenotypic switching of *Candida albicans* WO-1. *Infect. Immun.* 58:1174-79.

7. Newport, G., J. H. McKerrow, R. Hedstrom, M. Petitt, L. McGarrigle, P.J. Barr and N. Agabian. 1988. Cloning of the protease that facilitates infection by schistosome parasites. *J. Biol. Chem.* 263:13179-84.

8. Remold, H., H. Fasold, and F. Staib. 1968. Purification and characterisation of a proteolytic enzyme from *Candida albicans*. *Biochim. Biophys. Acta.* 167:399-406.

9. Slutsky B., J. Buffo, and D. R. Soll. 1985. High frequency switching of colony morphology in *Candida albicans*. *Science* 230:666-69.

10. Slutsky, B., M. Staebell, M. Pfaller, D. R. Soll. 1987. The "White-opaque transition": a second switching system in *Candida albicans*. *J. Bacteriol.* 169: 189-97.

11. Staib, F. 1965. Serum-proteins as nitrogen source for yeastlike fungi. *Sabouradia* 4:187-93.

12. Wright, R. J., A. Carne, A. D. Hieber, I. L. Lamont, G. W. Emerson, P. A. Sullivan. 1992. A second gene for a secreted aspartate proteinase in *Candida albicans*. *J. Bacteriol.* 174:7848-53.

Table 1. Putative Virulence Factors of *Candida albicans*

Immunomodulators (cell wall components (especially mannans)
thigmotropism
phenotypic switching
adhesins
few nutrient requirements
secreted hydrolytic enzymes
rapid growth rate
drug resistance
cell wall (resistance to dehydration/proteolytic enzymes)

Table 2. Differences between White and Opaque Cells

White cells	**Opaque cells**	**References**
round, budding cells similar to standard yeast cells	elongate, bean shaped, with 3X the volume of W cells	Slutsky, 1987
⇑ lipid and sterol content	⇓ lipid and sterol content	Ghanoum, 1990
⇓ resistance to antifungals	⇑ resistance to antifungals	Ghanoum, 1990
⇑ adhesion to buccal epithelial cells	⇓adhesion to buccal epithelial cells	Kennedy, 1988
⇑ resistance to candidacidal activity of neutrophils and oxidants	⇓resistance to candidacidal activity of neutrophils and oxidants	Kolotila, 1990
- surface pimple antigen	+ surface pimple antigen	Anderson, 1990

CHAPTER 4

Three Distinct Secreted Aspartyl Proteinases in *Candida albicans*.

Portions of the following chapter have been previously published in:

Journal of Bacteriology

October 1993

Volume 175, No.19

pp. 6126-6133

Theodore C. White, Shelley H. Miyasaki & Nina Agabian

Intercampus Program in Molecular Parasitology, University of California - San Francisco and Berkeley, and Department of Stomatology and the Oral AIDS Center, University of California, San Francisco. San Francisco, California 94143-1204.

DUKE LIBRARY

ABSTRACT

The secreted aspartyl proteinases of *Candida albicans* (products of the *SAP* genes) are thought to contribute to virulence through their effects on *Candida* adherence, invasion and pathogenicity. From a single strain of *C. albicans* (WO-1) which expresses a phenotypic switching system, three secreted aspartyl proteinases have been identified as determined by molecular weight and N-terminal sequence. Each of the three identified proteins represents the mature form of one of three distinct proteinase isoenzymes, two of which correspond to the recently cloned *SAP1* and *SAP2* genes (previously referred to as *CAP, PEP* or *PRA*). A genomic library was screened under low stringency hybridization conditions with a PCR fragment from *SAP1*. In addition to clones of *SAP1* and *SAP2*, a clone containing *SAP3*, a novel third secreted proteinase gene, was identified and sequenced. The three aspartyl proteinase isoenzymes differ in primary sequence and pI, suggesting that they may play different roles in virulence and pathogenesis. All three of these proteinases are expressed in the same strain. However, the pattern of proteinase expression is correlated with the switch phenotype of the cell. Opaque cells of strain WO-1 express Sap1 and Sap3 while white cells of the same strain express Sap2. The differential expression of three Sap proteinases may contribute to virulence in *Candida albicans*.

INTRODUCTION

Candida albicans and related species exist as commensal oral and vaginal flora in many, if not most, healthy individuals. Under certain conditions usually associated with a suppressed host immune system, *Candida* develops from a commensal to a pathogenic organism, causing oral, vaginal, and/or systemic candidiasis. The mechanism by which *Candida* becomes pathogenic has not been elucidated, although several factors have been characterized which contribute to the development of disease (reviewed by Cutler 1991).

Candida exists in at least two distinct morphological forms, spherical yeast forms which divide by budding and long hyphal forms (reviewed by Odds 1988). This morphological transition is considered a major factor in *Candida's* ability to avoid immune destruction. More recently, the ability of *Candida* to undergo phenotypic switching, expressed as colony morphologies on agar plates has become recognized as an important adaptive change in this organism (reviewed by Soll 1992). Switching between several different colony morphologies can occur at relatively high frequencies (10^{-2} to 10^{-4} per cell cycle). Several different switch systems have been described in *Candida albicans*, all of which may contribute to the organism's pathogenicity. Switch phenotypes have also been described in other pathogenic *Candida* spp. The switching system that is discussed in this paper is displayed in the WO-1 strain of *Candida albicans*, where switching results in two distinct cell morphologies, white (W) and opaque (O). White cells appear as normal round yeast forms,

UCLA

appear smooth in scanning electron microscopy (SEM), and form white colonies on agar. Opaque (O) cells are larger and oval in shape, contain large cytoplasmic vacuoles, possess a rough and punctate surface in SEM and grow as opaque colonies on agar (Anderson and Soll 1987, Slutsky et al. 1987).

Another characteristic of *Candida* which may contribute to its virulence is the secretion of aspartyl proteinases, first identified by Staib (1965). Originally designated as *CAP* (*Candida* Aspartyl Proteinase - a nomenclature which includes the genus name, *Candida*), this group of proteinases have been given a variety of labels. A recent publication (Morrow, Srikantha and Soll 1992) referred to the first cloned gene as *PEP1* because of its similarity to pepsinogen; however, the *Candida* proteinase resembles many different aspartyl proteinases and is not homologous to pepsinogen. The publication which reported the cloning of the second gene referred to both secreted aspartyl proteinases as *PRA* genes, because of their similarity to the PrA gene of *Saccharomyces cerevisiae* (Rüchel et al. 1992, Wright 1992). However, the PrA gene product is a vacuolar protein which is not secreted. While the secreted aspartyl proteinases of *Candida* are similar to PrA in sequence, they are not homologues. A related *Candida* aspartyl proteinase which is not secreted (Lott et al. 1989) is more closely related to the PrA proteinase. Our recent results (White and Agabian, unpublished data) have identified a *Saccharomyces* homologue of the *Candida* secreted aspartyl proteinase that is not PrA. For these reasons, we suggest that the *Candida* Secreted/Staib Aspartyl Proteinase be referred to as

SAP which is similar to the original designation, CAP. Accordingly, we shall refer to the previously cloned genes as *SAP1* (formerly *CAP*, *PEP1* and *PRA10,* Hube 1991) and *SAP2* (formerly *PRA11,* Wright 1992). (The *SAP* nomenclature was agreed upon at the 1993 ASM Meeting on Candida and Candidiasis.)

Candida proteinases have been biochemically characterized in a number of laboratories (reviewed by Douglas 1986 and Rüchel et al. 1992). The aspartyl proteinases from a variety of strains range in molecular weight between 40 and 45 kD, have a pI between 4 and 4.5, and some have been reported to be glycosylated. Characterization of the proteinases on the basis of substrate specificity, including bovine serum albumin, keratins, secretory IgA and extracellular matrix proteins, has yielded conflicting results. Recently, a genetic basis for this apparent variability has been suggested by the cloning of two genes, *SAP1* and *SAP2*, both encoding related secreted aspartyl proteinases (Hube et al. 1991 and Wright 1992).

The DNA sequence of the two *SAP* genes indicates that each proteinase is first synthesized as a precursor with an N-terminal extension (approximately 50 amino acids). The first half of the N-terminal extension encodes a hydrophobic signal sequence which by analogy to other systems targets the protein to the endoplasmic reticulum. The second half of the N-terminal extension probably encodes a "pro" peptide form of the proteinase. Both peptides appear to be removed from the mature protein by specific proteinase(s)

UCI LIBRARY

which cleaves after Lys-Arg residues, much like the *Kex2* cleavage of the a mating factor in *Saccharomyces cerevisiae* (Julius et al. 1984). However, the order of processing and the consequent effects on proteinase activity have not been determined. The mature forms of Sap1 and Sap2 are approximately 72% homologous and the predicted size of the proteins is very similar (36,249 Da and 36,341 Da, respectively).

An antiserum against isolated proteinase from W cells was prepared to examine the relationship between proteinase expression and proteinase activity during phenotypic switching. The antiserum recognizes three different protein bands on Western blots of culture supernatants from strain WO-1. Biochemical characterization of these different forms of proteinase has determined that each type represents a different isoenzyme of *Candida albicans* secreted proteinase, all secreted from a clonal strain (WO-1) of *Candida*. Cloning and sequencing of the previously unidentified third proteinase gene, which we call *SAP3*, demonstrates that each proteinase is the product of a different genetic locus. In this paper, criteria for differentiating between the three different proteinases are described which can be used to determine the presumptive role of each of these proteins in *Candida* virulence.

MATERIALS AND METHODS

Strains

The common laboratory strain, 3153A (ATCC # 28367), and strain WO-1, which switches colony morphologies (Slutsky et al. 1987), were generously

supplied by Dr. David Soll (U. of Iowa). Strain SS, which was obtained from Dr. Remo Morelli (San Francisco State), is a laboratory strain which forms hyphae at high frequency under standard conditions and which was originally isolated from a patient with systemic candidiasis (Blagoyevich 1979).

Growth Conditions and Media

Strains were routinely grown in YEPD (10 g yeast extract, 20 g peptone and 20 g dextrose per liter, autoclaved) and in YCB-BSA (23.4 g of yeast carbon base, 2 g yeast extract and 4 g bovine serum albumin per liter, pH to 5.0 and filter sterilized). YCB-BSA induces proteinase expression in most strains while YEPD does not. Cell cultures were inoculated using a single colony grown on a YEPD agar plate. Cultures were grown for 48 h to late stationary phase. The cells were pelleted using a table top centrifuge and the culture supernatants were stored at -20° C. White and opaque cells were maintained on YEPD agar plates containing phyloxine B (5 mg/mL).

Preparation of antisera

Culture supernatants from late stationary phase cultures (72 h for this preparation) of W cells grown in YCB-BSA were precipitated with ammonium sulfate (75%), centrifuged (10 min at 10,000 g), and dialyzed in Tris/EDTA/EGTA buffer (250 mM Tris HCl (pH 6.8), 10 mM EDTA, 5 mM EGTA containing 10 mg/mL phenylmethylsulfonyl floride, 13.3 mg/mL N-ethylmaleimide and 6.67 mg/mL diethylene triaminepentaacetic acid). The concentrated culture supernatants were electrophoresed through an SDS-polyacrylamide gel and the

region of the gel corresponding to 35 to 45 kD was excised. The excised gel slice was shredded and incubated in buffer (Tris/EDTA/EGTA containing 0.1% TWEEN 20) overnight at 4° C. The slurry was then loaded into an empty disposable column with a sintered glass disc at the bottom and the column was spun at 3000 g for 15 min. The eluate was lyophilized, suspended in water, dialyzed against Tris/EDTA/EGTA buffer and stored at -20° C. A white female New Zealand rabbit was inoculated subcutaneously with a 1:1 dilution of antigen in MF 59-400, an adjuvant developed and kindly provided by Chiron Corp. After four monthly injections, the rabbit was exsanguinated and the serum was stored at -20° C.

SDS-PAGE and Western Blots

Culture supernatants were mixed (1:1) with SDS gel loading buffer (8) containing 100 mM Dithiothreitol, boiled for 2 min and loaded on a 10% SDS polyacrylamide gel. Coomassie Blue staining was performed as described by Harlow and Lane (1988). For Western Blots, SDS-PAGE gels were electroblotted to nitrocellulose using 50 mM Tris HCl, 380 mM glycine, 0.1% SDS and 20% methanol in a Biorad Trans-blot Cell (Biorad, Hercules, Ca.) run overnight at 0.2 Amps at approximately 10° C. In all Western blots presented, the membranes were treated with periodate before reaction with the primary antibody. Periodate removes carbohydrate epitopes on the proteins immobilized to the nitrocellulose (Woodward, Young and Bloodgood 1985), eliminating high backgrounds caused by antibody response to the *Candida* mannans. Briefly, the

filters were washed in 50 mM sodium acetate (pH 4.5), incubated in 20 mM sodium m-periodate, 50 mM sodium acetate (pH 4.5) in the dark at room temperature, and then incubated in 50 mM sodium borohydride in PBS for 30 min at room temperature. The filters were then rinsed in TBST (Tris buffered saline with TWEEN = 10 mM Tris HCl (pH 8.0), 150 mM NaCl, and 0.05% TWEEN 20), blotted in 1% dried milk in TBST, incubated with the primary antibody (1/5000 dilution in TBST) for one h, washed three times in TBST, incubated for one h with the secondary antibody (1/5000 dilution of horseradish peroxidase (HRP) conjugated goat anti-rabbit IgG antibody (Zymed, So. San Francisco) in TBST) and washed three times. The HRP was detected with the ECL Western blotting detection reagents (Amersham, Arlington Hts., Il.) and the filters were exposed to X ray film for 10 s to 10 min depending on the intensity of the signal.

Preparative isoelectric focusing

The pI of the three proteinases was determined using the Rotofor Cell (Biorad, Hercules, Ca.) preparative isoelectric focusing apparatus. The Rotofor cell was assembled as described by the manufacturer using 2% ampholytes (Biolytes, pH 3-10, Biorad) and 1 mL of a mixture of W and O cell culture supernatants. The cell was run at 12 W for 6 h until the voltage had stabilized (usually 900 volts). Samples were collected across the gradient and analyzed by Western blot.

Protein Sequencing

Proteins to be sequenced were concentrated 10-fold using a Speed Vac Concentrator (Savant Instruments, Farmingdale, N.Y.) and electrophoresed on a 10% SDS polyacrylamide gel. The gel was blotted to Westran PVDF membrane (Schleicher & Schuell, Keene, N.H.) using CAPS buffer (10 mM CAPS (pH 11), 10% methanol). The Westran membrane was stained with Coomassie Blue, the desired band excised and sent to the Louisiana State University Medical Center Core Laboratories for protein sequencing. The sequencing was performed on a pulsed liquid protein sequencer (ABI model 477A) and the PTH amino acids were analyzed with an on-line microbore HPLC system (ABI model 120A) using Edman chemistry.

Nucleic Acid Manipulation

DNA isolation was performed as described (Ausubel 1989). Restriction enzyme digestions, gel electrophoresis, Southern blot hybridization and subcloning were performed as described (Maniatis 1982) using the Bluescript plasmid vector and the *E. coli* strain XL-1 Blue (Stratagene Inc., La Jolla, Ca.). Genomic libraries were constructed using SS genomic DNA partially digested with Sau3A and the λ FIX II system (Stratagene Inc.). Low stringency hybridization and washes were performed at 3 X SSC and 55°C. High stringency hybridizations were performed at 3 X SSC and 65°C and washed to 0.1 X SSC.

Three oligonucleotides based on the sequence of the pro-enzyme segment of the cloned gene were designed to be specific for each *SAP* gene and were used for Southern blot hybridization;

SAP1 specific = 5' CCAGTAGCAT TAACAGGAGT TTTAATGACA 3'

SAP2 specific = 5' TGACCATTAG TAACTGGGAA TGCTTTAGGA 3'

SAP3 specific = 5' TTGATTTCAC CTTGGGGACC AGTAACATTT 3'

The oligonucleotides were labeled with polynucleotide kinase and g-^{32}P-ATP (Maniatis 1982) and used to probe Southern blots of genomic DNA. Hybridizations and washes were performed at 42° C in 5 X SSC.

DNA Sequencing

The regions of the phage 15 and 17 encoding the mature termini of the *SAP* genes were sequenced using the *fmol* DNA Sequencing System (Promega), as per manufacturer's instructions, and a synthetic oligonucleotide directed to the first section of the active site of the genes.

The region of phage 15 that cross-hybridizes to the PCR fragment from *SAP1* was subcloned as a 1 kb *Eco*RI fragment (see Figure 5). The entire *Eco*RI fragment (shown to contain the 5' end of *SAP3*; Figure 5) was sequenced using Sequenase as per manufacturer's instructions (United States Biochemical) and synthetic oligonucleotides. To complete the sequence of the coding region, a 7 kb *Pst* I fragment containing the entire gene was subcloned and the 3' end of the gene was sequenced as above.

RESULTS

UofL Library

To investigate the expression of the secreted aspartyl proteinases in various strains of *Candida albicans*, stationary phase culture supernatants were examined from strains 3153A, SS and WO-1 (W cells and O cells separately) grown in the proteinase-inducing media, YCB-BSA. Figure 1 (left panel) shows the Coomassie blue staining pattern of an SDS-PAGE gel of the culture supernatants. A single major protein band (43 kD) is detected in culture.

A rabbit antiserum which reacts with purified W cell secreted aspartyl proteinase (see Materials and Methods) was used as a probe against Western blots of the same culture supernatants. In Figure 1(right panel), a duplicate panel from the same gel was blotted to nitrocellulose and reacted with rabbit antiserum. The major protein band of 43 kD in supernatants from strains 3153A, SS and the W cells of strain WO-1 reacts with the anti-proteinase antibody, as do the two major proteins of 40 and 41 kD in supernatants from O cells of strain WO-1. This suggests that the secreted aspartyl proteinase is the major secreted protein(s) in all strains tested. The differences in size of the proteinase from culture supernatants of O cells will be examined below.

In all Western blots presented, a faint 43 kD band is observed in O cells (see Figure 1 (right panel)). The presence of the 43 kD protein in O cells is highly variable and probably reflects heterogeneity due to switching of O cells to W cells in an O cell population of strain WO-1. Estimates suggest that the

number of W cells in the O cell population account for the amount of 43 kD protein in the O cell culture supernatant (White, Miyasaki and Agabian, unpublished data).

The antibody to the *Candida albicans* proteinase was also used to probe Western blots of culture supernatants from other yeast species (including *C. tropicalis*, *C. krusei*, *C. glabrata* and *Saccharomyces cerevisiae*) grown in the proteinase-inducing medium, YCB-BSA . While aspartyl proteinases are secreted by other pathogenic *Candida* under these conditions, the antibody did not cross react with these proteinases (White, Miyasaki and Agabian, unpublished data) . Neither did these yeasts appear to degrade the BSA present in the YCB-BSA medium at the same rate as the *Candida albicans* strain. In contrast, over 40 clinical isolates of *Candida albicans* have been tested for protease production by growth in YCB-BSA. All strains tested produce proteinase (detectable by Western blot of culture supernatants) and substantially degrade the BSA present in the medium (White, Miyasaki and Agabian, unpublished data).

The results discussed above suggest that three different forms of aspartyl proteinase (40, 41 and 43 kD) can be secreted by a single strain of *Candida* (WO-1). To determine if the secretion of these different proteinases is dependent on the growth state of the cells, a time course of proteinase production and cell number (measured by optical density) was performed for each of the three *Candida albicans* strains and for both W and O. Figure 2

UofI Library

shows that the proteinases are all secreted at approximately the same time in the growth phase of the cultures. The proteinases are first detectable in the media in late log phase and their levels plateau in the media as the cells reach stationary phase. The secreted proteinases persist in the media, without significant degradation for up to 5 days (White, Miyasaki and Agabian, unpublished data). The appearance of the proteinases is correlated with the depletion of the bovine serum albumin (BSA) from the media as the cells reach stationary phase. This result is predicted if proteinase is induced by the BSA which is the sole nitrogen source in the media (Morrow, Srikantha and Soll 1992). However, O cells produce proteinase in YEPD media (which does not contain BSA) with the same kinetics (24). No other cell type produces proteinase under these conditions. The 41 kD protein, which is frequently obscured by the 40 kD protein, appears later in the time course of growth in O cells (see Figure 2).

Preparative isoelectric focusing was used to determine the pI of the three observed proteinases. A mixture of culture supernatants from W cells and O cells was used in a preparative isoelectric focusing run. As shown in Figure 3, the 43 kD protein has an isoelectric point at approximately pH 4.25 and the 40 kD protein has an isoelectric point of approximately pH 4.0. Both of these values are consistent with the known pI of the secreted aspartyl proteinases (Ray, Payne and Morrow 1991 and Rüchel et al. 1992). Surprisingly, the 41 kD

UCSF LIBRARY

protein has an isoelectric point of pH 5.7, much higher than any previously observed pI for a secreted proteinase from *Candida.*

There are several possible explanations for the presence of three protein bands reacting with anti-proteinase antibody in Western blots. Proteinase activities with distinct characteristics have been observed under different growth conditions by several researchers (Rüchel et al. 1992) suggesting that the three protein bands may be proteinase isoenzymes. The proteinase genes that have been sequenced are known to contain a signal sequence and perhaps a propeptide form suggesting that the three proteins may be processing intermediates of a single proteinase. The proteinase may also be post-translationally processed (e.g. phosphorylation or glycosylation) or selectively degraded in the culture supernatants. To characterize the three proteinase types, culture supernatants were electrophoresed in an SDS-PAGE gel and prepared for microsequencing as described in the Materials and Methods. The N-terminal sequence of each of the three protein bands is shown in Figure 4. For each protein band, there is a unique but highly-related N-terminal sequence.

Two of the N-terminal protein sequences align completely with the predicted protein sequences of the two cloned proteinase genes, *SAP1* (Hube et al. 1991) and *SAP2* (Wright et al. 1992) (see Figure 4). The 43 kD protein from W cells corresponds to the predicted mature product of the *SAP2* gene and the 40 kD protein from O cells corresponds to the predicted mature product of the *SAP1* gene. The 41 kD protein sequence does not correspond to a previously

characterized gene product. A third gene has been cloned (see below) with a predicted protein sequence which matches the 41 kD protein N-terminal sequence, which we will refer to as *SAP3*. The three protein sequences that were obtained resemble the N-terminal sequences of several mature proteinases from unrelated *Candida albicans* strains that have been previously reported (see Figure 4 and references). Since the three protein bands were sequenced from the same strain, WO-1, this data clearly demonstrates that three different isoenzymes of aspartyl proteinase are secreted by the WO-1 strain of *Candida albicans* and that most protein sequences obtained to date can be grouped into these three proteinase types.

From the apparent molecular weights of the proteins (Figure 1) and the N-terminal sequence data (Figure 4), there appears to be a correspondence between the Sap isoenzyme and its apparent molecular weight. However, Sap isoenzymes might be modified post-translationally (e.g. glycosylation) under a variety of conditions resulting in varying molecular weights . To examine the correspondence of isoenzyme and molecular weight, the N-terminal sequences of the 43 kD proteinase produced by strains 3153A and SS in YCB-BSA media were determined and compared to the sequences from strain WO-1. The results (Figure 4) confirm that 43 kD proteinases in all three strains, 3153A, SS and WO-1 are the products of the *SAP2* gene locus.

The sequence of both alleles of the *SAP1* gene from strain WO-1 and from an unrelated strain, SS (N-terminal sequences shown in Figure 4) has been

determined (Miyasaki, White and Agabian 1994). The *SAP2* and *SAP3* N-terminal sequences (Figure 4) are distinctly different from the alleles of *SAP1* that have been sequenced. Therefore, *SAP2* and *SAP3* are not alleles of *SAP1* but are encoded by different genetic loci and thus different proteinase isoenzymes. To identify the gene for *SAP3* , a λ genomic library was screened with a PCR fragment from *SAP1* using low stringency hybridization conditions (Miyasaki, White and Agabian 1994). Two λ phage were identified which hybridize with *SAP1* at low stringency but not at high stringency. These two recombinant λ phage, 15 and 17, were isolated and the region encoding the mature N-terminus of the protein was sequenced for each phage. The sequences were translated into amino acid sequences which are listed in Figure 4. The deduced protein sequence of phage 15 corresponds to the protein sequence determined for the 41 kD protein and the deduced protein sequence of phage 17 corresponds to the protein sequence of the *SAP2* gene. In the *SAP1* and *SAP2* gene sequences, the last two amino acids of the processed signal and propeptide sequences are Lys-Arg, suggesting that these peptides are cleaved from the mature form of the proteinase by a proteinase mechanistically related to *KEX2* in *Saccharomyces cerevisiae* (Julius et al. 1989). The same amino acids, Lys-Arg, precede the N-terminus of the mature protein in Sap3, as determined by the nucleotide sequence of phage 15.

The entire *SAP3* gene from phage 15 was sequenced and is shown in Figure 5. The gene encodes a protein of 398 amino acids including a signal

sequence/proenzyme fragment of 58 amino acids. The *SAP3* gene is approximately 77% homologous to both the *SAP1* and *SAP2* genes while the *SAP1* and *SAP2* genes are approximately 77% homologous to each other, suggesting that all three genes diverged from a common ancestor at approximately the same time. The regions of strongest homology in all three genes include the two sections of the active site of the enzyme. The N-terminal extension of Sap3, containing the signal sequence, is separated from the mature form of the protein by the dipeptide Lys-Arg. However, the N-terminal extension of Sap3 does not contain a second Lys-Arg dipeptide that is found in the Sap1 and Sap2 extensions (see Table 1).

Comparison of the DNA sequences of *SAP3* and the published sequences of the *SAP1* and *SAP2* genes indicate that the proenzyme segment of each protein, between the signal sequence and the mature N-terminus, is the most divergent region of the proteins (White, Miyasaki and Agabian, unpublished data). Three oligonucleotides (30 bases in length) were designed to distinguish between the three genes. The oligonucleotides were used to determine the complexity of the protease genes by Southern blot analysis of *Eco*RI digested genomic DNA from three different strains. The results (Figure 6) show that the three oligonucleotides, representing the three different genes, are present at independent loci in all three strains, as defined by the unique DNA fragments that each recognizes in genomic Southern blots. The *SAP1* probe recognizes a single 2 kb DNA fragment containing the *SAP1* gene in strains

3153A and SS and recognizes two DNA fragments of 2 and 2.3 kb in strain WO-1 which represent an RFLP at the *SAP*1 gene (Miyasaki, White and Agabian 1994). The *SAP*2 probe recognizes the *SAP*2 gene on a 20 kb DNA fragment and cross-hybridizes to a 1.5 kb DNA fragment which may represent an unidentified SAP gene. The *SAP*3 probe hybridizes to the *SAP*3 gene on a 1.0 kb DNA fragment and to several higher molecular weight DNA fragments, suggesting that there are other related proteinase genes which have not yet been identified. The higher molecular weight bands from *SAP*3 are not due to partial digestion since the *SAP*1 blot, which used the same DNA, did not show evidence of partial digestion. These three oligonucleotide probes should be useful for the analysis of the three genes in most, if not all, strains of *Candida albicans* and may help to identify the related proteins in other *Candida* species.

DISCUSSION

The above results demonstrate that there are at least three isoenzymes of secreted aspartyl proteinase present in *Candida albicans* and that all three are secreted from at least one strain, WO-1. The three proteinase isoenzymes differ in primary sequence, pI, and pattern of expression (summarized in Table 1) and are products of three separate genetic loci (Figure 6). The differences between the three proteinases suggest that the three proteinases may have unique roles in the interaction between *Candida* and its host.

The proteinases are sufficiently distinct to be readily identified (Table 1). The most straightforward method of distinguishing the three proteinases is their

apparent molecular weight on SDS-PAGE using a polyclonal antiserum (such as ours) which recognizes all three proteinase isoenzymes. A previously reported monoclonal antibody may be specific to the Sap1 isoenzyme since it does not recognize proteinase in both W and O cells (Ray, Payne and Morrow 1991).

The three isoenzymes of secreted aspartyl proteinase can be easily identified based on their N terminal sequences (see Figure 4). For the purpose of this discussion, the three letter amino acid code is used and each isoenzyme is referred to by its corresponding genetic designation. Accordingly, over the length of the N-terminal sequence (15 amino acids) of the mature peptide, the three isoenzymes are identical in 8 out of 15 positions. While many amino acid differences exist between the three isoenzymes (see Figure 4), the most distinguishing feature of the three sequences in this region is position 8 where all three Saps differ in amino acid sequence; Sap1 has an Asn, Sap2 has a His and Sap3 has an Ile. Morrison et. al. (1993) have identified a fourth, unrelated Sap protein, the first 6 amino acids of which are highly divergent. However, at positions 3 and 8, this protein resembles Sap3.

Secreted aspartyl proteinases from other species of *Candida* have also been cloned. These include one gene from *Candida tropicalis* (Togni et al. 1991) and two tandemly-linked genes from *Candida parapsilosis* (DeViragh et al. 1993). The N-terminal sequences of all *Candida* secreted aspartyl proteinase genes are identical at positions 7, 10, and 14 ; and highly similar at positions 9 and 15 (See Figure 4). The predicted N-terminal sequence from the *C. tropicalis*

proteinase gene most closely resembles Sap3 of *C. albicans* (Figure 4, positions 3, 8 and 13). The predicted N-terminal sequences from the *C. parapsilosis* genes do not closely resemble any of the three *C. albicans* genes but may be more similar to Sap3 based on positions 8 and 13. The *C. tropicalis* and *C. parapsilosis* genes are more closely related to each other than to the *C. albicans* genes (Figure 4, positions 1, 8, 11 and 12).

The large size difference between the Sap proteins (40 kD, 41 kD and 43 kD) may be due to glycosylation. While the proteins are 77% similar in primary sequence, one major difference is the number of potential N-linked glycosylation sites (Asn - any amino acid - Ser or Thr) in the mature protein. Sap1 has no N-linked sites, Sap3 has one N-linked site at position 313, adjacent to the second active site, and Sap2 has 2 potential N-linked sites, one at 313 homologous to the Sap3 site, and an adjacent site at position 321 which is not conserved in the other proteins. The number of sites correlates with the apparent molecular weights of the proteins. However, attempts to confirm N-glycosylation of the SAP proteinases by treatment with N-glycosidase F or with endoglycosidase H in the presence or absence of SDS or N-octylglucoside have been unsuccessful. Chemical treatment with periodate or with trifluoromethanesulfonic acid (Oxford Glycosystems) which cleave both N and O linked carbohydrates also failed to cleave the Sap proteins, suggesting that if the proteins are glycosylated, the residues are not accessible to enzymatic or chemical cleavage.

The amounts of proteinase detected by Western blot of culture supernatants do not agree with the levels of specific activity that have been reported (Morrow, Srikantha and Soll 1992 and Ray, Payne and Morrow 1991). For instance, the protein levels of Sap2 in W cells and Sap1 in O cells are approximately the same, as determined by band intensity in Coomassie Blue stained SDS-PAGE (Figure 1). Yet it has been reported that O cells have 20-fold higher levels of proteinase activity (Morrow, Srikantha and Soll 1992 and Ray, Payne and Morrow 1991). One likely explanation for this discrepancy is that the three isoenzymes have different substrate specificities and that the biochemical assay of proteinase activity is biased in favor of detection of one isoenzyme. In addition, the pH of the growth media (such as YCB-BSA) varies greatly depending on which cell type (W or O) is in culture (White and Agabian, unpublished data). Therefore, secreted proteinases with different pI and inactivation spectra may be inactivated in media of differing pH before being assayed.

An intracellular, non-secreted aspartyl proteinase gene from *Candida albicans* has been described (Lott et al. 1989). No region of the predicted protein sequence from that gene resembles the N-terminal sequences of the secreted aspartyl proteinase isoenzymes. Neither does the gene contain a signal sequence nor a correctly placed Lys-Arg dipeptide which might function as a cleavage site in a pro-form of the enzyme.

Our detection of the Sap1 and Sap3 isoenzymes in culture supernatants from O cells of strain WO-1 is consistent with the results of Morrow and colleagues (1992). In that study, *SAP1* cDNAs were identified as differentially transcribed mRNAs, present in O cells and absent from W cells. When these cDNAs were used as probes, *SAP1* mRNA was detected in Northern blots of O cells while little, if any, *SAP1* mRNA was detected from W cells.

The identification of three distinct proteinase isoenzymes in *Candida albicans* and the observed expression of all three isoenzymes from a single strain (WO-1) under different switch phenotypes allows us to address the signals and mechanism involved in the switch between W and O cells in strain WO-1 and will affect the genetic analysis of the role of the proteinases in virulence. It should be noted that the pattern of proteinase expression, and the isoenzymes expressed, at the site of infection and disease may not resemble the pattern of expression generated in any of the growth conditions currently in use. We are currently attempting to determine *in situ* which proteinases are expressed in regions of oral and vaginal lesions. This will be crucial in evaluating the role of each of the three proteinases in pathogenesis.

ACKNOWLEDGMENTS

We thank Dr. David Soll, University of Iowa, for providing strain WO-1, and for his endless support in our entry into the field of Candida research; Dr. Remo Morelli, San Francisco State University, for kindly providing strain SS; Chiron Corp. for supplying the adjuvant used in the preparation of antiserum; and the members of our laboratory for many helpful discussions. This work was supported in part by a grant from the N.I.H. (#1RO1AI33317) and a John D. and Catherine T. MacArthur Foundation Award (#84-85) to N.A., a Universitywide AIDS Research Program Young Investigator Award from the State of California (#K92-SF-011) to T.C.W. and an N.I.H. Dentist-Scientist Award (5K15DE00334) to S.H.M.

REFERENCES - CHAPTER 4

1. Anderson, J. M. and D. R. Soll. 1987. Unique Phenotype of Opaque Cells in the White-Opaque Transition of *Candida albicans*. *J. Bacteriol.* 169: 5579-88.
2. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, K. Struhl, eds. 1989. Current Protocols in Molecular Biology. New York, John Wiley & Sons.
3. Blagoyevich, P. G. 1979. Masters thesis. San Francisco State University.
4. Cutler, J. E. 1991. Putative Virulence Factors of *Candida albicans*. *Ann. Rev. Microbiol.* 45:187-218.
5. DeViragh, P. A., D. Sanglard, G. Togni, R. Falchetto, M. Monod. 1993. Cloning and sequencing of two *Candida parapsilosis* genes encoding acid proteinases. *J. Gen. Microbiol.* 139:335-42.
6. Douglas, L. J. 1986. *Candida* Proteinases and candidosis. *CRC Critical Rev. in Biotech.* 8:121-9.
7. Ganesan, K., A. Banerjee and A. Datta. 1991. Molecular Cloning of the Secretory Acid Proteinase Gene from *Candida albicans* and Its Use as a Species-Specific Probe. *Infect. Immun.* 59:2972-7.
8. Harlow, E., and D. Lane. 1988. Antibodies, A Laboratory Manual. Cold Spring Harbor, N.Y. Cold Spring Harbor Laboratory.

9. Hube, B., C. J Turver, F. C. Odds, H. Eiffert, G. J. Boulnois, H. Köchel, and R. Rüchel. 1991. Sequence of the *Candida albicans* gene encoding the secretory aspartate proteinase. *J. Med. Vet. Mycol.* 29:129-32.

10. Julius, D.,. A. Blake. L. Blair, R. Kunisawa and J. Thorner. 1984. Isolation of the putative structural gene for the Lysine-Arginine-cleaving endopeptidase required for processing of yeast prepro-a-factor. *Cell* 37:1075-1089.

11. Lott, T.J., L. S. Page, P. Boiron, J. Benson and E. Reiss. 1989. Nucleotide sequence of the *Candida albicans* aspartyl proteinase gene. *Nucl. Acids Res.* 17:1779.

12. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning, A Laboratory Manual. Cold Spring Harbor, N.Y. Cold Spring Harbor Laboratory.

13. Miyasaki, S. H., T. C. White and N. Agabian. 1994. A fourth secreted aspartyl proteinase gene (*SAP*4) and a *CARE*2 repetitive element are located upstream of the *SAP*1 gene in *Candida albicans. J. Bacteriol.* 176:1702-10.

14. Morrison, C. J., S. F. Hurst, S. L. Bragg, R. J. Kuykendall, H. Diaz, J. Pohl and E. Reiss. 1993. Heterogeneity of the Purified Extracellular Aspartyl Proteinase from *Candida albicans:* Characterization with Monoclonal Antibodies and N-Terminal Amino Acid Sequence Analysis. *Infect. Immun.* 61:2030-6.

15. Morrow, B., T. Srikantha and D. R. Soll. 1992. Transcription of the Gene for a pepsinogen, *PEP1*, Is Regulated by White-Opaque Switching in *Candida albicans. Mol. Cell. Biol.* 12:2997-3005.

16. Odds, F. C. 1988. *Candida* and Candidosis; A Review and Bibliography., Philadelphia, Baillière Tindall.

17. Ray, T. L., C. D. Payne and B. J. Morrow. 1991. *Candida albicans* acid proteinase: characterization and role in candidiasis. In: Structure and Function of the Aspartic Proteinases. New York, B.M. Dunn (ed.) Plenum Press,

18. Rüchel, R., F. De Bernardis, T. L. Ray, P. A. Sullivan and G.T. Cole. 1992. *Candida* acid proteinases. *J. Med. and Vet. Mycol.* 30, supp. 1:123-32.

19. Slutsky, B., M. Staebell, J. Anderson, L. Risen, M. Pfaller and D. R. Soll. 1987. "White-Opaque Transition"; a Second High-Frequency Switching System in *Candida albicans. J. Bacteriol.* 169:189-97.

20. Soll, D. R. 1992. High-frequency switching in *Candida albicans. Clin. Microbiol. Rev.* 5:183-203.

21. Staib, F. 1965. Serum-proteins as nitrogen source for yeast-like fungi. *Sabouraudia* 4:187-93.

22. Togni, G., D. Sanglard, R. Falchetto, M. Monod. 1991. Isolation and nucleotide sequence of the extracellular acid proteinase gene (ACP) from the yeast, *Candida tropicalis. FEBS Letters* 286:181-5.

23. White, T. and N. Agabian. Unpublished data.

24. White, T.C., S. H. Miyasaki, and N. Agabian. Unpublished data.

25. Woodward, M. P., W. W. Young Jr. and R. A. Bloodgood. 1985. Detection of Monoclonal Antibodies Specific for Carbohydrate Epitopes Using Periodate Oxidation. *J. Immunol. Methods* 78:143-53.

26. Wright, R. J., A. Carne, A. D. Hieber, I. L. Lamont, G. W. Emerson and P. A. Sullivan. 1992. A Second Gene for a Secreted Aspartate Proteinase in *Candida albicans*. *J. Bacteriol.* 174:7848-53.

27. Yamamoto, T., K. Nohara, K. Uchida, H. Yamaguchi. 1992. Purification and characterization of Secretory Proteinase of *Candida albicans*. *Microbiol. Immunol.* 36:637-64.

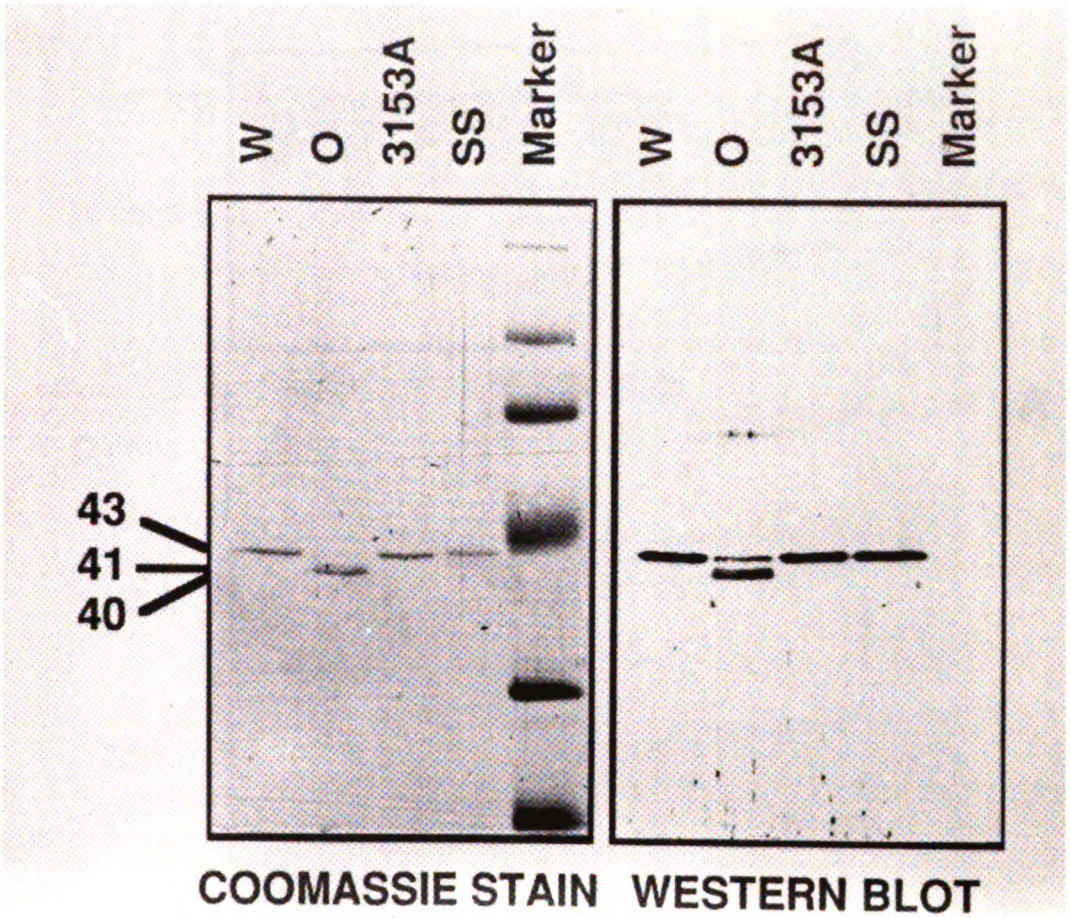


Figure 1

DETECTION OF THREE PROTEINASES

Western Blot of an SDS-PAGE of culture supernatants of strains 3153A, SS, and W and O cells from strain WO-1. The left panel in the figure shows a Coomassie Blue stain of one half of an SDS-PAGE gel. The marker lane contains approximately 1 µg of each protein band. The size of each band (from the top) are: 208, 101, 71, 44, 29,18+15 kDa. The right panel in the figure shows a Western blot of the second half of the same gel probed with the anti-proteinase antibody (see Materials and Methods).

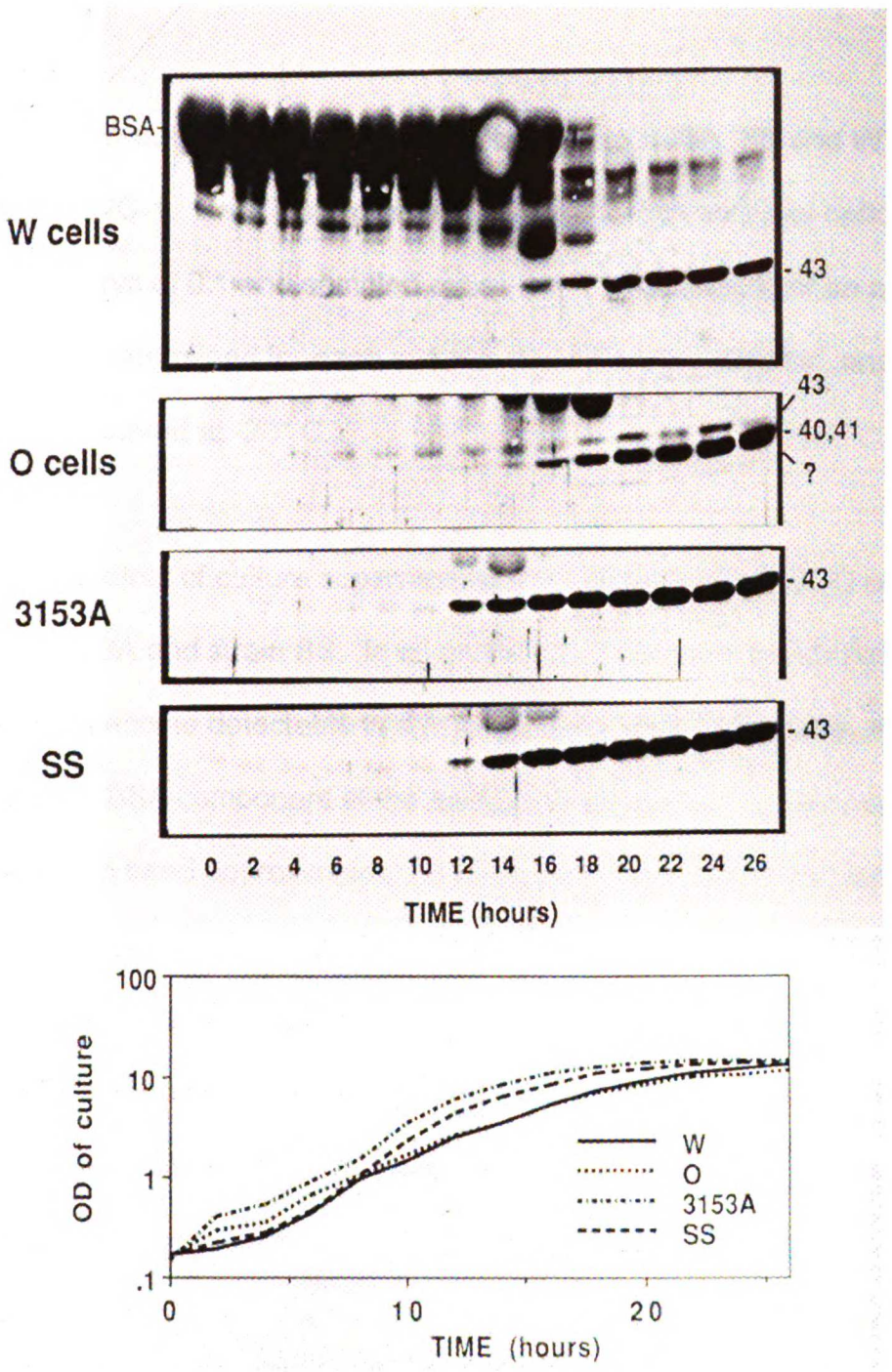

BSA
W cells
43
O cells
43
40,41
?
3153A
43
SS
43
0 2 4 6 8 10 12 14 16 18 20 22 24 26
TIME (hours)
A
100
10
1
.1
OD of culture
W
O
3153A
SS
0
10
20
TIME (hours)
B

Figure 2

Figure 2 Legend. EXPRESSION OF THREE PROTEINASES DURING GROWTH.

Growth curves were constructed for strains 3153A, SS and W and O cells from strain WO-1. Cells were grown in YCB-BSA from an initial optical density (OD) at 600 nm of 0.1 and sampled every two h. The OD_{600} of an appropriate dilution was determined for each sample, the cells were pelleted, and the culture supernatant stored at -20° C.

(A) Western blots of culture supernatants from W cells of WO-1, O cells of WO-1, strain 3153A and strain SS. In all cases, the proteinase bands (labeled 40, 41 and 43 kD) become detectable in mid to late log phase of the cells, at the same time that the BSA component of the medium is degraded. In the O cell Western blot, a protein band approximately 38 kD in size is also detected, possibly a degradation product of one of the proteinases. In all panels, a faint band is visible at 43 kD throughout the time course. This band is a degradation product of the BSA component of the media. The BSA component of the medium is shown in the W cell panel. The BSA is detected in the Western in part because of BSA contamination of the antigen that was used to prepare the antiserum, and in part because the secondary antibody recognizes mammalian albumin.

(B) shows growth of the four cultures as determined by the OD at 600 nm.

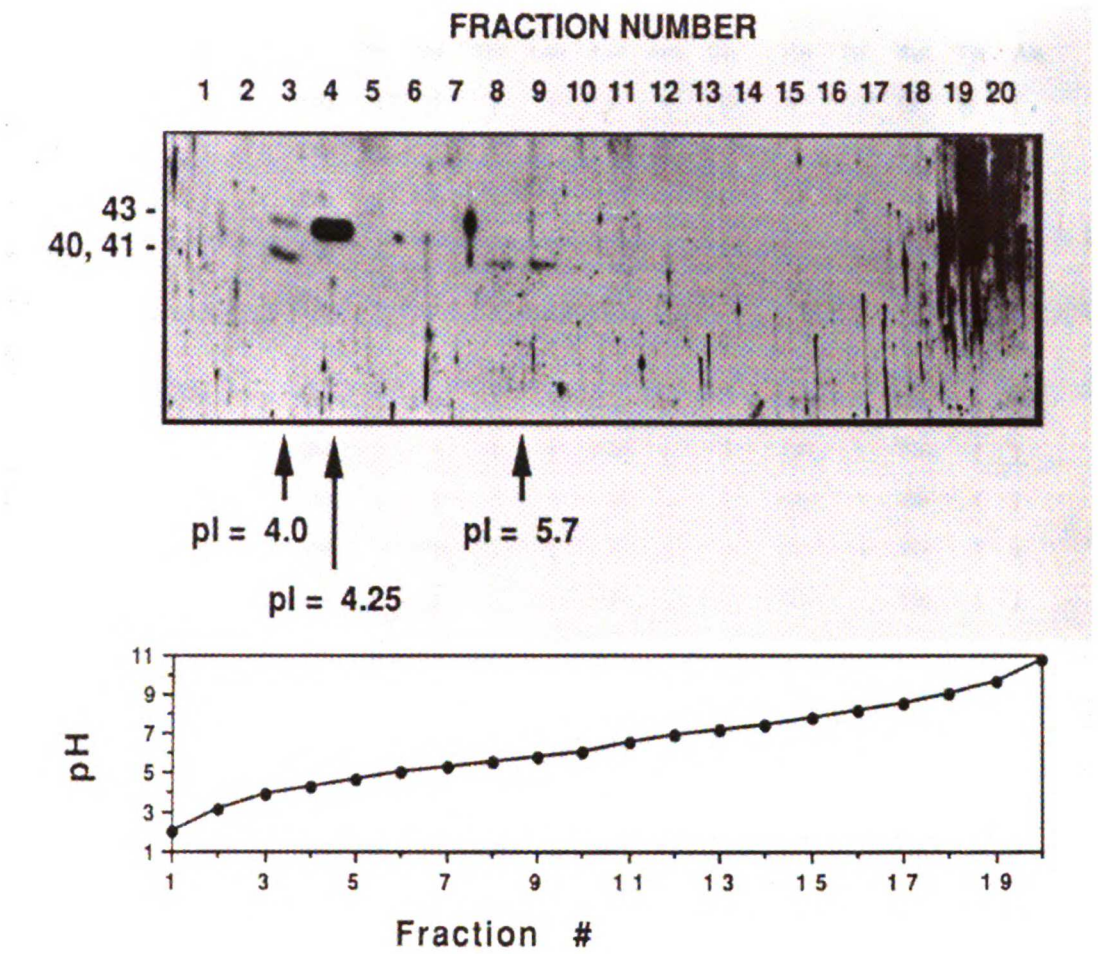


Figure 3

Isoelectric focusing of the three isoenzymes. A mixture of W and O culture supernatants were subjected to preparative isoelectric focusing (see Materials and Methods). Fractions were collected and stored at -20° C for analysis. The top panel shows the Western blot of a sample of each of the fractions. The proteinase bands (40, 41 and 43 kDa), are visible in fractions 3, 4, 8 and 9. The pH for each of the fractions is indicated below the Western.

Figure 4

SAP1 Sequences:																
10231 DNA	Gln	Ala	Ile	Pro	Val	Thr	Leu	Asn	Asn	Glu	His	Val	Ser	Tyr	Ala	9
WO-1, 40 kD PROTEIN	|	|	Leu	|	|	|	|	|	|	|	X	|	(|)	(|)	|	•
WO-1, 2.4 allele, DNA	|	|	Leu	|	|	|	|	|	|	|	|	|	|	|	|	|	13
WO-1, 2.6 allele, DNA	|	|	Leu	|	|	|	|	|	|	|	|	|	|	|	|	|	13
SS, 2.4 allele, DNA	|	|	Leu	|	|	|	|	|	|	|	|	|	|	|	|	|	13
SS, 2.8 allele, DNA	|	|	Ile	|	|	|	|	|	|	|	|	|	|	|	|	|	13
SAP2 Sequences:																
10261 DNA	|	|	Val	|	|	|	|	His	|	|	Gln	|	Thr	|	|	26
WO-1, 43 kD PROTEIN	|	|	Val	|	|	|	|	(His)	|	|	Gln	|	Thr	|	|	•
3153A, 43 kD PROTEIN	X	|	Val	|	|	|	|	X	|	|	Gln	|	Thr	|	|	•
SS, 43 kD PROTEIN	|	|	Val	X	|	|	|	X	|	|	Gln	|	Thr	|	|	•
SC5314 PROTEIN	|	|	Val	|	|	|	|	His	|	|	Gln	|	Thr	|	|	7
CBS 2730 PROTEIN	|	|	Val	|	|	|	|	His	|	|						14
#114 PROTEIN	|	|	Val	|	|	|	|	His	|	|	Gln	|	Thr	|	|	27
Phage 17 DNA	|	|	Val	|	|	|	|	His	|	|	Gln	|	Thr	|	|	•
SAP3 Sequences:																
WO-1, 41 kD PROTEIN	X	X	Val	|	|	Lys	|	Ile	|	|	Gln	|	(|)	|	|	•
#74 PROTEIN	Glu	|	Val	|	|	|	|	Ile	|	|						22
Phage 15 DNA	|	Thr	Val	|	|	Lys	|	Ile	|	|	Gln	|	|	|	|	•
Other *SAP* Sequence:																
CBS2730	Gly	Thr	Val	Gln	Thr	Ser	|	Ile	|	|							14
Other Species Sesquences:																
C. tropicalis	Ser	Asp	Val	|	Thr	|	|	Ile	|	|	Gly	Pro	|	|	|	22
C. parapsilosis (left)	Ser	Ser	Pro	Ser	Ser	Pro	|	Tyr	Phe	|	Gly	Pro	|	|	Gly	5
C. parapsilosis (right)	Asp	Ser	Ile	Ser	Leu	Ser	|	Ile	|	|	Gly	Pro	|	|	|	5
Amino Acid Position:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	

Figure 4 Legend N-TERMINAL SEQUENCES OF SECRETED ASPARTYL PROTEINASES.

The N-terminal sequences of secreted aspartyl proteinases of *Candida albicans* can be grouped into three classes corresponding to the three isoenzymes described in this paper. Each sequence is listed with the strain from which the sequence was derived. The designation DNA after the strain signifies that the protein sequence was determined by translation of the DNA sequence. A vertical line (|) in the sequence represents identity with the top sequence (10231 DNA) . An (X) in the sequence represents a residue that could not be determined by protein sequencing. An amino acid in parentheses represents a residue in which the protein sequence is not definitive. The reference for each sequence is listed at the right of the figure. An asterisk (*) in the references indicates that the sequence was determined in this paper.

```
-229  GAATTCAACTTGGTGCCGTTATCAAAACAACCACACTTGTGAGAACCTCCTTCCCAGTAAAACTGAAGTAGGATAGAGT
-150  AGCTCTCATTGGGAAAGGGGTATTACCATGATTGAAGAATAGTGAACAATATATAAATAAGGCTGGATAACCTTCTCTA
-71   AAATTATGGATTGGAACATTTCTAATTCAATCAATTAACTTCCACACACACATACACATATATATAACAA

1    ATG TTT TTA AAA
     M   F   L   K

13   AAT ATC TTT ATT GCT CTT GCT ATT GCT TTA TTA GCT GAT GCT ACT CCA ACA ACT TTC AAC
     N   I   F   I   A   L   A   I   A   L   L   A   D   A   T   P   T   T   F   N

73   AAT TCT CCA GGG TTT GTT GCT TTG AAT TTT GAT GTT ATC AAA ACT CAT AAA AAT GTT ACT
     N   S   P   G   F   V   A   L   N   F   D   V   I   K   T   H   K   N   V   T
                                                             ======================

133  GGT CCC CAA GGT GAA ATC AAT ACC AAC GTC AAC GTC AAG AGA CAA ACT GTT CCA GTT AAA
     G   P   Q   G   E   I   N   T   N   V   N   V   K   R   Q   T   V   P   V   K
     ==================================

193  TTA ATT AAT GAA CAA GTT AGT TAT GCT TCT GAT ATT ACT GTT GGT TCC AAT AAA CAA AAA
     L   I   N   E   Q   V   S   Y   A   S   D   I   T   V   G   S   N   K   Q   K

253  TTA ACT GTT GTT ATT GAT ACT GGA TCA TCT GAT TTA TGG GTT CCT GAT TCT CAA GTT TCA
     L   T   V   V   I   D   T   G   S   S   D   L   W   V   P   D   S   Q   V   S

313  TGT CAA GCT GGT CAA GGA CAA GAT CCA AAT TTT TGT AAA AAT GAA GGA ACT TAT TCC CCA
     C   Q   A   G   Q   G   Q   D   P   N   F   C   K   N   E   G   T   Y   S   P

373  AGT TCT TCA AGT AGT TCT CAA AAT TTG AAT AGT CCA TTT AGT ATT GAA TAT GGT GAT GGA
     S   S   S   S   S   S   Q   N   L   N   S   P   F   S   I   E   Y   G   D   G

433  ACT ACT TCA CAA GGG ACA TGG TAT AAA GAT ACT ATT GGA TTT GGT GGT ATT TCT ATC ACA
     T   T   S   Q   G   T   W   Y   K   D   T   I   G   F   G   G   I   S   I   T

493  AAG CAA CAA TTT GCC GAT GTT ACT AGT ACA TCA GTT GAT CAA GGG ATT TTA GGG ATT GGT
     K   Q   Q   F   A   D   V   T   S   T   S   V   D   Q   G   I   L   G   I   G

553  TAT AAA ACT CAT GAA GCT GAA GGT AAT TAT GAT AAT GTT CCT GTG ACT TTA AAA AAT CAA
     Y   K   T   H   E   A   E   G   N   Y   D   N   V   P   V   T   L   K   N   Q

613  GGA ATT ATT TCT AAA AAT GCT TAT TCA CTT TAT CTT AAT TCA AGA CAA GCC ACT AGT GGA
     G   I   I   S   K   N   A   Y   S   L   Y   L   N   S   R   Q   A   T   S   G

673  CAA ATT ATT TTT GGT GGT GTT GAT AAT GCT AAA TAT AGT GGG ACA TTG ATT GCT TTA CCA
     Q   I   I   F   G   G   V   D   N   A   K   Y   S   G   T   L   I   A   L   P

733  GTT ACT TCT GAT AAT GAA TTA AGA ATT CAT TTG AAT ACT GTA AAA GTT GCT GGA CAA TCC
     V   T   S   D   N   E   L   R   I   H   L   N   T   V   K   V   A   G   Q   S

793  ATT AAT GCT GAT GTT GAT GTT TTG TTG GAT TCA GGT ACT ACC ATT ACT TAT TTA CAA CAA
     I   N   A   D   V   D   V   L   L   D   S   G   T   T   I   T   Y   L   Q   Q

853  GGT GTT GCT GAT CAA GTG ATT AGT GCT TTT AAT GGT CAA GAA ACT TAT GAT GCT AAT GGT
     G   V   A   D   Q   V   I   S   A   F   N   G   Q   E   T   Y   D   A   N   G

913  AAT CTT TTC TAT CTT GTT GAT TGT AAT TTG TCA GGA TCA GTT GAT TTT GCT TTT GAT AAA
     N   L   F   Y   L   V   D   C   N   L   S   G   S   V   D   F   A   F   D   K

973  AAT GCT AAA ATT TCC GTT CCA GCT TCT GAA TTT ACT GCT CCA TTA TAC ACT GAA GAT GGT
     N   A   K   I   S   V   P   A   S   E   F   T   A   P   L   Y   T   E   D   G

1033 CAA GTT TAT GAT CAA TGT CAA CTT CTT TTT GGA ACT AGT GAT TAT AAC ATT CTT GGT GAT
     Q   V   Y   D   Q   C   Q   L   L   F   G   T   S   D   Y   N   I   L   G   D

1093 AAT TTC TTG AGA TCA GCC TAT ATT GTT TAT GAT TTG GAT GAT AAT GAA ATT TCA TTA GCT
     N   F   L   R   S   A   Y   I   V   Y   D   L   D   D   N   E   I   S   L   A

1153 CAA GTT AAG TAT ACT ACT GCT TCT AAC ATT GCT GCT CTT ACT TAGGGGGAAGTACTCCTC
     Q   V   K   Y   T   T   A   S   N   I   A   A   L   T   *

     TGGAGATTGATTCTTATTTTTGATTCTATTCAATAGAACATTTCGGTTTAGCTTGGTTTATTATTGTTAGACTTTATAGCT
1213 TCT TTT TGT TTA CTT TTT ACT TTT ACT TTA TTA G
1295
```

Figure 5

Figure 5 Legend SEQUENCE OF THE *SAP3* GENE.

The sequence of the *SAP3* gene is shown, together with the deduced amino acid sequence of the protein. The sequence begins at the first of two *Eco*RI sites in the sequence (underlined). The putative N-terminus of the mature protein, which matches the protein sequence of the 41 kD protein, has a double line above the sequence. The two sections of the active site of the proteinase which are highly conserved between the Sap proteins have a single line above the sequence. The entire sequence is 1556 nucleotides in length. The Genbank Assession Number is L22358.

TABLE 1. Characteristics of the *SAP* Gene Products

Characteristic:	Sap1	Sap2	Sap3
Apparent Molecular Weight (10^3)	40 kD	43 kD	41 kD
Predicted Molecular Weight (10^3)	36.2 kD	36.3 kD	36.5 kD
Isoelectric point	pI 4.0	pI 4.25	pI 5.7
Signal sequence	50 a.a.	56 a.a.	58 a.a.
Lys-Arg Dipeptides in Signal Sequence	2	2	1
Potential N linked glycosylation sites[a]	0	2	1
Expression in Strain WO-1 in YCB-BSA	+ (O)	+ (W)	+ (O)[b]
Expression in Clinical Strains in YCB-BSA[c]	-	+	-

[a]Sites (N-X-S/T) in the mature protein.

[b] Sap3 expression has been observed in W cell types. However, expression is at low levels.

[c]White, Agabian and Miyasaki unpublished data

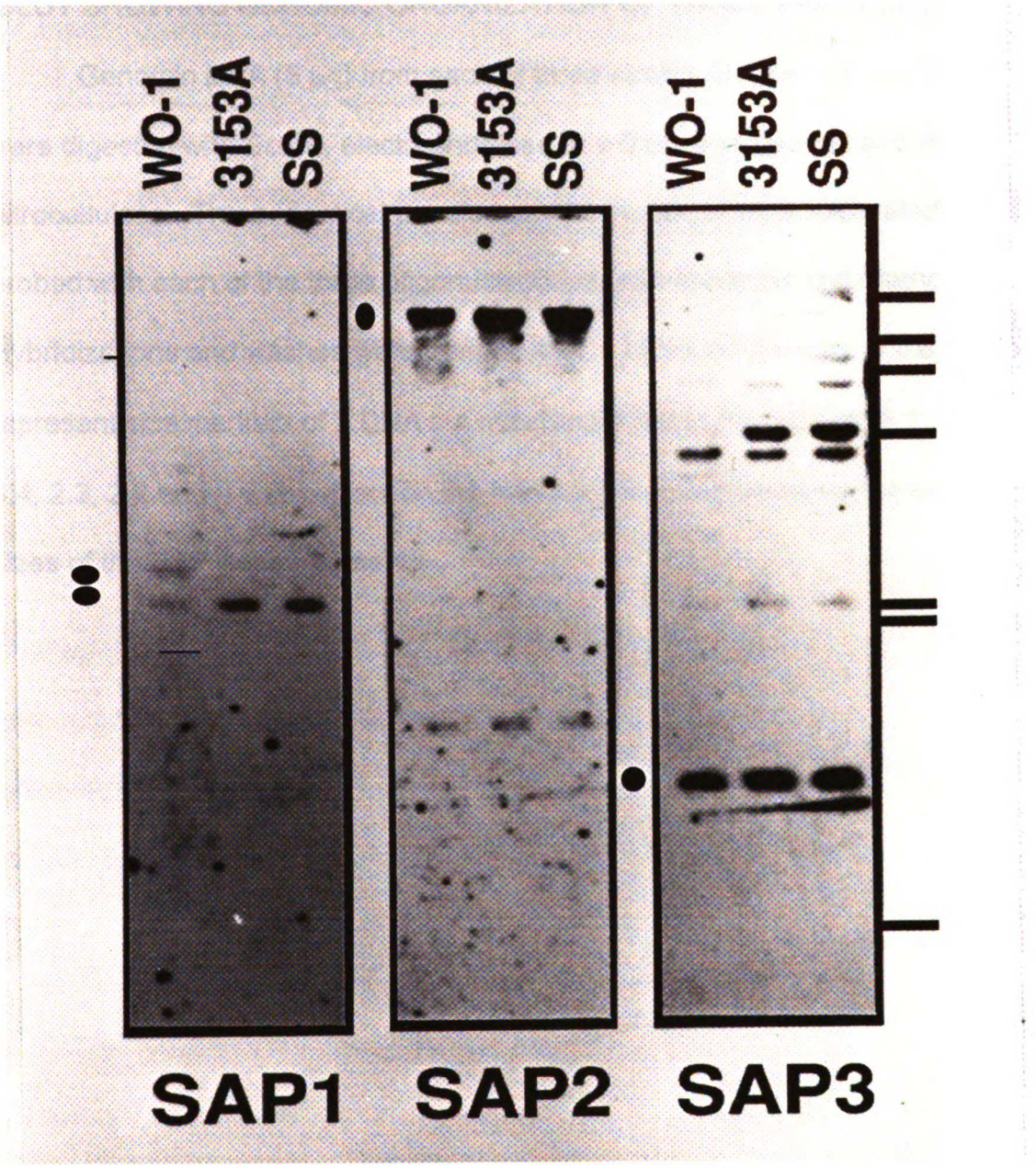


Figure 6

Figure 6 Legend

BLOT SHOWING GENOMIC ORGANIZATION OF THREE PROTEINASES.

Genomic DNA (5 μg) from each of three strains (3153A, SS and WO-1) were digested with EcoRI, electrophoresed in a 0.8% agarose gel and blotted to nitrocellulose. Three identical panels from the same gel were separated and probed with each of the three oligonucleotides (see Materials and Methods). Hybridizations and washes were done at 42°C. Lines on the right of the figure represent size markers of λ DNA cut with HindIII (sizes from top = 23.1, 9.4, 6.6, 4.4, 2.3, 2.2 and 0.6 kb). Dots on the left of each panel represent the expected sizes of the SAP gene fragments.

CHAPTER 5

A Fourth Secreted Aspartyl Proteinase Gene (*SAP*4) and a *CARE*2 Repetitive Element are Located Upstream of the *SAP*1 Gene in *Candida albicans*.

Portions of this chapter have been previously published in:

Journal of Bacteriology

1994

Volume 176, No. 6

pp. 1702-1710

Shelley H. Miyasaki, Theodore C. White, and Nina Agabian.

Intercampus Program in Molecular Parasitology, Departments of Pharmaceutical Chemistry and Stomatology, and the Oral AIDS Center, University of California at San Francisco and Berkeley, San Francisco, California 94143

UCSF LIBRARY

Our studies identified 3 different *SAP* genes and further established that these were differentially expressed in the white and opaque switch phenotypes in the WO1 strain. Because transcription and expression of Saps 1 and 3 are strictly correlated with switch phenotype, the analysis of their expression provides us with a basis for beginning a study which addresses, at the molecular level, the transcriptional regulation of *SAP* gene expression and its relationship to phenotypic switching. Chapter 5 is the story of this investigation.

ABSTRACT

Candida albicans secreted aspartyl proteinases (Sap), products of the *SAP* genes, are presumed to act as virulence factors. In the *C. albicans* strain WO-1, the ability to secrete Sap1 is regulated with switch phenotype, another putative virulence factor. *Kpn*I restriction fragment length polymorphisms differentiate between several distinct *SAP1* alleles in laboratory and clinical strains. Both *SAP1* alleles from strain WO-1 along with their 5' and 3' flanking regions were cloned and sequenced as were both alleles from another strain, SS. The SS strain was chosen for further study because the *SAP1* genes from this strain were not expressed under any conditions. The 5' flanking regions were remarkably similar in all four of the sequenced alleles over approximately 1500 nt. RFLP analysis revealed a correlation between the expression of Sap1 and the presence of a 2.6 kb allele of the *SAP1* gene; which was present only in the WO1 strain. Together these characteristics suggested to us the possibility that the 2.6 allele may be the functional allele resulting in Sap 1 expression during switching. S1 analysis revealed that both alleles of WO-1 are transcribed. Characterization of the one allele from strain WO-1 identified a 284 nucleotide insertion flanked by 8 bp direct repeats that shows homology to the *CARE2* repetitive element and that is not present in the other alleles. Characterization of the *SAP1* alleles also identified a fourth *SAP* gene (*SAP4*) that includes an extended leader sequence. *SAP4* is positioned upstream, in

tandem to *SAP1*, in all strains tested and may encode another closely related Sap proteinase.

INTRODUCTION

Candida albicans is one of the most common causes of opportunistic infections which afflict immunocompromised patients. Factors which may contribute to the conversion of this normally commensal organism to a more virulent form include a morphological transformation between yeast and hyphal cells, as reviewed by Cutler (1991) and the ability of a *Candida* species to express alternative colony and cell morphologies, a phenomenon which has been referred to as phenotypic switching (reviewed by Soll (1992).) The phenotypic switching system of strain WO-1 has been characterized in detail (Slutsky et al. 1987 and Soll, Morrow and Srikantha 1993) and is easily manipulated, making it an important laboratory tool for examining the relationship between switch phenotype and the expression of virulence characteristics. In this strain, the cells switch reversibly and at high frequency between cellular morphologies, those which are predominantly spherical and smooth, and others which are larger, ovoid in shape and whose surface is rough in appearance. Each of these cell types gives rise to a distinct colony phenotype on solid media, appearing as white (W) and opaque (O), respectively. The switch phenotype further affects properties associated with virulence such as cell permeability, adhesiveness, drug susceptibility, and sensitivity to neutrophils and oxidants (reviewed by Soll, Morrow and Srikantha 1993). The

ability to switch between morphologic forms seems, therefore, to be a reflection of more pervasive changes in the spectrum of proteins expressed by each phenotype, a feature which is particularly interesting in the context of virulence as each distinct cell type may be able to respond differentially to the host.

The secretion of proteolytic enzymes by *C. albicans* was first noted by Staib in 1965 and since that time, a *C. albicans* secreted aspartyl proteinase (Sap) activity has been linked with virulence through indirect observation. It has been reported that Sap activity is directly proportional to the virulence of several different *Candida* species, the more virulent species such as *C. albicans* producing more Sap than *C. glabrata,* for example (Rüchel, Uhlemann and Boning 1983). Immunocytological studies have detected Sap antigens in tissues of infected patients and animals and Sap-deficient mutants of *C. albicans* are less able to adhere to host tissues and are less virulent in experimental animal models (Cutler 1991, Ghannoum and Abu Elteen 1986, Kwon-Chung et al. 1985, Macdonald and Odds 1980 and Rüchel et al. 1983). Sap activities purified from different *C. albicans* isolates also cleave a number of appropriate host substrates in vitro, including keratin, collagen, albumin, hemoglobin, IgA and secretory IgA (Hattori et al. 1981, Kaminishi et al. 1986, Rüchel 1981, Rüchel, Tegeler and Trost 1982, and Staib 1965). Thus, many possible roles for Sap in virulence have been proposed, including assisting in adhesion to and invasion of the host or in interfering directly with host defense mechanisms (Borg and Rüchel 1988). The ability to express a Sap activity was recently associated with

phenotypic switching in the white-opaque strain WO-1 (Morrow, Srikantha and Soll 1992), thereby providing a possible connection between two putative virulence factors.

Determining the specific role of Sap in pathogenesis has been difficult for several reasons. In particular, studies characterizing Sap activity have been inconsistent, reporting differences in pH optima, inhibition profiles and substrate specificity (Rüchel, Tegeler and Trost 1982). These differences suggested that more than a single type of Sap might exist, a perception which has recently been confirmed through molecular genetic studies which have identified several different aspartyl proteinase genes in *C. albicans*, *SAP1* by Hube et al. (1991), *SAP2* by Wright et al. (1992) and *SAP3* by White et al. (1993). The presence of multiple loci in the *C. albicans* genome which encode similar aspartyl proteinases suggests that much of the literature which reports the induction of Sap expression by various substrates and under different growth conditions may in fact reflect the differential expression of each of these loci or alleles within a given locus.

At this time, little is known concerning the differential expression of these three Sap proteinases under various conditions. Wright et. al. (1992) showed that *SAP1* and *SAP2* transcripts are produced by strain ATCC 10261 in media containing yeast extract and glucose supplemented with BSA, although they showed that the *SAP2* mRNA was more abundant. Morrow, Srikantha and Soll (1992) showed that the *SAP1* gene is expressed exclusively in O cells and not in

W or hyphal cells of strain WO-1. In addition, they showed that the *SAP1* gene is not expressed in budding or hyphal cells of strain 3153A or in cells of two clinical isolates. *SAP1* is expressed in O cells under a variety of conditions including defined media and serum induction (Morrow, Srikantha and Soll 1992). Recently, they have shown that transcription of *SAP1* and a second opaque-specific gene, *Op4*, is immediately inhibited by shifting the temperature from 25°C to 37°C, a shift which results in the mass conversion of O cells to W cells (Morrow et al. 1993). This inhibition of *SAP1* and *Op4* transcription can be reversed by lowering the temperature before the second semisynchronous cell cycle. Our Western blot analysis of culture supernatants, and subsequent N terminal protein sequencing (White, Miyasaki and Agabian 1993) confirmed that the Sap1 protein is O-specific and identified a second O-specific Sap protein, Sap3. This Western analysis also showed that Sap2 is the predominant protein expressed in strain 3153A, in strain SS, and in W cells of strain WO-1. These Western analyses were conducted in proteinase-inducing media to maximize the production of Sap proteinases. The correlation between the *SAP1* transcription results of Morrow, Srikantha and Soll (1992) and the Western blot results (White, Miyasaki and Agabian 1993) suggest that the pattern of differential proteinase expression is not altered by growth in these different media.

Our studies were initiated with the *SAP1* gene locus, the only *SAP* locus which had been identified when these studies began. We found a correlation between restriction fragment length polymorphisms (RFLP) in the *SAP1* alleles

and the O-specific expression of *SAP1*. Comparisons of the Southern hybridization patterns of WO-1and two other *C. albicans* strains revealed an RFLP pattern (designated 2.6) specific to WO-1 cells. The *SAP1* alleles and their flanking regions were cloned from WO-1 and from SS. These four alleles account for the major RFLP patterns of *SAP1* in laboratory and most clinical strains of *C. albicans*. S1 analysis demonstrated that both alleles of *SAP1* are transcribed in O cells. Analysis of the upstream flanking regions of the 2.6 allele identified a 284 bp insertion with homology to part of the *CARE2* element (14), bounded by an 8 bp repeat, suggesting that *CARE2* is indeed a transposable element. Sequence analysis of the upstream regions also identifed a fourth *SAP* gene (*SAP4*) positioned in tandem with the *SAP1* gene.

MATERIALS AND METHODS

Strains, media and plasmids

The *C. albicans* strains used were strains 3153a (ATCC 28367) and WO-1 (Slutsky et al. 1987), both generously provided by Dr. David Soll (University of Iowa), and strain SS (Blagoyevich 1979), provided by Dr. Remo Morelli (San Francisco State University). Cultures were maintained at 25°C on YEPD medium (10 g yeast extract, 20 g peptone and 20 g dextrose per liter) and transferred weekly. The WO-1 strain was maintained separately as both W and O phenotypes on YEPD plates containing phyloxine B (5 mg/ml). Cultures were grown in YCB-BSA medium (23.4 g yeast carbon base, 2 g yeast extract and 4 g bovine serum albumin per liter, pH 5.0 and filter sterilized) to induce proteinase

transcription. The *Escherichia coli* Bluescript plasmid vector (Stratagene Inc., La Jolla, Calif.) was used to subclone the *SAP1* alleles and was transformed into the *E. coli* strain XL1-blue for sequencing.

SAPAS probe

The DNA probe used for screening genomic libraries and for performing Southern analyses was a 655-bp internal fragment of the *SAP1* gene, amplified from strain SS genomic DNA by the polymerase chain reaction (Saiki et al. 1988). Oligonucleotide primers derived from the *SAP1* sequence published by Hube et al. (1991) were used to amplify the internal *SAP1* fragment. These primers were:

SM-1: 5'-CAA GCC ATC CCA GTT ACT TTA AA-3'

SM-2: 5'-CCA GAA TCT AAA AGA ACA TCG ATA TTA CC-3'

This probe contains both active site domains and extends from nucleotide 277 to nucleotide 932 of the Hube sequence (1991). This fragment will be referred to as "SAPAS" for *SAP1* active site probe.

Southern blot analysis

Southern analyses of restriction enzyme-digested genomic DNA from the three strains WO-1, SS and 3153a were performed as described by Sambrook, Fritsch and Maniatis (1989). Restriction enzymes used included *Eco*RI, *Bam*HI and *Kpn*I (New England Biolabs, Beverly, Massachusetts). We also performed Southern analyses using *Eco*RI- and *Kpn*I-digested phage DNA isolated from

genomic DNA libraries of strains WO-1 and SS as described below. All Southern blots were probed with the ^{32}P-labeled SAPAS probe.

Cloning and DNA sequence analysis of the *SAP1* alleles from WO-1 and SS strains

To isolate the alleles corresponding to the *SAP1* gene from both the WO-1 and SS strains, we screened libraries of genomic DNA fragments digested with *Sau*IIIA from both strains in the λ phage vector FIX II (Stratagene Inc., La Jolla California), and size-selected libraries of genomic DNA fragments digested with *Kpn*I and with *Bam*H1from the WO-1 strain in the plasmid vector pBluescript.

At high stringency, the SAPAS probe hybridizes exclusively to the *SAP1* alleles. Hybridization at high stringency (65°C in 3X SSC buffer (20X SSC: 3 M NaCl, 0.3 M sodium citrate, pH 7.0) with a final wash in 0.1X SSC) and low stringency (50°C in 3X SSC buffer with all washes in 3X SSC) yielded two sets of clones. Recombinant DNA was prepared from these clones. This DNA was digested separately with *Eco*RI and *Kpn*I, Southern blotted and hybridized with the SAPAS probe. Representative *Kpn*I and *Eco*RI fragments from each group were transferred to the Bluescript plasmid vector for restriction digest mapping and sequence analysis. The nucleotide sequence of the four alleles obtained from clones that hybridized at high stringency was determined by the dideoxy chain termination method as described by Sanger, Nicklen and Coulson (1977) with synthetic oligonucleotide primers. All other recombinant DNA procedures were carried out according to previously published protocols (Sambrook, Fritsch

and Maniatis 1989). Sequence analysis was performed using programs from the Molecular Biology Information Resource (MBIR, Baylor College of Medicine, Houston, Texas).

RNA isolation

C. albicans cells were grown in 5 ml of either YEPD or YCB-BSA and harvested at mid-log phase. Cells were washed in ice water and suspended in 2.5 ml of RNA lysis buffer (0.5 M NaCl-1% sodium dodecyl sulfate (SDS)-0.2 M Tris-Cl (pH. 8)-10 mM EDTA). An equal volume of phenol, and 1.5 volumes of glass beads was added and the entire mixture was vortexed at high speed for 2 min. The mixture was placed on ice for 2 min, and again vortexed for 2 min. After centrifugation at 11,950 x *g* for 10 min in a Sorvall SS-34 rotor at 4°C, the aqueous phases were collected and extracted with phenol. The mixture was again centrifuged (all remaining centrifugations and manipulations were done at room temperature) and the aqueous phase extracted with phenol-chloroform-isoamyl alcohol (25:24:1, vol/vol). After a final centrifugation for 5 min, the aqueous phase was collected and extracted a final time with chloroform-isoamyl alcohol (24:1, vol/vol). The aqueous phase was ethanol precipitated overnight at -20°C. Determinations of the A_{260} were made to assess RNA concentration.

S1 nuclease protection analyses

S1 nuclease protection analyses of allelic expression (1) were performed as described (Bernards et al. 1981), using 50 mg of total RNA. The RNA was pelleted and dried with 50,000 cpm of a 2-kb probe labeled at the 5' end with

^{32}P by using polynucleotide kinase. The probe consisted of an *Eco*RI fragment composed of the first 870 bp encoding the *SAP1* gene and approximately 1.2 kb of upstream sequence. This mixture was suspended in 20 ml S1 hybridization solution (80% formamide-0.4 M NaCl-1.0 mM EDTA-40 mM piperazine-N,N'-bis(2-ethane-sulfonic acid) (Pipes) buffer, pH 6.4) and hybridized while submerged overnight at 37°C. Then 300 ml of S1 reaction mix (20 U/ml S1 nuclease-20 mg/ml denatured salmon sperm DNA-4.5 mM zinc acetate-0.280 M NaCl-30 mM sodium acetate, pH 4.5) was added and the reaction mixture incubated at 37°C for 30 min. The reaction was terminated with the addition of 77 ml of S1 stop mix (50 mM EDTA-2.5 mM ammonium acetate with 20 mg *E. coli* tRNA), and the mixture was precipitated overnight in 2 volumes of 100% ethanol at -20°C. The reaction mixture was electrophoresed on a 7% polyacrylamide gel, dried, and autoradiographed.

RNA transcripts representing the 2.4 and 2.6 alleles of WO-1 were synthesized to provide positive controls for the S1 analysis. The pBluescript clones were linearized by digestion with *Hind*III. The linearized clones were used as templates in a T3 RNA polymerase reaction. The reaction mixture, containing 2.5 ml of 10X transcription buffer (400 mM Tris-Cl (pH 7.9)-60 mM $MgCl_2$-20 mM spermidine-100 mM dithiothreitol-0.5 M NaCl), 1 mg of linearized DNA template, 1 ml of RNAsin (40 U/l), 10 U of T3 RNA polymerase and H_2O to a total volume of 25 ml, was incubated at 37°C for 30 min. Then 75 ml of H_2O was added and the entire mixture was spun through a P-30 column (Biorad,

Richmond, California). RNA concentration was determined by A_{260} absorption. 5 ng of transcript was used in the control S1 reactions.

RESULTS

Our interest in proteinase virulence factors in *C. albicans* originated from the reported observation that an aspartyl proteinase was secreted primarily by O cells (and not W cells) in vivo at lesional sites in kidneys, liver and spleen of mice experimentally infected by WO-1 cells made by Ray, Payne and Morrow (1991). This is consistent with the results of Morrow, Srikantha and Soll (1992) in which *SAP1* is expressed exclusively in O cells in defined and serum induced media, and with our results (White, Miyasaki and Agabian 1993) in which *SAP1* and *SAP3* are expressed exclusively in O cells in proteinase inducing media. This suggested that *SAP1*, then considered the principal secreted aspartyl proteinase, might be associated with either switch phenotype or the expression of virulence characteristics in *C. albicans* infections. Accordingly, we initiated an analysis of aspartyl proteinase gene expression in the WO-1 switch phenotype and two unrelated strains, SS, originally isolated from a patient with systemic candidiasis and 3153a, a strain which possesses a different phenotypic switching system.

Restriction fragment length polymorphism mapping distinguishes three alleles of *SAP1*

Using the amino acid sequence of *SAP1* reported by Hube et al.(1991), a PCR derived probe (SAPAS) for the *SAP1* gene was synthesized as described

in Materials and Methods. Southern blots containing genomic DNA isolated from WO-1, 3153a and SS cells and digested with the restriction endonuclease *Kpn*I detected two Southern patterns specific to the SAPAS probe (Figure 1). The first hybridization pattern, common to both the SS and 3153a strains, consisted of *Kpn*I fragments of 2.8 kb, 2.4 kb and 0.4 kb; the second pattern was found only in the WO-1 strain and consisted of *Kpn*I fragments of 2.6 kb, 2.4 kb and 0.4 kb. Three distinct *SAP1* alleles could be identified from this Southern analysis by the presence or absence of *Kpn*I sites (Figure 2A). For convenience these three classes of alleles will be referred to by the size of their largest *Kpn*I fragments as the 2.4 allele from WO-1, SS and 3153a, the 2.6 allele from WO-1 and the 2.8 alleles from SS and 3153a. The same three alleles were consistently identified in clinical isolates of *C. albicans* (White, Miyasaki and Agabian, unpublished observations). In Figure 1, the differences in hybridization signals between the alleles is due to the length of the probe which hybridizes to each *KpnI* fragment (see Figure 2A). Thus, the 2.8 *KpnI* fragment hybridizes more intensely in Figure 1 because the SAPAS probe hybridizes to a longer region of the 2.8 *Kpn*I fragment than the 2.6 or 2.4 *Kpn*I fragments. The allelic pattern for strain WO-1 was the same for DNA isolated from either W or O cells (White, Miyasaki and Agabian, unpublished observations). Further restriction analysis using *Hin*d III, *Eco*RI, *Eco*RV, *Bgl*II, *Bam*HI, *Cla*I, *Kpn*I, and *Xmn*I (White, Miyasaki and Agabian, unpublished observations) established that all

fragments represent polymorphisms of the single copy *SAP*1 gene in *C. albicans*.

DNA sequence analysis of *SAP1* alleles and their flanking regions

To understand the molecular signals that might be operating in the differential transcription of the *SAP1* gene in each of the laboratory strains, genomic DNA libraries prepared from either WO-1 or SS strains were screened with the SAPAS probe at high stringency. Several clones encoding sequences homologous to SAPAS were isolated from each strain and characterized in detail. Clones which correspond to the 2.4 and 2.6 alleles of WO-1 and the 2.4 and 2.8 alleles of SS were each sequenced in their entirety. Tables 1 and 2 summarize the differences among these four alleles of *SAP1*. Within the coding regions of the *SAP1* alleles (Table 1), there are very few sequence differences. Of these, only the Tyr to Asn change at position 391 in the 2.6 allele of WO-1 represents a substitution that alters the nature of the amino acid residue; the remaining differences result mostly in silent or conservative changes. Approximately 1600 bases of the 2.6 allele nucleotide sequence upstream of the gene are shown in Figure 3. Table 2 compares this sequence with that upstream of the remaining 3 alleles as well as approximately 140 bases of 3' flanking sequence. We have numbered the nucleotides of the *SAP1* gene and flanking regions according to conventional standards, with the A of the initiation AUG numbered 1. For ease of comparison, we have included the numbers (where available) of these nucleotides as they correspond to the *SAP1*

sequence as published by Hube et al. (1991). It is remarkable that there are very few significant differences in the untranscribed 5' and 3' flanking regions between WO-1 and SS. The 3' flanking regions are essentially identical over 140 nucleotides. In the 5' flanking regions, with the exception of the insertion at position -1539 to -1255 of the 2.6 allele of WO-1, there are no nucleotide changes that are found uniquely in the WO-1 alleles which might encode cis-acting regulatory elements for *SAP1* expression in the WO-1 and not the SS strains in which *SAP1* expression has not been detected (White, Miyasaki and Agabian 1993). In the putative promoter regions, at position -166 to -174 there are a variable number of C residues, 9 and 11 in the WO-1 2.6 and 2.4 alleles respectively, and 23 (CT-21C) and 26 in the SS 2.8 and 2.4 alleles respectively. This duplication of a C tract 5' of the SS alleles is the only region proximal to the promoter that may be significant and of interest to functionally test for its effect upon transcription.

The 284 bp insert found only in the 2.6 allele of WO-1 is located 1,255 bp upstream of the start of translation. A Genbank search with this insert sequence revealed that over a span of 200 bp there was 94% homology with a repetitive *C. albicans* DNA sequence, *CARE2*, whose function is unknown (Lasker et al. 1992). The 284 bp insert upstream of the 2.6 allele is flanked at both ends by an 8 bp direct repeat of the presumed target sequence at the site of insertion.

Each of the four 5' flanking sequences has been used to search the transcription factors database (Ghosh 1992) using the FASTA program (Pearson

and Lipman 1988); no possible binding sites were identified for known transcription factors over the entire 5' flanking region. Figure 2B presents a limited restriction map and compilation of the major differences in the approximately 2.6 kb of sequence determined for each of the four alleles and their flanking regions.

Transcription of both 2.4 and 2.6 alleles of strain WO-1 in proteinase-inducing media

In comparing the upstream sequences of the four different *SAP1* alleles, the flanking region of the 2.6 allele is the most distinct. We asked therefore if the differences among the 5' flanking regions of the 2.6 and 2.4 alleles of strain WO-1 might affect the transcription of *SAP1* in these cells. Taking advantage of the several base pair differences in the coding regions of the 2.4 and 2.6 alleles of WO-1 (see Table 1), S1 analysis (diagramed in Figure 2C and 2D) was used to determine which *SAP*1 alleles of strain WO-1 are transcribed for *SAP1* gene expression .

In this experiment, DNA probes representing the 2.6 and 2.4 alleles of WO-1 were labeled with ^{32}P at an *Eco*RI site within the coding region of the *SAP1* gene (Figure 2C). These DNA probes extend into the 5' flanking regions of each allele. Each probe was hybridized to total RNA from O cells grown in the proteinase-inducing media, YCB-BSA. Hybridization of the probes to synthetic transcripts of the 2.6 and 2.4 alleles of WO-1 provided controls. The hybridization mixtures were digested with S1 nuclease using conditions in which

only DNA/RNA hybrids survive the digestion. The length of the largest resulting DNA fragment represents the distance from the 5' end of the RNA (the start of transcription) to the internal label (see Figure 2C). The result is seen in Figure 4 as the intense band of approximately 800 nucleotides in length in lanes 3 and 7. The bands near the top of the control lanes 1,2,5 and 6 most likely represent breathing of the DNA/RNA hybrid under S1 conditions and/or nuclease degradation of the control transcripts.

S1 nuclease can also attack and cleave at mismatches in the RNA/DNA hybrid (see Figure 2D), although the cleavage is at a low frequency. The length of these shorter DNA fragments (Figure 4, lanes 3 and 7) represents the distance from the mismatch to the internal label. From DNA sequence analysis (Table1), there are 3 mismatches between the 2.4 and 2.6 alleles of strain WO-1 upstream of the labeled EcoRI site (position 748). These include positions 210, 391 and 480, which are expected to generate labeled fragments of 538, 357 and 268 respectively, as the labeled EcoRI site is present at position 748. All three mismatches are present on one strand of RNA encoded by the other allele, suggesting that the shortest fragment should be the only fragment observed. However, S1 nuclease cleaves one base-pair mismatches at very low frequencies. Therefore, under ideal conditions, bands representing all three fragments, should be present. However, the sequence surrounding the mismatch and the S1 conditions influence the recognition of one base-pair mismatches. The expected 538 and 268 bands were not observed, even in the

control transcript lanes. However, the 357 band was observed (Figure 4 see arrows, lanes 2 and 5) when the 2.4 DNA probe was hybridized with the 2.6 T3 transcript and when the 2.6 DNA probe was hybridized with the 2.4 T3 transcript. The presence of bands of the same size in lanes 3 and 7 (Figure 4) demonstrates that mRNA from both the 2.4 and 2.6 alleles of WO-1 are present in total RNA, therefore, both alleles are transcribed. There is a slight size difference between bands seen in the total RNA lanes (lanes 3 and 7) and the control transcript lanes (lanes 2 and 5). This difference is a common problem and is due to the different amounts of RNA loaded into each lane (50 mg of total RNA vs. 50 ng of unlabeled control transcript) which results in a slight alteration in the mobility of the banding pattern. The difference can not be corrected by running carrier RNA in the control lanes (White, Miyasaki and Agabian, unpublished observations).

Identification and sequencing of the *SAP4* gene

During the subcloning and sequencing of fragments upstream of the *SAP1* gene, we identified a fourth *SAP* gene (see restriction map (Figure 2B) and complete sequence (Figure 5). A comparison of the predicted protein product of the *SAP4* gene with the other *SAP* gene products (Fig 6A) shows that the mature proteins are highly conserved in size and in sequence, especially surrounding the two portions of the active site. The hydrophobic signal sequences in the percursor proteins of the four gene products are also highly conserved (first 20 amino acids). The most divergent region of the four genes is

the region 3' of the signal sequences and immediately 5' of the start of the mature proteins, in which *SAP4* has a 24 amino acid insertion, which includes two additional Lys-Arg dipeptide residues which may be involved in processing of this extended precursor.

A phylogenetic tree (Figure 6B) shows that Sap1 and Sap2 are the most closely related Sap proteins, and that the *SAP4* gene product is the most divergent Sap identified to date. Southern analysis shows that the *SAP4* gene is present upstream of the *SAP1* gene in all laboratory and clinical strains tested (White, Miyasaki and Agabian, unpublished observations). To date, Northern analysis has not detected the expression of *SAP4* in any cell line or switch system under any of a variety of growth conditions (White, Miyasaki and Agabian, unpublished observations).

DISCUSSION

Restriction endonuclease and DNA sequence analysis of the three *C. albicans* strains tested here revealed an RFLP pattern which differentiated among three different allelic forms of *SAP1*. The two allelic forms, 2.4 and 2.8, differ because of a single nucleotide change which results in the presence or absence of a *Kpn*I site within the gene. The third allelic form, 2.6, is the result of a transposable element insertion 1255 nucleotides upstream of the gene.

The observation that WO-1 cells have an unique *SAP1* allele, 2.6, coupled with the demonstration that *SAP1* is uniquely expressed in O cells which contain this allele (Morrow, Srikantha and Soll 1992 and White, Miyasaki and Agabian 1993), led us to explore the possibility that the pattern of Sap1 expression in WO-1 cells might be related specifically to a feature of the 2.6 allele or its flanking regions. While this hypothesis was proven incorrect, the cloning, sequencing and comparison of these four *SAP1* alleles and their flanking regions from the WO-1 and SS strains of *C. albicans* has resulted in many interesting observations.

The regions encoding the *SAP1* gene alleles and their 5' and 3' flanking regions are highly homologous. The Tyr to Asn change at position 391 of the 2.6 allele aside, the structural gene sequences are essentially identical except for minor conservative or silent mutations in sequence. The sequence similarity of all four coding regions of the *SAP1* alleles establishes that these are each true alleles of *SAP1* and not pseudogenes or other closely related aspartyl

proteinases. Despite the high degree of homology between these alleles, the DNA sequence information obtained allowed us to design probes which differentiate *SAP1* alleles directly by S1 endonuclease protection analysis. Using these probes, we were able to show that both the 2.6 and 2.4 alleles of *SAP1* are transcribed in O cells grown in proteinase inducing media.

Surveying the sequences 5' of the ATG for approximately -200 nucleotides, there are several potential TATAA-like sequence elements which may serve as promoters (Figure 3). Nothing is known about spacing or sequence requirements of functional TATAA elements in *C. albicans.* However, in the related yeast *Saccharomyces cerevisiae,* TATAA elements in the regulatory regions of genes are commonly found anywhere between 40 and 120 bp upstream of the mRNA initiation site. We note two potential TATAA-like elements at -92 and -71. Intergenic sequences in the *C. albicans* genome are A-T rich, thus identification of putative promoter sequences requires functional analysis. In the RNA 5' to the AUG, the *SAP1* sequence 5' AAACAAUGUUU 3' shares some homology with the consensus *Saccharomyces cerevisiae* translation initiator sequence, 5' -A/YAA/U-AAAUGUCU- 3', where Y is C or U. Two small differences, a C for the A or U at position -3 and a U replacing the C at +5, should not be enough to cause a major alteration in translational efficiency, since studies of mutational analyses performed both 5' and 3' to the AUG of the *HIS4* gene in *S. cerevisiae* showed no more than a two fold decrease caused by the change in context (Cigan, Pabich and Donahue 1988).

Both the 2.6 and 2.4 alleles are transcribed in WO-1, while no *SAP1* expression is found in strain SS (White, Miyasaki and Agabian 1993). Since the differences between the WO-1 alleles and the SS alleles are overlapping, we surmise that differences in the upstream flanking regions are not important in the differential transcription of *SAP1* in these two strains. The only striking difference in sequence is found at position -166 where each of the alleles contains a polyC tract varying from 9-26 nucleotides in length. Those in the SS strain are over twice the length of those in WO-1 and represent the only candidate region where any significant differences in potential cis-acting regulatory sequences might exist.

Approximately 1.1 kb of DNA from a different strain of *C. albicans* has been sequenced and designated *CARE2* (*C. albicans* repetitive element). Regions of this sequence appear to be mobile, based on alternating genomic hybridization patterns of the element in different isolates (Lasker et al. 1992). Analysis of the DNA sequence upstream of the 2.6 allele revealed a 284 bp insertion at position -1539 to -1255. The center of this 284 bp sequence is highly homologous to a region of the 1.1 kb of DNA characterized as *CARE2*. However, large regions at the ends of the 284 bp sequence have diverged from the *CARE2* sequence. The 284 bp sequence is flanked by 8 bp direct repeats. The 8 nucleotides of direct repeats are present at the same position in each of the other alleles sequenced, indicating that this is a typical duplication of a target sequence characteristic of transposable elements. A sequence

arrangement of several hundred nucleotides bounded by short direct repeats is reminiscent of delta insertion elements in *S. cerevisiae*. The delta elements of *Saccharomyces* are found at both ends of the TY1 transposons; during transposition, when a TY1 element is excised by recombination, a delta element is often left behind, bounded by direct repeats of the target sequence (Roeder and Fink 1983). Based on these observations, it would appear that the 284 bp sequence found upstream of the 2.6 allele represents an analogous solo delta-type sequence in *C. albicans*. In addition to strain WO-1, three clinical isolates has been found to contain the 2.6 allele polymorphism (White, Miyasaki and Agabian 1993). In one isolate, oligonucleotides from the 284 bp *CARE2* sequence hybridizes with the 2.6 kb *Kpn*I fragment from this clinical strain, indicating that a similar insertion occurs upstream of *SAP1* in an independent *C. albicans* isolate (White, Miyasaki and Agabian, unpublished observations). The occurrence of this insertion at the site of these virulence factors raises the possibility that, like other transposable elements, a *CARE2*-like element might contribute to the regulation of expression of *SAP1*.

Assuming that the alteration in poly C tract at position -166 to -174 is not involved in the differential transcription of *SAP1*, what then can we surmise about the potential regulation of *SAP1* expression in WO-1, 3153a and SS strains? Since the 5' and 3' flanking regions are essentially identical in two different *C. albicans* isolates (WO-1 and SS), one of which expresses the gene (WO-1; Morrow, Srikantha, Soll, 1992, White, Miyasaki, Agabian, 1993) and the

other which does not (White, Miyasaki, Agabian, 1993), one must conclude that there are different trans-acting elements provided by WO-1 and SS which interact with similar cis-acting elements upstream of *SAP1*. These could either be transcriptional activators, being expressed only in O cells, or conversely transcriptional repressors expressed in W cells and in other strains. In addition to the WO switching system, there are at least three different switching systems described for *C. albicans*, reviewed by Soll (1992). One interpretation of the *SAP1* expression studies (Morrow, Srikantha and Soll 1992 and White, Miyasaki and Agabian 1993) is that SS and 3153a cells remained in a switch phase which was not permissive for *SAP1* expression, similar to the W cell phenotype. Presumably, if a phenotypic switch could be induced in 3153a and SS strains under comparable conditions, then the *SAP1* alleles would be transcribed.

The tandem linkage of *SAP1* and *SAP4* genes may be the result of a tandem duplication of a *SAP* gene on that chromosome. Close inspection of the four protein sequences in Figure 6A and the phylogenetic tree (Figure 6B) shows that *SAP4* is the most divergent of the four proteins, suggesting that *SAP4* is not a recent duplication of *SAP1*.

Recently, the existance of another *SAP* gene was inferred from a hybridization signal to chromosome 3 on a chromosome blot (Magee et al. 1993). Signals on chromosomes 6 and R were attributed to the presence of *SAP1* and *SAP2* on their respective chromosomes. The presence of *SAP4* linked in tandem with *SAP1* raises the possibility that still other *SAP* genes are

linked in tandem with the already identified *SAP1, 2,* and*3* genes. Our preliminary hybridizations and sequencing of cloned flanking regions have not identified any further *SAP* genes. However, at least one more *SAP* gene remains to be identified - the gene which encodes a protein containing the N terminal sequence GTVQTSLINE which was identified by Morrison et. al. (1993).

ACKNOWLEDGMENTS

We thank Dr. David Soll, University of Iowa, for many helpful discussions and for providing strains WO-1 and 3153a; Dr. Remo Morelli, San Francisco State University, for providing strain SS; our colleagues in the laboratory for many helpful comments; and Ms. Evangeline Leash for editing the manuscript.

This work was supported in part by NIH #1RO1AI-33317, NIH # PO1-DE-O7946, and a John D. and Catherine T. MacArthur Foundation Award # 84-85 to N.A., University of California UniversityWide AIDS Research Program grant from the State of California #K92-SF-011 to T.C.W., and an NIH Dentist-Scientist Award (5K15DE00334) to S.H.M.

REFERENCES - CHAPTER 5

1. Berk, A. J., and P. A. Sharp. 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease digested hybrids. *Cell* 12:721-32.
2. Bernards, A., L. H. T. van der Ploeg, A. C. C. Frosch, P. Borst, J. C. Boothroyd, S. Coleman, and G. A. M. Cross. 1981. Activation of trypanosome surface glycoprotein genes involves a duplication-transposition leading to an altered 3' end. *Cell* 27:497-505.
3. Blagoyevich, P. G. 1979. Master's thesis. San Francisco State University.
4. Borg, M., and R. Rüchel. 1988. Expression of extracellular acid proteinase by proteolytic *Candida* spp. during experimental infection of oral mucosa. *Infect. Immun.* 56:626-31.
5. Cigan, A. M., Pabich, E. K., and Donahue, T. F. 1988. Mutational analysis of the *HIS4* translational initiator region in *Saccharomyces cerevisae*. *Mol. Cell. Biol.* 8:2964-75.
6. Cutler, J. E. 1991. Putative virulence factors of *Candida albicans*. Ann. *Rev. Microbiol.* 45:187-218.
7. Feng, D. F. and R. F. Doolittle. 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25:351-60.
8. Ghannoum, M., and K. Abu Elteen. 1986. Correlative relationship between proteinase production, adherence and pathogenicity of various strains of *Candida albicans*. *Infect. Immun.* 24:407-13.

9. Ghosh, D. 1992. TFD the transcription factor database. *Nucleic Acids Res.* 20(Supplement):2091-3.

10. Hattori, M., K. Yoshiura, M. Negi, and H. Ogawa. 1981. Keratinolytic proteinase produced by *Candida albicans*. *Sabouraudia* 22:175-83.

11. Hube, B., C. J. Turver, F. C. Odds, H. Eiffert, G. J. Boulnois, H. Kochel, and R. Rüchel. 1991. Sequence of the *Candida albicans* gene encoding the secretory aspartate proteinase. *Med. Vet. Mycol.* 29:129-32.

12. Kaminishi, H., Y. Hagihara, S. Hayashi, and T. Cho. 1986. Isolation and characteristics of collagenolytic enzyme produced by *Candida albicans*. *Infect. Immun.* 53:312-6.

13. Kwon-Chung, D. J., D. Lehman, C. Good, and P. T. Magee. 1985. Genetic evidence for role of extracellular proteinase in virulence of *Candida albicans*. *Infect. Immun.* 49:571-5.

14. Lasker, B. A., L. S. Page, T. J. Lott, and G. S. Kobayashi. 1992. Isolation, characterization and sequencing of *Candida albicans* repetitive element 2. *Gene* 116:51-7.

15. Macdonald, F., and F. C. Odds. 1980. Inducible proteinase of *Candida albicans* in diagnostic serology and in the pathogenesis of systemic candidosis. *J. Med. Microbiol.* 13:423-35.

16. Magee, B.B., B. Hube, R.J. Wright, P.J. Sullivan and P.T. Magee. 1993. The genes encoding the secreted aspartyl proteinases of *Candida*

albicans constitute a family with at least three members. *Infect. Immun.* 61:3240-3.

17. Morrison, C. J., S. F. Hurst, S. L. Bragg, R. J. Kuykendall, H. Diaz, J. Pohl and E. Reiss. 1993. Heterogeneity of the Purified Extracellular Aspartyl Proteinase from *Candida albicans:* Characterization with Monoclonal Antibodies and N-Terminal Amino Acid Sequence Analysis. *Infect. Immun.* 61: 2030-6.
18. Morrow, B., T. Srikantha, and D. R. Soll. 1992. Transcription of the gene for a pepsinogen, *PEP*1, is regulated by white-opaque switching in *Candida albicans*. *Mol. Cell. Biol.* 12:2997-3005.
19. Morrow, B., T. Srikantha, J. Anderson and D. R. Soll. 1993. Coordinate regulation of two opaque-phase-specific genes during white-opaque switching in *Candida albicans*. *Infect. Immun.* 61:1823-8.
20. Pearson, W.R., and D.J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85:2444-8.
21. Ray, T. L., C. D. Payne, and B. J. Morrow. 1991. *Candida albicans* acid proteinase: Characterization and role in candidiasis. *In* B. M. Dunn (ed.), Structure and function of the aspartic proteinases. New York. Plenum Press.
22. Roeder, G. S., and G. R. Fink. 1983. Transposable elements in yeast, p. 299-328. *In* J. A. Shapiro (ed.), Mobile genetic elements, New York. Academic Press.

23. Rüchel, R. 1981. Properties of a purified proteinase from the yeast *Candida albicans*. *Biochim. Biophys. Acta* 659:99-113.

24. Rüchel, R., R. Tegeler, and M. Trost. 1982. A comparison of secretory proteinases from different strains of *Candida albicans*. *Sabouraudia* 20:233-44.

25. Rüchel, R., K. Uhlemann and B. Boning. 1983. Secretion of acid proteinases by different species of the genus Candida. *Zbl. Bakt. Hyg., I. Abt. Orig. A.* 255:537-48.

26. Rüchel, R., F. DeBernardis, T. L. Ray, P. A. Sullivan, and G. T. Cole. 1992. *Candida* acid proteinases. *Med. Vet. Mycol.* 30 (Supplement 1):123-32.

27. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Sharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer directed enzymatic amplification of DNA with thermostable DNA polymerase. *Science* 239:487-94.

28. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor, N. Y., Cold Spring Harbor Laboratory.

29. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-7.

30. Slutsky, B., M. Staebell, J. Anderson, L. Risen, M. Pfaller, and D. Soll. 1987. "White-opaque transition," a second high-frequency switching system in *Candida albicans*. *J. Bacteriol.* 169:198-7.

31. Soll, D. R. 1992. High-frequency switching in *Candida albicans*. *Clin. Microbiol. Rev.* 5:183-203.

32. Soll, D.R., Morrow, B. and Srikantha, T. 1993. High frequency phenotypic switching in *Candida albicans*. *Trends in Genetics* 9:61-5.

33. Staib, F. 1965. Serum-proteins as nitrogen source for yeast-like fungi. *Sabouraudia* 4:187-193.

34. White, T. C., S. H. Miyasaki, and N. Agabian Unpublished observations.

35. White, T. C., S. H. Miyasaki, and N. Agabian. 1993. Three distinct secreted aspartyl proteinases in *Candida albicans*. *J. Bacteriol.* 175:6126-33.

36. Wright, R. J., A. Carne, A. D. Hieber, I. L. Lamont, G. W. Emerson, and P. A. Sullivan. 1992. A second gene for a secreted aspartate proteinase in *Candida albicans*. *J. Bacteriol.* 174:7848-53.

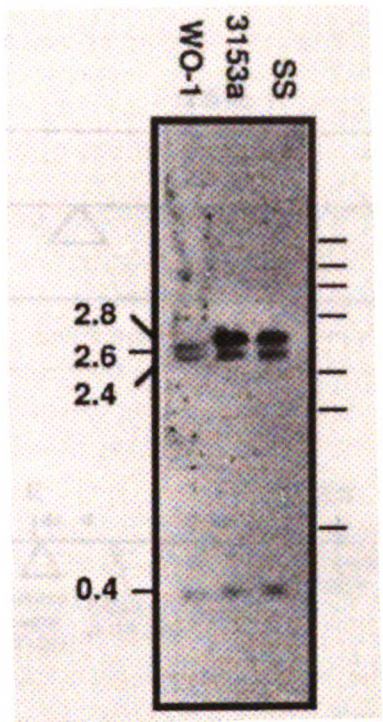


Figure 1

RFLP mapping distinguishes three alleles in three laboratory strains. Southern analysis of *Kpn*I-digested genomic DNA from SS, 3153a and WO-1 strains hybridized with the ^{32}P-labeled SAPAS probe. 10 µg of genomic DNA from the three strains (3153a, SS and WO-1) was digested with *Kpn*I, electrophoresed on a 0.8% agarose gel and blotted to nitrocellulose. The blot was hybridized at high stringency. Genomic DNA used in each lane is indicated above the lane. Numbers to the left of the panel indicate sizes of the genomic fragments. λ size markers are indicated by lines on the right of the panel. The size of each marker (from the top) is 23.1, 9.4, 6.7, 4.3, 2.3, 2.0, and 0.5 kb. The intensity of hybridization of each band is related to the length of the fragment that hybridizes to the SAPAS probe (see Fig. 2A).

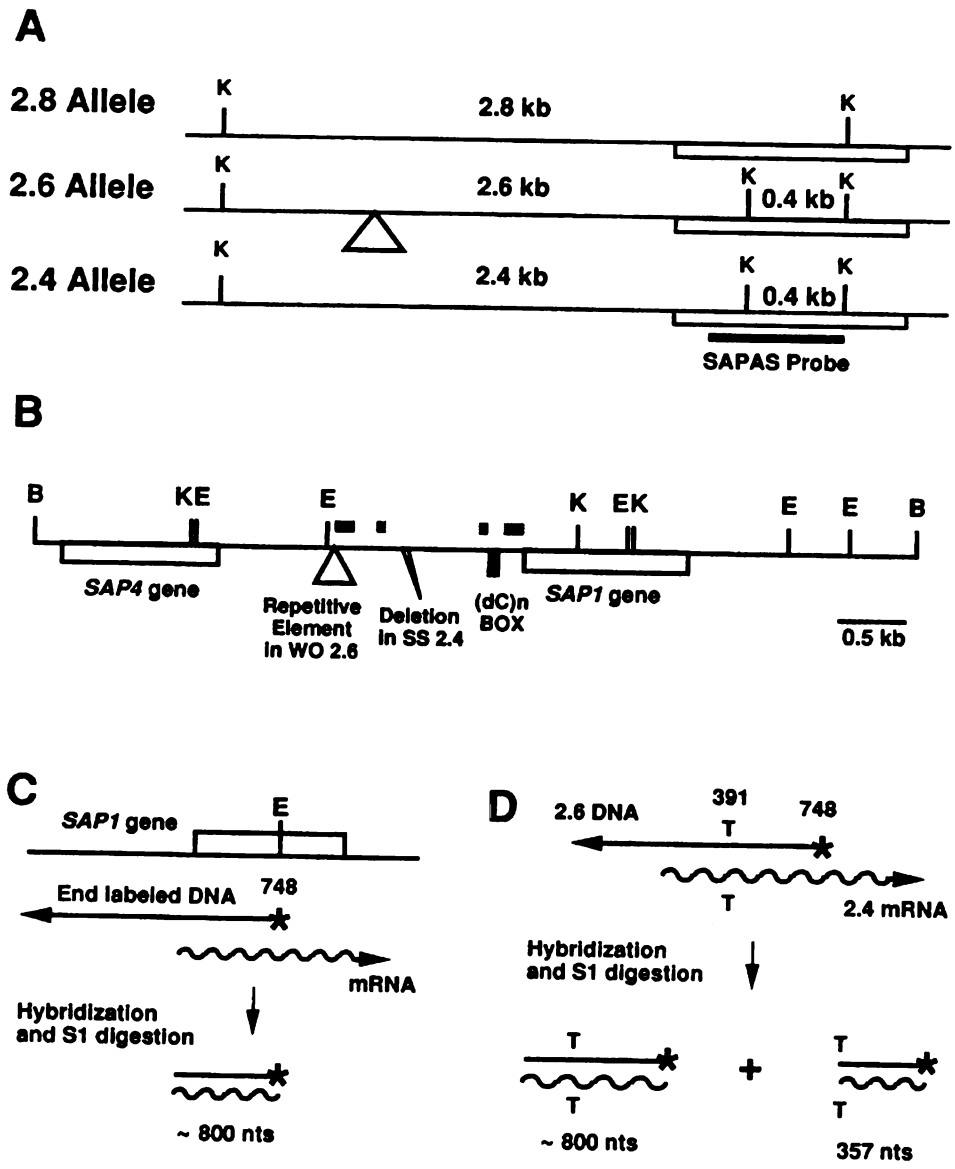

A
2.8 Allele
2.6 Allele
2.4 Allele
K
2.8 kb
2.6 kb
2.4 kb
0.4 kb
SAPAS Probe
B
B
KE
E
K
EK
E
E
B
SAP4 gene
Repetitive Element in WO 2.6
Deletion in SS 2.4
(dC)n BOX
SAP1 gene
0.5 kb
C
SAP1 gene
E
End labeled DNA
748
mRNA
Hybridization and S1 digestion
~ 800 nts
D
2.6 DNA
391
T
748
2.4 mRNA
Hybridization and S1 digestion
~ 800 nts
357 nts

Figure 2

Figure 2 Legend Restriction maps and summary of major allelic differences.

A. A schematic drawing of the *Kpn*I (K) restriction fragments which are used to identify the three *SAP1* alleles. The position of the SAPAS probe relative to these fragments is shown as a bar below the three allele maps. The 2.6 allele contains an insert located approximately 1255 nucleotides upstream of the gene.

B. Major insertions and deletions in the *SAP1* alleles and a limited restriction map of the region which includes *SAP1* and *SAP4* (to scale). The bar labeled "(dC)n box" shows the position of the series of deoxycytidines, which vary in number between alleles. DNA secondary structures in the 5' flanking regions of *SAP1* include two short inverted repeats (-1024 to -1044 and -234 to -255) and two large inverted repeats (-13 to -147 and -1218 to -1590). All inverted repeat pairs are shown by shaded boxes. B = *Bam*HI, E = *Eco*RI, K = *Kpn*I.

C. Schematic outline of the S1 analysis. Cloned DNA is digested with *Eco*RI (E in figure at position 748 in the *SAP1* gene) and 5' end labeled with polynucleotide kinase. Total RNA, which contains *SAP1* mRNA is hybridized to the labeled DNA. After hybridization, the DNA/RNA hybrids are digested with S1 nucelase which degrades single stranded DNA and RNA, leaving the DNA/RNA hybrid. The resulting end-labeled DNA fragment (~ 800 nucleotides in length), which is detected on denaturing polyacrylamide gels, represents the distance from the internal label at the EcoRI site to the start of transcription (the 5' end of the mRNA). Arrowheads on DNA and RNA indicate the 5' to 3' direction of the strand.

D. Schematic drawing of an S1 experiment containing mismatches. End labeled DNA from a clone of the 2.6 allele contains a T at position 391 (based on positions in Table 1). RNA from allele 2.4 contains a T at position 391. When hybridized together, the DNA/RNA hybrid contains a mismatch at position 391 since the two T residues can not base pair. This mismatch is detected by S1, although at a low frequency. Therefore, most of the DNA/RNA hybrids are ~ 800 nts. as expected. However, a small proportion will be cleaved at the T/T mismatch by S1 nuclease, resulting in a labeled fragment of 357 nts. Similar diagrams can be constructed for DNA from the 2.4 allele hybridized with RNA from the 2.6 allele and for each of the nucleotide positions in which the 2.4 and 2.6 alleles differ in sequence (see Table 1).

TABLE 1. Differences between coding regions in alleles from three strains of *C. albicans*

Position[a]	Nucleotide					Coding change
	10231	WO-1 allele	WO-1 allele	SS 2.8 allele	SS 2.4 allele	
21 (147)	C	C	C	T	C	Silent
24 (150)	C	T	T	T	T	Silent
157 (283)	A	C	C	A	C	Ile/Leu
159 (285)	C	C	C	T	C	
210 (336)	T	T	A	T	T	Silent
383 (509)	G	C	C	C	C	Ser/Thr
384 (510)	C	C	C	T	C	
390 (516)	T	C	C	C	C	Silent
391 (517)	T	A	T	T	T	Tyr/Asn
444 (570)	T	C	C	T	T	Silent
480 (606)	C	T	C	C	C	
663 (789)	C	T	T	T	T	Silent
882 (1008)	T	C	C	C	T	Silent
956 (1082)	C	C	T	C	C	Ala/Val
981 (1107)	A	G	G	G	G	Silent

[a]Numbering system based on sequence of WO-1 allele; corresponds to numbering in Figure 3. Numbers in parentheses represent the corresponding nucleotides in the sequence reported by Hube et al. for ATCC 10231 (Hube, 1991).

TABLE 2. Sequence differences between flanking regions of alleles from three strains of *C. albicans*

Position[a]	10231	WO-1 allele	WO-1 allele	SS 2.8 allele	SS 2.4 allele
5' - Flanking region					
-1562		A	TT	T	TT
-1553		C	G	G	G
-1539 to -1255[b]		Insert	—	—	—
-1251		T	C	C	C
-1212		T	C	C	T
-1107		G	A	A	A
-1102		C	A	A	A
-858 to --875[b]		—	—	—	Deleted
-588		A	A	G	A
-489		A	G	G	G
-166 to -174[c]		9C	11C	CT-21C	26C
-133		T	T	G	T
-110 (17)	G	G	G	G	A
-23 (104)	T	C	C	C	C
3'- Flanking region					
1270 (1396)	T	T	T	C	C
1272 (1398)	T	C	C	C	C

[a]Numbering system is based on sequence of WO-1 2.6 allele and corresponds to numbering in Figure 3. Numbers in parentheses represent the corresponding nucleotides in the sequence reported by Hube et al. for strain 10231 (Hube, 1991). The sequence of strain 10231 as reported by Hube et al. (Hube, 1991) extends only to position -126.

[b]Dashes represent allelic sequences which do not contain insertions or deletions.

[c]Values represent poly (C) tracts of the indicated lengths. CT-21C represents a CT dinucleotide followed by 21 C nucleotides.

Figure 3. Sequence of the 5' flanking region of the WO-1 2.6 allele. The sequence ends at the ATG (in large letters) at the start of the coding region. The underlined deoxycytidines present at positions -166 to -174 vary in number among the different alleles. The underlined sequence beginning at positions -857 to -875 is deleted only from the SS 2.4 allele. The 284-bp insert present in only the WO-1 2.6 allele is underlined and spans positions -1568 to -1255. The double lines above the sequence indicate the 8-bp direct repeats (ATATTTTG) which flank the 284-bp insert. The completed sequences for all four alleles of *SAP*1 have been deposited in GenBank under accession numbers L12449 through L12452.

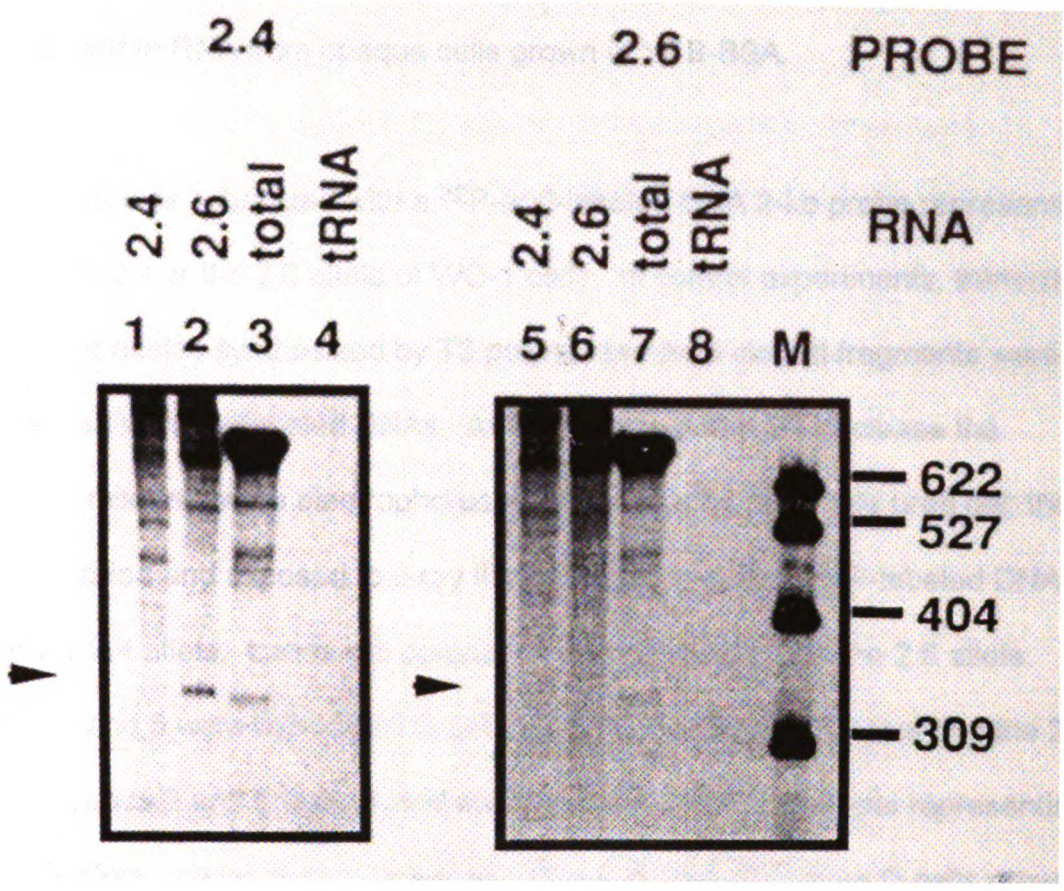
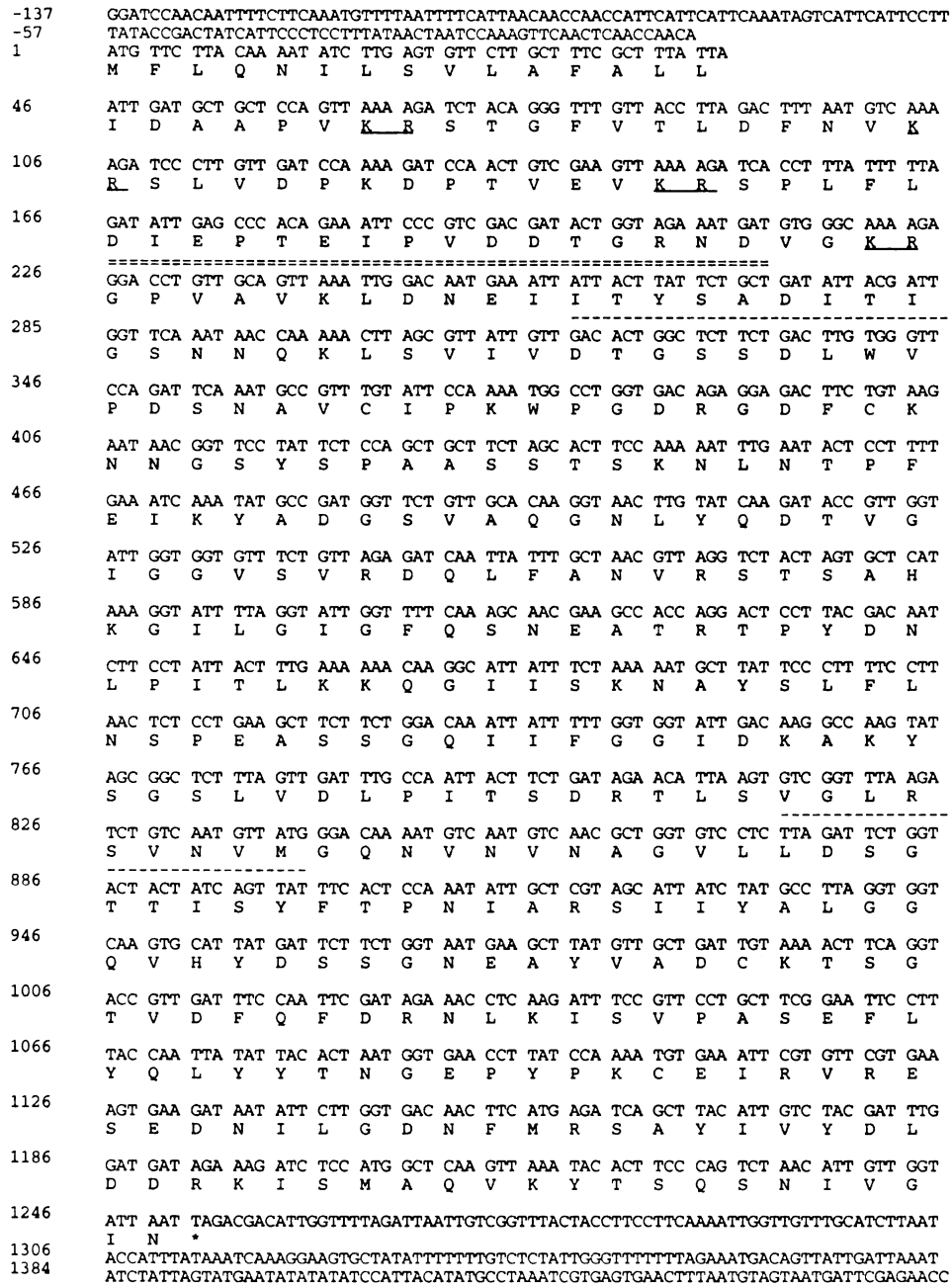


Figure 4

Figure 4 Legend Both the 2.4 and the 2.6 *SAP1* alleles of strain WO-1 are transcribed in RNA from opaque cells grown in YCB-BSA.

Total RNA was hybridized with a ^{32}P-end-labeled DNA 2-kb probe representing either the 2.4 or the 2.6 allele of WO-1 cells. In control experiments, transcripts of the two alleles synthesized by T3 polymerase from cloned fragments were hybridized with the labeled DNAs. After incubation with S1 nuclease the reaction mixtures were electrophoresed on a 7% polyacrylamide urea gel; the gel was dried and exposed to x-ray film. Lanes 1-4 contain ^{32}P-labeled DNA from the 2.4 allele. Lanes 5-8 contain ^{32}P-labeled DNA from the 2.6 allele. Lanes 1 and 5 were hybridized to unlabeled T3 transcripts representing the 2.4 allele. Lanes 2 and 6 were hybridized to unlabeled T3 transcripts representing the 2.6 allele. Lanes 3 and 7 were hybridized to total RNA from O cells grown in YCB-BSA. Lanes 4 and 8 were hybridized to *E. coli* tRNA as a control. Lane M represents pBR322 marker DNA digested with *Msp*I and labeled with ^{32}P; the sizes of the marker bands (in bp) are listed on the right. Arrows point out the mismatches at position 357, indicating that both alleles are transcribed in total RNA from O cells grown in YCB-BSA. The slight size difference between bands in the total RNA lanes and the control transcript lanes is due to the different amounts of RNA loaded into each lane (50 μg of total RNA vs. 50 ng of unlabeled control transcript).



-137 GGATCCAACAATTTTCTTCAAATGTTTTAATTTTCATTAACAACCAACCATTCATTCATTCAAATAGTCATTCATTCCTT
-57 TATACCGACTATCATTCCCTCCTTTATAACTAATCCAAAGTTCAACTCAACCAACA
1 ATG TTC TTA CAA AAT ATC TTG AGT GTT CTT GCT TTC GCT TTA TTA
M F L Q N I L S V L A F A L L

46 ATT GAT GCT GCT CCA GTT AAA AGA TCT ACA GGG TTT GTT ACC TTA GAC TTT AAT GTC AAA
I D A A P V K R S T G F V T L D F N V K

106 AGA TCC CTT GTT GAT CCA AAA GAT CCA ACT GTC GAA GTT AAA AGA TCA CCT TTA TTT TTA
R S L V D P K D P T V E V K R S P L F L

166 GAT ATT GAG CCC ACA GAA ATT CCC GTC GAC GAT ACT GGT AGA AAT GAT GTG GGC AAA AGA
D I E P T E I P V D D T G R N D V G K R
==========

226 GGA CCT GTT GCA GTT AAA TTG GAC AAT GAA ATT ATT ACT TAT TCT GCT GAT ATT ACG ATT
G P V A V K L D N E I I T Y S A D I T I
----

285 GGT TCA AAT AAC CAA AAA CTT AGC GTT ATT GTT GAC ACT GGC TCT TCT GAC TTG TGG GTT
G S N N Q K L S V I V D T G S S D L W V

346 CCA GAT TCA AAT GCC GTT TGT ATT CCA AAA TGG CCT GGT GAC AGA GGA GAC TTC TGT AAG
P D S N A V C I P K W P G D R G D F C K

406 AAT AAC GGT TCC TAT TCT CCA GCT GCT TCT AGC ACT TCC AAA AAT TTG AAT ACT CCT TTT
N N G S Y S P A A S S T S K N L N T P F

466 GAA ATC AAA TAT GCC GAT GGT TCT GTT GCA CAA GGT AAC TTG TAT CAA GAT ACC GTT GGT
E I K Y A D G S V A Q G N L Y Q D T V G

526 ATT GGT GGT GTT TCT GTT AGA GAT CAA TTA TTT GCT AAC GTT AGG TCT ACT AGT GCT CAT
I G G V S V R D Q L F A N V R S T S A H

586 AAA GGT ATT TTA GGT ATT GGT TTT CAA AGC AAC GAA GCC ACC AGG ACT CCT TAC GAC AAT
K G I L G I G F Q S N E A T R T P Y D N

646 CTT CCT ATT ACT TTG AAA AAA CAA GGC ATT ATT TCT AAA AAT GCT TAT TCC CTT TTC CTT
L P I T L K K Q G I I S K N A Y S L F L

706 AAC TCT CCT GAA GCT TCT TCT GGA CAA ATT ATT TTT GGT GGT ATT GAC AAG GCC AAG TAT
N S P E A S S G Q I I F G G I D K A K Y

766 AGC GGC TCT TTA GTT GAT TTG CCA ATT ACT TCT GAT AGA ACA TTA AGT GTC GGT TTA AGA
S G S L V D L P I T S D R T L S V G L R
----

826 TCT GTC AAT GTT ATG GGA CAA AAT GTC AAT GTC AAC GCT GGT GTC CTC TTA GAT TCT GGT
S V N V M G Q N V N V N A G V L L D S G
----

886 ACT ACT ATC AGT TAT TTC ACT CCA AAT ATT GCT CGT AGC ATT ATC TAT GCC TTA GGT GGT
T T I S Y F T P N I A R S I I Y A L G G

946 CAA GTG CAT TAT GAT TCT TCT GGT AAT GAA GCT TAT GTT GCT GAT TGT AAA ACT TCA GGT
Q V H Y D S S G N E A Y V A D C K T S G

1006 ACC GTT GAT TTC CAA TTC GAT AGA AAC CTC AAG ATT TCC GTT CCT GCT TCG GAA TTC CTT
T V D F Q F D R N L K I S V P A S E F L

1066 TAC CAA TTA TAT TAC ACT AAT GGT GAA CCT TAT CCA AAA TGT GAA ATT CGT GTT CGT GAA
Y Q L Y Y T N G E P Y P K C E I R V R E

1126 AGT GAA GAT AAT ATT CTT GGT GAC AAC TTC ATG AGA TCA GCT TAC ATT GTC TAC GAT TTG
S E D N I L G D N F M R S A Y I V Y D L

1186 GAT GAT AGA AAG ATC TCC ATG GCT CAA GTT AAA TAC ACT TCC CAG TCT AAC ATT GTT GGT
D D R K I S M A Q V K Y T S Q S N I V G

1246 ATT AAT TAGACGACATTGGTTTTAGATTAATTGTCGGTTTACTACCTTCCTTCAAAATTGGTTGTTTGCATCTTAAT
I N *

1306 ACCATTTATAAATCAAAGGAAGTGCTATATTTTTTTGTCTCTATTGGGTTTTTTTAGAAATGACAGTTATTGATTAAAT
1384 ATCTATTAGTATGAATATATATATCCATTACATATGCCTAAATCGTGAGTGAACTTTAATGTAGTAATGATTCGAGAACC

Figure 5

Figure 5 Legend Sequence of the *SAP4* gene and deduced amino acid sequence of the protein.

The sequence begins at a *Bam*HI site shown in Figure 2B. The putative N terminus of the mature protein, which is highly conserved with the N termini of the other *SAP* proteins has a double line above the sequence. The two sections of the active site of the proteinase have a single line above the sequence. The four Lys-Arg dipeptides in the precursor peptide are underlined. The entire sequence is 1,618 nucleotides in length. The GenBank accession number is L25388.

A

Precursor Peptides:

```
sap1  MFLKNIFIALAIALLVDASP..AKRSPGFVTLDFDVIK.....................T....PV.NATG.QEG....K.VKR
sap2  -----------------T-ttT---A---A---S-V-.....................-pkaf--T-..-.---....KTS--
sap3  ---------------A--T-ttFNN-----A-N-----.....................-hk.....-V--P-GEintnVN---
sap4  ---Q--LSV--F---I--A-..V---T-------N-.-rslvdpkdptvevkrsplfldiep-ei..--DD.--......rnDVG--
```

Mature Proteins:

```
                                        ---------
sap1  QAIPVTLNNEHVSYAADITIGSNKQKFNVIVDTGSSDLWVPDASVTCDKPRPGQSADFCKGKGIYTPKSSTTSQNLGSPFYIGYGDGSSS
sap2  --V----H--Q-T------V---N--L---------------VN-D-QVTYSD-T-----Q--T-D-SG-SA--D-NT--K---------
sap3  -TV--K-I--Q----S---V------LT-VI-----------SQ-S-QAG.Q--DPN---NE-T-S-S--SS----N---S-E----TT-
sap4  GPVA-K-D--IIT-S--------N--LS--------------SNAV-IPKW--DRG----NN-S-S-AA-S--K--NT--E-K-A---VQ

sap1  SQGTLYKDTVGFGGASITKQVFADITKTSIPQGILGIGYKTNEA.AGDYDNVPVTLKNQGVIAKNAYSLYLNSPNAATGQIIFGGVDKAK
sap2  --------------V--KN--L--VDS---D-----V--------.G-S---------K----------------D------------N--
sap3  ----W----I----I-----Q---V-S--VD----------H--.E-N------------I-S----------RQ-TS---------N--
sap4  A--N--Q----I--V-VRD-L--NVRS--AHK------FQS---tRTP---L-I---K--I-S------F-----E-SS--------I----

                                        ---------
sap1  YSGSLIAVPVTSDRELRITLNSLKAVGKNIN.GNIDVLLDSGTTITYLQQDVAQDIIDAFQAELKLDGQGHTFYVTDCQTSGTVDFNFDN
sap2  -------L----------S-G-VEVS--T--tD-V----------------L-DQ--K--NGK-TQ-SN-NS--EV--NL--D-V---SK
sap3  ---T---L-----N----H--TV-VA-QS--.ADV--------------G--DQV-S--NGQETY-AN-NL--LV--NL--S---A--K
sap4  -----VDL-I----T-SVG-R-VNVM-Q-V-.V-AG---------S-FTPNI-RS--Y-LGGQVHY-SS-NEA--A--K-------Q--R

sap1  NAKISVPASEFTAPLSYANGQPYPKCQLLLGISDANILGDNFLRSAYLVYDLDDDKISLAQVKYTSASNIAALT
sap2  -----------A-S-QGDD----D-----FDVN--------------I------NE-----------S-S---
sap3  ---------------YTED--V-DQ----F-T--Y------------I------NE---------T--------
sap4  -L---------LYQ-Y-T--E-----EIRVRE-ED-------M----I------R---M-------Q---VGIN
```

B

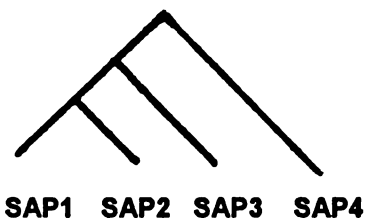


FIGURE 6

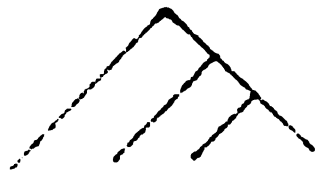


Figure 6 Legend Comparison of the deduced protein sequences of the four *SAP* genes.

A. Alignment of protein sequences. A dash (-) represents a residue identical to the *SAP1* sequence on the top. A dot (.) indicates a space in the sequence introduced to maximally align the sequences. Lys-Arg dipeptides are underlined in the sequences of the precursor peptides. The sections of the active site in the mature protein have a line above the sequence.

B. Phylogenic tree of the mature portions of the four Sap proteins. A phylogenetic tree, constructed using the Feng and Doolittle program (Feng and Doolittle, 1987) based on the protein sequences. A similar tree is constructed from an alignment of the DNA sequences (White, Miyasaki and Agabian, unpublished observations). The lengths of the lines are proportional to the estimated genetic distances.

CHAPTER 6

Future Directions

At their inception, these studies were designed to examine the relationship between important putative virulence factors, the secreted aspartyl proteinase, and the phenomenon of phenotypic switching in *C. albicans*. Our studies have led to a number of interesting and unanticipated findings which will guide our future directions. Rather than a single Sap, our studies with *SAP1* contributed directly to the elucidation of a large family of *SAP* genes in *C. albicans*, all of which appear to be uniquely regulated not only with respect to the phenotypic state of the cell, but also by nutritional and other environmental conditions. The finding of an association between the cellular phenotype of the cell and the expression of specific proteinases has spurred interest in both as potential virulence markers. The effect of switch phenotype upon a number of other potential virulence factors, including cell permeability, adhesiveness, drug susceptibility and sensitivity to host defense factors has led to the theory that specific cell phenotypes might be more virulent than others in the context of a given host environment. Since each *C. albicans* isolate apparently has the ability to cause disease in an immunocompromised host, each strain must be capable of responding to triggers from the host environment. As such, the adaptation of *C.albicans* might include changes in cellular phenotype, along with the expression of the appropriate Saps and other virulence factors. Phenotypic

switching may be accompanied by multiple events including transcription and translation of a number of gene products which might in some cases be related to virulence.

In studying the *SAP1* gene, we have been intrigued by two features of this gene and other members of its large family: first, the structural gene products are extremely similar--in most cases there is not enough difference between the Sap proteins to expect them to be easily assayed either immunologically or indirectly through mRNA hybridization; second, the upstream prepeptide and non-coding *cis* regulatory regions of each of these genes are very different, suggesting that each member of the family is independently regulated, thus ensuring proteinase expression under a wide range of conditions in the host cell. We began our search for potential *cis*-acting transcriptional regulatory regions using *SAP*1 and in the process defined some specific sequences (eg. the *CARE*2-like element, the cytidine rich upstream region and the potential translation initiator sequence)which differentiated this member of the gene family. The identification of *SAP*3 and *SAP*4 led to the realization that a family of secreted aspartyl proteinases existed in *C.albicans*. Determination of the number of Saps made it clear that traditional knock-out experiments to define the roles of the various proteinases would not be easy or clear-cut. With this in mind we are constructing a reporter system in which putative regulatory regions can be tested in conjunction with different phenotypes in the WO1 strain.

Reporter systems are powerful genetic tools which allow the rapid and potentially quantitative analysis of gene function. In *C. albicans* these systems would be extremely useful because diploidy and the lack of a sexual stage in this organism makes it difficult to perform mutations *in vivo.* My planned reporter system invokes the use of an auxotrophic *C. albicans* WO1 strain and a fluorescent reporter gene. The construct which I have made, (pSMYGFP - plasmid map shown in Figure 1) is based upon a similar *C. albicans* promoter construct used by David Soll (ref. Srikantha, 1995). The main difference lies in my use of a different reporter gene, a mutant of the *Aequorea victoria* green fluorescent protein gene (yEGFP). This same mutant was generated and characterized by Brendan Cormack (1997) and has been proven to work in *C. albicans*, under control of the *C. albicans* maltose promoter. All of the CTG codons have been mutated to code for leucine instead of serine, to avoid the *C. albicans* alternate codon usage problem (Tuite, Bower and McLaughlin 1986 and White 1995) in the y*GFP*3 mutant. In addition, my construct has two *C. albicans* autonomously replicating sequences (ARS) sequences (Pla et al., 1995, Cannon et al., 1990) as well as the *AMP* gene for replication in *Escherichia coli* and the *URA3* gene (Gillum, Tsay, Kirsch, 1984) for selection and plasmid maintenance after transformation into a *ura- C. albicans* strain. Multiple cloning sites are present both upstream and downstream of the *yGFP*3, to facilitate placement of *SAP* gene 3' flanking sequences as well as promoter sequences.

In order to utilize this construct I will need a mutant auxotrophic *C. albicans* strain. My interest in phenotypic switching would be well served with a mutant WO1 isolate. Unfortunately, the only auxotrophic mutant of WO1 which exists is an *ade2* mutant (McEachern, personal communication). The *ADE2* gene is an inconvenient one to use in constructs as it is large (5 kb), and has unfortunate restriction sites within the coding sequence. Therefore I are using the protocol refined by Fonzi and Irwin (1993). to create a mutant WO1 isogenic *ura-* strain.

To construct a WO1 *ura-* isogenic strain, a recombinatorial mutagenesis scheme is being used to generate sequential mutations in both alleles of the target *URA3* gene of a *C. albicans* WO-1 isolate. The plasmid (p*Ura3*Δ::λ) generously supplied by W. Fonzi, contains a *C. albicans URA*3 gene in which the majority of the *URA*3 coding sequences have been replaced with the lamda immunity region. This plasmid is then linearized within the *URA3* sequence, and subsequently transformed into *C. albicans*. Because it has *URA*3 DNA on either end it preferentially targets the *C. albicans URA*3 gene for integration. (The integrative transformation efficiency has been reported a 3 to 6 X 10^{-6} transformants per viable spheroplast (Fonzi and Irwin, 1996)) The transformants are then plated out to obtain approximately 2.5 X10^{4} spheroplasts per plate with 25 plates. These are allowed to mature into colonies from which genomic DNA is subsequently prepared.

A PCR reaction is then performed using DNA from each of the 25 plates, using 2 primers. One derived from the native *URA*3 gene and the other from within the lambda immunity locus. Thus, only DNA which has undergone an integrative transformation event can be amplified. Once a plate is identified which contains the desired transformant, serial dilutions are plated and alternated with PCR amplification until a single colony is identified. This transformant then undergoes a second transformation with the same pURA3::λ DNA. However this second time, selection is made much easier with the use of 5FOA (fluoroorotic acid) which selects against heterozygotes or *URA*+ revertants. This should result in the selection of a *ura*- double knock-out isolate. I am currently in the process of isolating a single knockout transformant, and hope to have the double knockout mutant soon.

Ultimately with these methods I will have a system for testing putative regulatory regions when fused with the yGFP3 reporter construct. Once *cis*-acting regulatory regions have been identified, comparisons can be made with potential regulatory regions from other putative virulence factors. With more than one of these in hand, experiments can be designed to trap and identify *trans*-acting regulatory proteins/nucleic acids. The ultimate goal is to begin to understand the inter-regulation of a number of virulence factors as well as what triggers their expression. These types of studies can potentially lead to the development of novel antifungal therapies, some of which may be directed at modulating the host environment as well as attenuating the fungal response.

Candida has been called the ultimate opportunist, and what we learn about its transcriptional and translational regulatory systems, may enable us to extend our knowledge of other opportunistic infections. Characterizing these proteolytic enzymes gives us information to design new drug therapies, totally different from the kinds available to us now. Finally, learning about these mechanisms of fungal virulence lets us appreciate the complexity of the interactions between the host and the parasite and reminds us that it is not a static situation. In fact the organism adapts quite well to the changing host status and we must strive to keep up with *Candida's* adaptive ability.

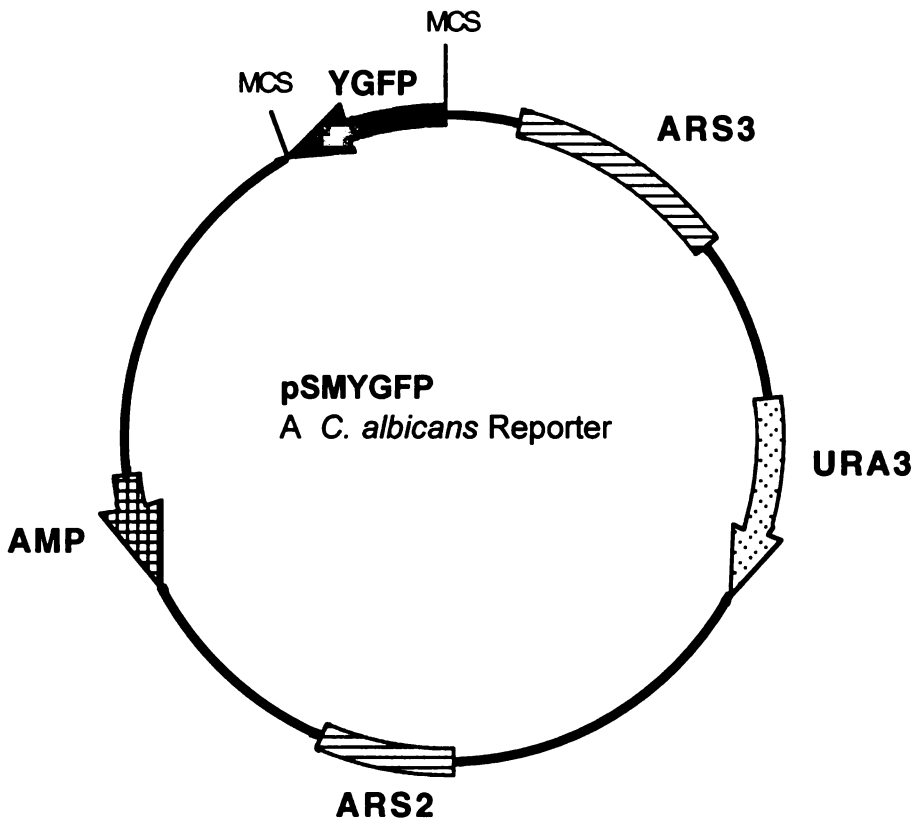
REFERENCES - CHAPTER 6

1. Cannon, R. D., F. Howard, M. G. Shepherd. 1990. Isolation and Nucleotide sequence of an autonomously replicating sequence (*ARS*) element functional in *Candida albicans* and *Saccharomyces cereviseiae*. *Mol. Gen. Genet.* 221:210-218.
2. Cormack, B. P., G. Bertram, M. Egerton, N. A. R. Gow, S. Falkow and A. J. P. Brown. 1997. Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene expression in *Candida albicans*. *Microbiology* 143:303-11.
3. Fonzi, W. A. and Irwin, M. Y., 1993. Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* 134:717-28.
4. Gillum, A. M., E. Y. H. Tsay and D. R. Kirsch. (1984) Isolation of the *Candida albicans* gene for orotidine-5'-phosphate decarboxylase by complementation of *S. cerevisiae ura*3 and *E. coli pyr*F mutations. *Mol. Gen. Genet.* 198:179-82.
5. Pla, J., R. M. Pérez-Díaz, F. Navarro-García, M. Sánchez and C. Nombela. 1995. Cloning of the *Candida albicans HIS*1 gene by direct complementation of a *C. albicans* histidine auxotroph using an improved double-*ARS* shuttle vector. *Gene* 165:115-20.
6. Srikantha, T., Chandrasekhar, A. and Soll, D.R. 1995. Functional analysis of the promoter of the phase-specific *WH*11 gene of *Candida albicans*. *Mol. Cell. Biol.* 15:1797-1805.

7. Tuite, M. F., P. A. Bower and C. S. McLaughlin. 1986. A novel suppressor tRNA from the dimorphic fungus *Candida albicans*. *Biochim. Biophys. Acta* 866:26-31.

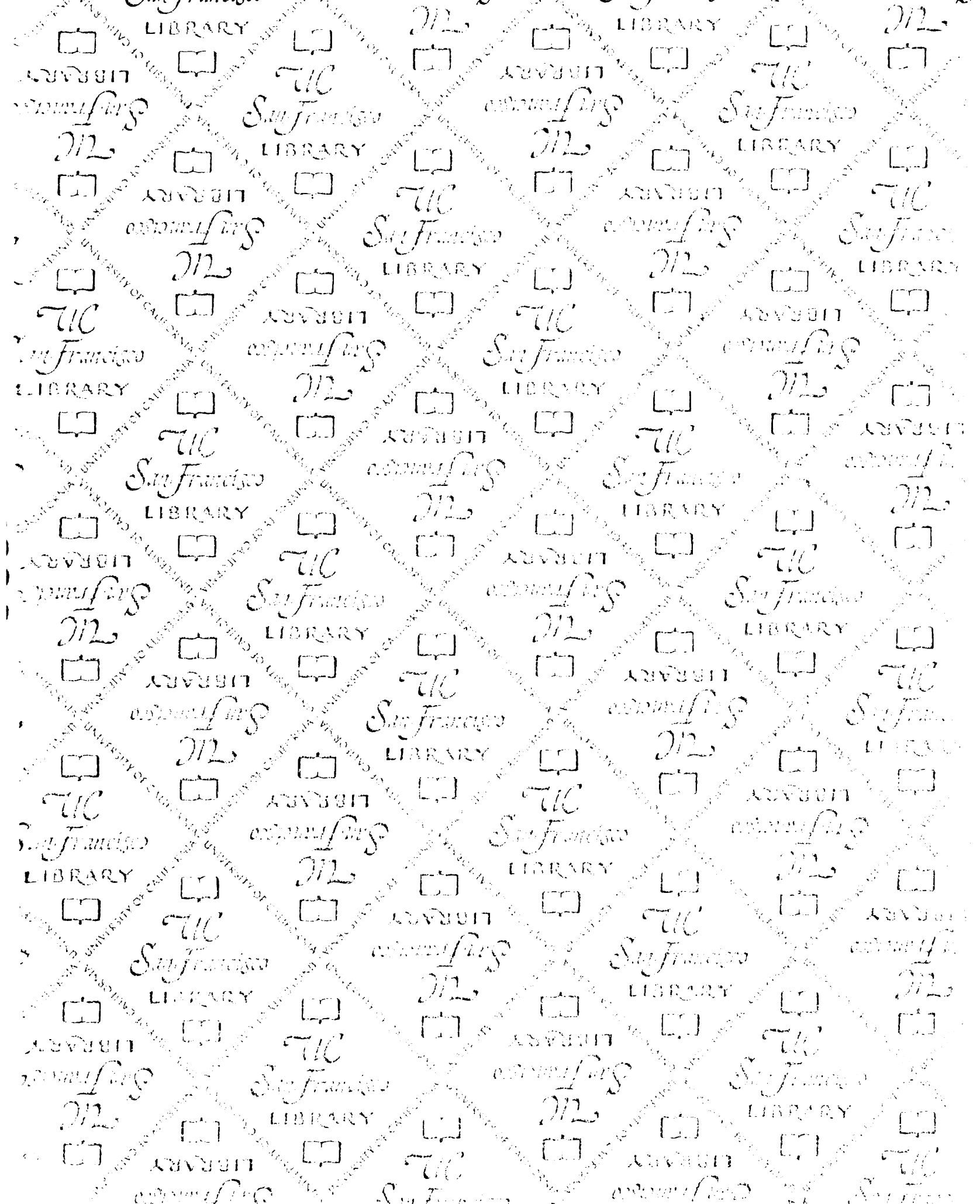
8. White, T. C., L. E. Andrews, D. Maltby and N. Agabian. 1995. The "universal" leucine codon CTG in the secreted aspartyl proteinase 1 (*SAP*1) gene of *Candida albicans* encodes a serine in vivo. *J. Bacteriol.* 177:2953-5.

Plasmid Map pSMYGFP Reporter Construct



Abbreviations:
ARS - *Candida albicans* autonomously replicating sequence (Cannon,Jenkinson and Shepherd 1990; Pla et al. 1995)
AMP - *E. coli* Ampicillin gene
yGFP - yEGFP mutant *A. victoria* green fluorescent protein gene (Cormack 1997)
URA - *C. albicans URA3* gene (Gillum, Tsay and Kirsch 1984)
MCS - multiple cloning site

Black single lines denote pBluescript KS II+ sequences (Stratagene, La Jolla, California)


UC
San Francisco
LIBRARY

UC San Francisco LIBRARY

For reference

Not to be taken from the room.