JOURNAL OF BACTERIOLOGY, Oct. 1993, p. 6126–6133
0021-9193/93/196126-08$02.00/0

Vol. 175, No. 19

Three Distinct Secreted Aspartyl Proteinases in *Candida albicans*

THEODORE C. WHITE, SHELLEY H. MIYASAKI, AND 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

Received 7 April 1993/Accepted 30 July 1993

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 polymerase chain reaction 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 *C. albicans*.

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, *C. albicans* develops from a commensal to a pathogenic organism, causing oral, vaginal, and/or systemic candidiasis. The mechanism by which *C. albicans* becomes pathogenic has not been elucidated, although several factors which contribute to the development of disease have been characterized (reviewed by Cutler [4]).

C. albicans exists in at least two distinct morphological forms, spherical yeast forms which divide by budding and long hyphal forms (reviewed by Odds [16]). This morphological transition is considered a major factor in the ability of *C. albicans* to avoid immune destruction. More recently, the ability of *C. albicans* 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 [20]). Switching between several different colony morphologies can occur at relatively high frequencies (10^{-2} to 10^{-4} per cell cycle). Several different switch systems, all of which may contribute to the organism's pathogenicity, have been described for *C. albicans*. Switch phenotypes have also been described for other pathogenic *Candida* spp. The switching system that is discussed in this article is displayed in the WO-1 strain of *C. albicans*, in which switching results in two distinct cell morphologies, white (W) and opaque (O). W cells appear as normal round yeast forms, appear smooth in scanning electron microscopy, and form white colonies on agar. O cells are larger and oval, contain large cytoplasmic vacuoles, possess a rough and punctate surface in scanning electron microscopy, and grow as opaque colonies on agar (1, 19).

Another characteristic of *C. albicans* which may contribute to its virulence is the secretion of aspartyl proteinases, first identified by Staib (21). Originally designated CAP (*Candida* aspartyl proteinase), this group of proteinases has been given a variety of labels. A recent article (15)z referred to the first cloned gene as *PEP1* because of its similarity to the pepsinogen gene; however, the *Candida* proteinase resembles many different aspartyl proteinases and is not homologous to pepsinogen. The article which reported the cloning of the second gene referred to both secreted aspartyl proteinase genes as *PRA* genes, because of their similarity to the PrA gene of *Saccharomyces cerevisiae* (26). However, the PrA gene product is a vacuolar protein which is not secreted. While the secreted aspartyl proteinases of *C. albicans* are similar to PrA in sequence, they are not homologs. A related *Candida* aspartyl proteinase which is not secreted (11) is more closely related to the PrA proteinase. We have recently (23) identified a *Saccharomyces* homolog 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* [9]) and *SAP2* (formerly *PRA11* [26]). (The *SAP* nomenclature was recently agreed upon at the 1993 American Society for Microbiology Meeting on Candida and Candidiasis.)

Candida proteinases have been biochemically characterized in a number of laboratories (reviewed in references 6 and 18). The aspartyl proteinases from a variety of strains range in molecular mass between 40 and 45 kDa and 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 (BSA), keratins, secretory immunoglobulin A, 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

* Corresponding author.

SAP2, both encoding related secreted aspartyl proteinases (9, 26).

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 "propeptide" form of the proteinase. Both peptides appear to be removed from the mature protein by a specific proteinase(s) which cleaves after Lys-Arg residues, much like the *KEX2* cleavage of the α mating factor in *S. cerevisiae* (10). 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 sizes of the proteins are very similar (36,249 versus 36,341 kDa).

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 (immunoblots) 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 *C. albicans* secreted proteinase, all secreted from a clonal *Candida* strain (WO-1). 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 article, criteria for differentiating between the three different proteinases, which can be used to determine the presumptive role of each of these proteins in *Candida* virulence, are described.

MATERIALS AND METHODS

Strains. The common laboratory strain, 3153A (ATCC 28367), and strain WO-1, which switches colony morphologies (19), were generously supplied by David Soll (University of Iowa). Strain SS, which was obtained from Remo Morelli (San Francisco State University), is a laboratory strain which forms hyphae at a high frequency under standard conditions and which was originally isolated from a patient with systemic candidiasis (3).

Growth conditions and media. Strains were routinely grown in YEPD (10 g of yeast extract, 20 g of peptone, and 20 g of dextrose per liter; autoclaved) and in YCB-BSA (23.4 g of yeast carbon base, 2 g of yeast extract, and 4 g of BSA per liter; pH to 5.0; filter sterilized). YCB-BSA induces proteinase expression in most strains, while YEPD does not. Cell cultures were inoculated by using a single colony grown on a YEPD agar plate. Cultures were grown for 48 h to late stationary phase. The cells were pelleted by using a tabletop centrifuge, and the culture supernatants were stored at −20°C. W and O cells were maintained on YEPD agar plates containing phyloxine B (5 μg/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 [ethylene glycol-bis(β-aminoethyl ether)-*N,N,N',N'*-tetraacetic acid] containing 10 μg of phenylmethylsulfonyl fluoride, 13.3 μg of *N*-ethylmaleimide, and 6.67 μg of diethylene triaminepentaacetic acid per ml}. The concentrated culture supernatants were electrophoresed through a sodium dodecyl sulfate (SDS)-polyacrylamide gel, and the region of the gel corresponding to 35 to 45 kDa 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 3,000 × *g* for 15 min. The eluate was lyophilized, suspended in water, dialyzed against Tris-EDTA-EGTA buffer, and stored at −20°C. A female New Zealand White 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-polyacrylamide gel electrophoresis (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 an SDS–10% polyacrylamide gel. Coomassie blue staining was performed as described elsewhere (8). For Western blots, SDS-polyacrylamide gels were electroblotted to nitrocellulose by using 50 mM Tris HCl, 380 mM glycine, 0.1% SDS, and 20% methanol in a Bio-Rad Trans-blot Cell (Bio-Rad, Hercules, Calif.) run overnight at 0.2 A 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 (25), eliminating high background levels 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 phosphate-buffered saline 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 (1/5,000 dilution in TBST) for 1 h, washed three times in TBST, incubated for 1 h with the secondary antibody (1/5,000 dilution of horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G antibody [Zymed, South San Francisco, Calif.] in TBST), and washed three times. The horseradish peroxidase was detected with the ECL Western blotting detection reagents (Amersham, Arlington Heights, Ill.), 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 pIs of the three proteinases were determined by using the Rotofor Cell (Bio-Rad) preparative isoelectric focusing apparatus. The Rotofor cell was assembled as described by the manufacturer by using 2% ampholytes (Biolytes [pH 3 to 10]; Bio-Rad) 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 V). Samples were collected across the gradient and analyzed by Western blot.

Protein sequencing. Proteins to be sequenced were concentrated 10-fold by using a Speed Vac Concentrator (Savant Instruments, Farmingdale, N.Y.) and electrophoresed on an SDS–10% polyacrylamide gel. The gel was blotted to a Westran polyvinylidine difluoride membrane (Schleicher & Schuell, Keene, N.H.) by using CAPS buffer {10 mM (3-[cyclohexylamino]-1-propanesulfonic acid) [pH 11], 10% methanol}. The Westran membrane was stained with Coomassie blue, and the desired band was 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 amino acids were analyzed with an on-line microbore high-performance liquid chromatography system (ABI model 120A) using Edman chemistry.

Nucleic acid manipulation. DNA was isolated as described elsewhere (2). Restriction enzyme digestions, gel electrophoresis, Southern blot hybridization, and subcloning were performed as described elsewhere (12) by using the Bluescript plasmid vector and *Escherichia coli* XL-1 Blue (Stratagene Inc., La Jolla, Calif.). Genomic libraries were constructed by using SS genomic DNA partially digested with *Sau*3A and the λ FIX II system (Stratagene Inc.). Low-stringency hybridization and washes were performed with 3× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at 55°C. High-stringency hybridizations were performed with 3× SSC at 65°C and washed to 0.1× SSC.

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

SAP1, 5′CCAGTAGCATTAACAGGAGTTTTAATGACA3′
SAP2, 5′TGACCATTAGTAACTGGGAATGCTTTAGGA3′
SAP3, 5′TTGATTTCACCTTGGGGACCAGTAACATTT3′

The oligonucleotides were labeled with polynucleotide kinase and [γ-^{32}P]ATP (12) and used to probe Southern blots of genomic DNA. Hybridizations and washes were performed at 42°C in 5× SSC.

DNA sequencing. The regions of phages 15 and 17 encoding the mature termini of the *SAP* genes were sequenced by using the *fmol* DNA Sequencing System (Promega) according to the 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 polymerase chain reaction (PCR) fragment from *SAP1* was subcloned as a 1-kb *Eco*RI fragment (see Fig. 5). The entire *Eco*RI fragment (shown to contain the 5′ end of *SAP3* [Fig. 5]) was sequenced by using Sequenase according to the 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 described above.

RESULTS

To investigate the expression of the secreted aspartyl proteinases in various strains of *C. albicans*, stationary-phase culture supernatants from strains 3153A, SS, and WO-1 grown in the proteinase-inducing medium, YCB-BSA, were examined (W cells and O cells were examined separately). Figure 1A shows the Coomassie blue staining pattern of an SDS-PAGE gel of the culture supernatants. A single major protein band (43 kDa) was detected in culture supernatants obtained from strain 3153A, from strain SS, and from W cells from strain WO-1, while two major bands of slightly smaller sizes (40 and 41 kDa) were found in culture supernatants of O cells from strain WO-1. Together, these bands represent the major protein(s) secreted into the culture medium by each of these strains. The molecular masses of these proteins range between 40 and 43 kDa, the expected size of the secreted aspartyl proteinase. From Fig. 1A, it is clear that the cells have completely degraded the BSA which was present in the growth media as a source of nitrogen.

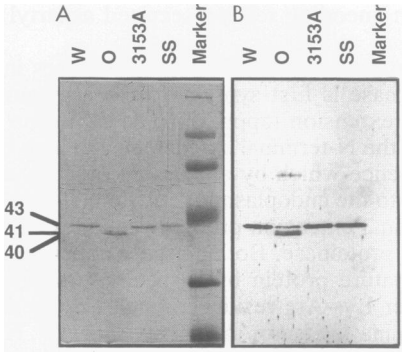

FIG. 1. Detection of three proteinases. A Western blot of an SDS-polyacrylamide gel of culture supernatants of strain 3153A, of strain SS, and of W and O cells from strain WO-1 is shown. (A) A Coomassie blue stain of one half of an SDS-PAGE gel. The marker lane contains approximately 1 μg of each protein band; the sizes of the bands (from the top) are 208, 101, 71, 44, 29, and (18 + 15) kDa. (B) A Western blot of the second half of the same gel probed with the antiproteinase antibody (see Materials and Methods).

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 Fig. 1B, a duplicate panel from the same gel was blotted to nitrocellulose and reacted with rabbit antiserum. The major protein band of 43 kDa in supernatants from strain 3153A, from strain SS, and from W cells of strain WO-1 reacts with the antiproteinase antibody, as do the two major proteins of 40 and 41 kDa 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 the sizes of the proteinases from culture supernatants of O cells will be examined below.

In all Western blots presented, a faint 43-kDa band is observed for O cells (Fig. 1B). The presence of the 43-kDa 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 accounts for the amount of the 43-kDa protein in the O-cell culture supernatant (24).

The antibody to the *C. albicans* proteinase was also used to probe Western blots of culture supernatants from other yeast species (including *Candida tropicalis*, *Candida krusei*, *Candida glabrata*, and *S. cerevisiae*) grown in the proteinase-inducing medium, YCB-BSA. While aspartyl proteinases are secreted by other pathogenic *Candida* spp. under these conditions, the antibody did not cross-react with these proteinases (24). Neither did these yeasts appear to degrade the BSA present in the YCB-BSA medium at the same rate as the *C. albicans* strain. In contrast, over 40 clinical isolates of *C. albicans* have been tested for proteinase production by growth in YCB-BSA. All strains tested produce proteinase (detectable by a Western blot of culture supernatants) and substantially degrade the BSA present in the medium (24).

The results discussed above suggest that three different forms of aspartyl proteinase (40, 41, and 43 kDa) can be secreted by a single *Candida* strain (WO-1). To determine whether the secretion of these different proteinases is depen-

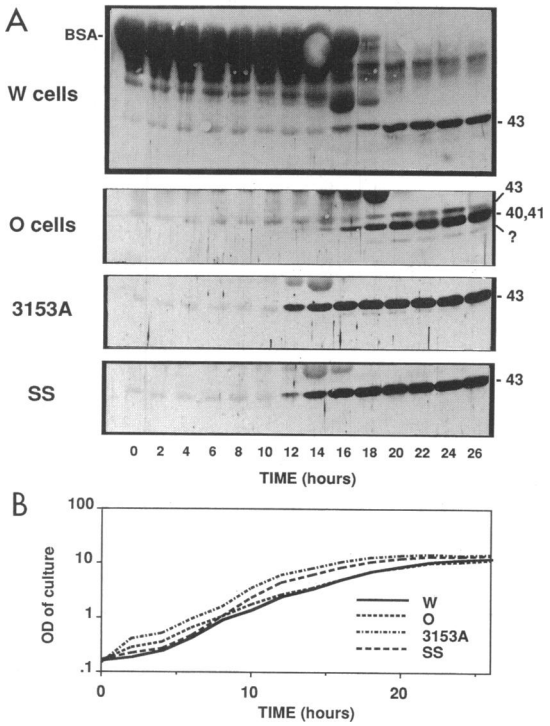

FIG. 2. Expression of three proteinases during growth. Growth curves were constructed for strain 3153A, strain SS, and W and O cells from strain WO-1. Cells were grown in YCB-BSA from an initial optical density (OD) at 600 nm of 0.1 and sampled every 2 h. The OD at 600 nm of an appropriate dilution was determined for each sample, the cells were pelleted, and the culture supernatant was 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 (40, 41, and 43 kDa) 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 kDa in size is also detected, possibly a degradation product of one of the proteinases. In all panels, a faint band is visible at 43 kDa throughout the time course. This band is a degradation product of the BSA component of the media. The BSA component of the medium is indicated in the W-cell panel. The BSA is detected in the Western blot 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) Growth of the four cultures as determined by the OD at 600 nm.

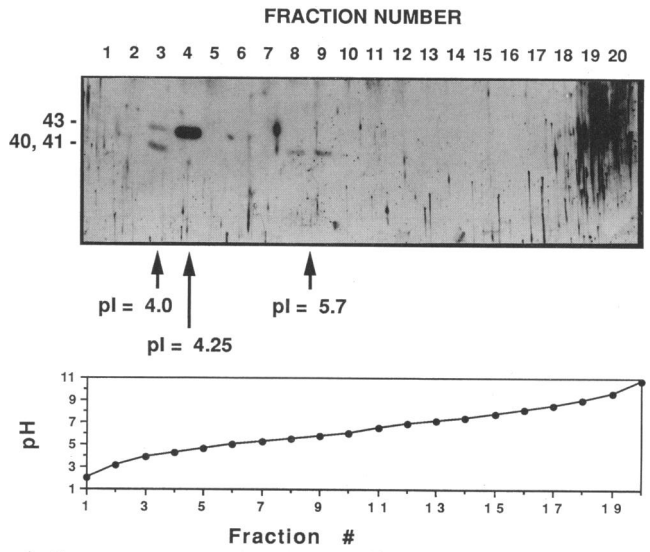

FIG. 3. Isoelectric focusing of the three isoenzymes. A mixture of W- and O-culture supernatants was subjected to preparative isoelectric focusing (see Materials and Methods). Fractions were collected and stored at −20°C for analysis. The Western blot of a sample of each of the fractions is shown at the top. 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.

dent on the growth state of the cells, a time course of proteinase production and cell number (measured by optical density) was determined for each of the three *C. albicans* strains and for both W and O cells. Figure 2 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 (24). The appearance of the proteinases is correlated with the depletion of the 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 (18). However, O cells produce proteinase in YEPD medium (which does not contain BSA) with the same kinetics (24). No other cell type produces proteinase under these conditions. The 41-kDa protein, which is frequently obscured by the 40-kDa protein, appears later in the time course of growth in O cells (Fig. 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 Fig. 3, the 43-kDa protein has an isoelectric point at approximately pH 4.25 and the 40-kDa 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 (17, 18). Surprisingly, the 41-kDa protein has an isoelectric point of pH 5.7, much higher than any previously observed pI for a secreted proteinase from *Candida* spp.

There are several possible explanations for the presence of three protein bands reacting with antiproteinase antibody in Western blots. Proteinase activities with distinct characteristics have been observed under different growth conditions by several researchers (18), 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 posttranslationally 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 Materials and Methods. The N-terminal sequence of each of the three protein bands is shown in Fig. 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* (9) and *SAP2* (26) (Fig. 4). The 43-kDa protein from W cells corresponds to the predicted mature product of the *SAP2* gene, and the 40-kDa protein

															Ref
***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 Sequences:																
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	

FIG. 4. N-terminal sequences of secreted aspartyl proteinases. The N-terminal sequences of secreted aspartyl proteinases of *C. albicans* can be grouped into three classes corresponding to the three isoenzymes described in this article. Each sequence is listed with the strain from which the sequence was derived. "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. Parentheses around an amino acid represent a residue in which the protein sequence is not definitive. The reference for each sequence is listed on the right. An asterisk indicates that the sequence was determined in this study.

from O cells corresponds to the predicted mature product of the *SAP1* gene. The 41-kDa 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-kDa-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 *C. albicans* strains that have been previously reported (Fig. 4). Since the three protein bands were sequenced from the same strain, WO-1, these data clearly demonstrate that three different isoenzymes of aspartyl proteinase are secreted by the WO-1 strain of *C. 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 (Fig. 1) and the N-terminal sequence data (Fig. 4), there appears to be a correspondence between the Sap isoenzyme and its apparent molecular weight. However, Sap isoenzymes might be modified posttranslationally (e.g., glycosylation) under a variety of conditions resulting in various molecular weights. To examine the correspondence of isoenzyme and molecular weight, the N-terminal sequences of the 43-kDa proteinase produced by strains 3153A and SS in YCB-BSA media were determined and compared with the sequences from strain WO-1. The results (Fig. 4) confirm that 43-kDa proteinases in all three strains, 3153A, SS, and WO-1, are the products of the *SAP2* gene locus.

The DNA sequences of both alleles of the *SAP1* gene from strain WO-1 and from an unrelated strain, SS, have been determined (13). The N-terminal sequences of these alleles are shown in Fig. 4. The *SAP2* and *SAP3* N-terminal sequences (Fig. 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, encode different proteinase isoenzymes. To identify the gene for *SAP3*, a λ genomic library was screened with a PCR fragment from *SAP1* under low-stringency hybridization conditions (13). Two λ phages which hybridize with *SAP1* at low stringency but not at high stringency were identified. These two recombinant λ phages, 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 Fig. 4. The deduced protein sequence of phage 15 corresponds to the protein sequence determined for the 41-kDa 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 mechanisti-

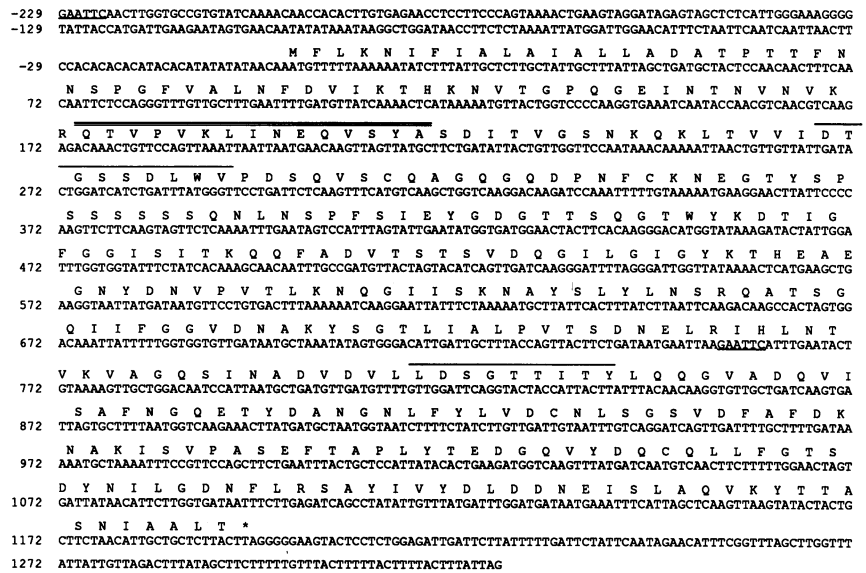

FIG. 5. Sequence of the *SAP3* gene. The sequence of the *SAP3* gene is shown, together with the deduced amino acid sequence of the protein. The sequence begins at the first of two *Eco*RI sites in the sequence (underlined). The putative N terminus of the mature protein, which matches the protein sequence of the 41-kDa protein, has a double line above the sequence. The two sections of the active site of the proteinase which are highly conserved among the Sap proteins have a single line above the sequence. The entire sequence is 1,556 nucleotides in length. The GenBank assession number is L22358.

cally related to *KEX2* in *S. cerevisiae* (10). 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 Fig. 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 the *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 indicates that the proenzyme segment of each protein, between the signal sequence and the mature N terminus, is the most divergent region of the proteins (24). Three oligonucleotides (30 bases in length) were designed to distinguish between the three genes. The oligonucleotides were used to determine the complexity of the proteinase genes by Southern blot analysis of *Eco*RI-digested genomic DNAs from three different strains. The results (Fig. 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 a restriction fragment length polymorphism at the *SAP1* gene (13). The *SAP2* probe recognizes the *SAP2* 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 *SAP3* probe hybridizes to the *SAP3* 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 *SAP3* are not due to partial digestion, since the *SAP1* blot,

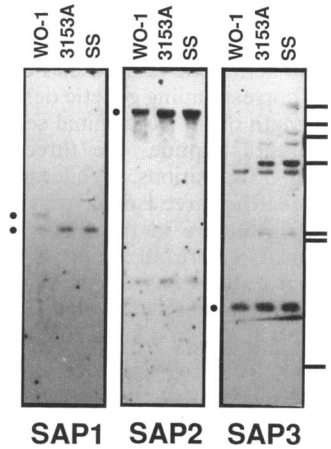

FIG. 6. Genomic blot showing genomic organization of three proteinases. Genomic DNA (5 μg) from each of three strains (3153A, SS, and WO-1) was digested with *Eco*RI, electrophoresed in a 0.8% agarose gel, and blotted to nitrocellulose. Three identical panels from the same gel were separated and probed with each of the three oligonucleotides (see Materials and Methods). Hybridizations and washes were done at 42°C. Lines on the right represent size markers of λ DNA cut with *Hin*dIII (sizes from the 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.

TABLE 1. Characteristics of the *SAP* gene products

Isoenzyme	Mol wt (10^3)		pI	Signal sequence		No. of potential N-linked glycosylation sites[a]	Expression[b] in:	
	Apparent	Predicted		No. of amino acids	No. of Lys-Arg dipeptides		WO-1	Clinical strains[c]
Sap1	40	36.2	4.0	50	2	0	+ (O)	−
Sap2	43	36.3	4.25	56	2	2	+ (W)	+
Sap3	41	36.5	5.7	58	1	1	+ (O)[d]	−

[a] Number of N-X-S/T sites in the mature protein.
[b] In YCB-BSA.
[c] See reference 23.
[d] Low-level expression in W cells has also been observed.

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 *C. 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 *C. 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 (Table 1) and are products of three separate genetic loci (Fig. 6). The differences between the three proteinases suggest that the three proteinases may have unique roles in the interaction between *C. albicans* and its host.

The proteinases are sufficiently distinct to be readily identified (Table 1). The most straightforward method of distinguishing the three proteinases is by their apparent molecular weights with 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 (17).

The three isoenzymes of secreted aspartyl proteinase can be easily identified on the basis of their N-terminal sequences (Fig. 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 at 8 of 15 positions. While many amino acid differences between the three isoenzymes exist (Fig. 4), the most distinguishing feature of the three sequences in this region is position 8, where all three isoenzymes differ in amino acid sequence; Sap1 has an Asn, Sap2 has a His, and Sap3 has an Ile. Morrison et al. (14) have identified a fourth, unrelated Sap protein, the first six amino acids of which are highly divergent. However, at positions 3 and 8, this protein resembles Sap3.

Secreted aspartyl proteinases from other *Candida* species have also been cloned. These include one gene from *C. tropicalis* (22) and two tandemly linked genes from *Candida parapsilosis* (5). 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 (Fig. 4). The predicted N-terminal sequence of the *C. tropicalis* proteinase gene most closely resembles Sap3 of *C. albicans* (Fig. 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 on the basis of positions 8 and 13. The *C. tropicalis* and *C. parapsilosis* genes are more closely related to each other than to the *C. albicans* genes (Fig. 4, positions 1, 8, 11, and 12).

The large size difference between the Sap proteins (40, 41, and 43 kDa) 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 two potential N-linked sites, one at position 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 cleaves 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 (15, 17). 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-polyacrylamide gels (Fig. 1). Yet, it has been reported that O cells have 20-fold higher levels of proteinase activity (15, 17). 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 medium (such as YCB-BSA) varies greatly depending on which cell type (W or O) is in the culture (23). Therefore, secreted proteinases with different pI and inactivation spectra may be inactivated in media of differing pH before being assayed.

An intracellular, nonsecreted aspartyl proteinase gene from *C. albicans* has been described previously (11). 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 et al. (15). 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 (RNA) blots of O cells while little, if any, *SAP1* mRNA was detected from W cells.

The identification of three distinct proteinase isoenzymes in *C. albicans* and the observed expression of all three isoenzymes from a single strain (WO-1) under different switch phenotypes allow 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 under 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 David Soll, University of Iowa, for providing strain WO-1 and for his endless support in our entry into the field of *Candida* research; 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 NIH (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 NIH Dentist-Scientist award (5K15DE00334) to S.H.M.

REFERENCES

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**:5579–5588.
2. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.).** 1989. Current protocols in molecular biology, p. 13.11.1–13.11.4. John Wiley & Sons, New York.
3. **Blagoyevich, P. G.** 1979. M.S. thesis. San Francisco State University, San Francisco, Calif.
4. **Cutler, J. E.** 1991. Putative virulence factors of *Candida albicans*. Annu. Rev. Microbiol. **45**:187–218.
5. **DeViragh, P. A., D. Sanglard, G. Togni, R. Falchetto, and M. Monod.** 1993. Cloning and sequencing of two *Candida parapsilosis* genes encoding acid proteinases. J. Gen. Microbiol. **139**: 335–342.
6. **Douglas, L. J.** 1986. *Candida* proteinases and candidosis. Crit. Rev. Biotechnol. **8**:121–129.
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**:2972–2977.
8. **Harlow, E., and D. Lane.** 1988. Antibodies: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
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**:129–132.
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-α-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. Nucleic Acids Res. **17**:1779.
12. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Miyasaki, S. H., T. C. White, and N. Agabian.** Submitted for publication.
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**:2030–2036.
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. Bailliere Tindall, Philadelphia.
17. **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. Plenum Press, New York.
18. **Rüchel, R., F. De Bernardis, T. L. Ray, P. A. Sullivan, and G. T. Cole.** 1992. *Candida* acid proteinases. J. Med. Vet. Mycol. **30**(Suppl. 1):123–132.
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**:189–197.
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**:187–193.
22. **Togni, G., D. Sanglard, R. Falchetto, and M. Monod.** 1991. Isolation and nucleotide sequence of the extracellular acid proteinase gene (ACP) from the yeast, *Candida tropicalis*. FEBS Lett. **286**:181–185.
23. **White, T. C., 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**:143–153.
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**: 7848–7853.
27. **Yamamoto, T., K. Nohara, K. Uchida, and H. Yamaguchi.** 1992. Purification and characterization of secretory proteinase of *Candida albicans*. Microbiol. Immunol. **36**:637–664.