THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1994 by The American Society for Biochemistry and Molecular Biology, Inc.
Vol. 269, No. 37, Issue of September 16, pp. 22945–22951, 1994
Printed in U.S.A.

Identification of a Putative Transcription Factor in *Candida albicans* That Can Complement the Mating Defect of *Saccharomyces cerevisiae ste12* Mutants*

(Received for publication, March 18, 1994)

Krishnamurthy Malathi, Kaliannan Ganesan, and Asis Datta‡
From the School of Life Sciences, Jawaharlal Nehru University, New Delhi-110 067, India

We have isolated an acid proteinase-related gene, *ACPR*, from *Candida albicans* using a partial clone (Ganesan, K., Banerjee, A., and Datta, A. (1991) *Infect. Immun.* 59, 2972–2977) as a probe. Sequencing of the full-length gene revealed an open reading frame that can encode a protein of 699 amino acids. The deduced NH_2-terminal amino acid sequence did not correspond with that determined from the purified secretory acid proteinase; however, the encoded protein is antigenically related to secretory acid proteinase and has a putative active site for acid proteinase. Interestingly, the amino acid sequence of the NH_2-terminal 215 residues of Acprp is highly similar to the DNA binding domain of Ste12p of *Saccharomyces cerevisiae*. Gel retardation experiments showed that this region of Acprp, like Ste12p, could bind to *S. cerevisiae* pheromone response elements, suggesting that Acprp has a function similar to Ste12p. Chimeric constructs composed of *S. cerevisiae STE12* and *C. albicans ACPR* genes complemented the mating defect of *S. cerevisiae* a or α *ste12* mutants. Our results suggest the presence of a signal transduction system in *C. albicans* similar to that of *S. cerevisiae* mating pathway.

Among the clinically important *Candida* species, *Candida albicans* is the most pathogenic. It is an opportunistic pathogen that causes candidiasis in human beings and other warm blooded animals. It occurs naturally as a diploid and lacks a sexual phase in its life cycle. When grown in a medium containing bovine serum albumin or other proteins as a sole nitrogen source, *C. albicans* and some other *Candida* species secrete acid proteinases (Banerjee *et al.*, 1991; Ruchel, 1981; Staib, 1965, 1969). The ability to secrete acid proteinases has been correlated to the pathogenicity of *C. albicans* (McDonald and Odds, 1983).

We have previously isolated a partial clone from a λ gt11 genomic expression library using affinity-purified proteinase antibodies. The partial clone was shown to be specific for various strains of *C. albicans*; *Candida tropicalis* and *Candida parapsilosis* were shown to have related genes (Ganesan *et al.*, 1991). We report here the isolation and characterization of the full-length gene. The deduced NH_2-terminal amino acid sequence of the encoded protein did not correspond with that determined from purified secretory acid proteinase. This indicated that this gene encoded a cross-reacting protein product, and therefore, we refer to this as acid proteinase-related protein (Acprp).[1] The primary sequence revealed a tripeptide sequence (Asp-Thr-Gly), which is characteristic of active sites of acid proteinases. Surprisingly, the NH_2-terminal 200-amino acid region of Acprp was very similar to the corresponding region of Ste12p of *Saccharomyces cerevisiae*. Ste12p is a DNA binding protein involved in *S. cerevisiae* mating. It binds to a heptamer sequence TGAAACA, referred to as pheromone response element (PRE), which is present in multiple copies in the regulatory regions of the genes that respond to pheromones (Dolan *et al.*, 1989; Errede and Ammerer, 1989). We have demonstrated the ability of Acprp to bind PRE *in vitro*. Domain swap experiments suggest that *ACPR* gene can complement the mating defect of *ste12* mutants of *S. cerevisiae* strains. The results indicate the existence of a signal transduction system in *C. albicans* analogous to that of *S. cerevisiae* mating pathway.

EXPERIMENTAL PROCEDURES

Strains, Growth Conditions, and Materials—Bacterial strains used in this study were *Escherichia coli* DH5α and LE392. *E. coli* cells were grown at 37 °C in LB medium (2.5% Luria broth base). Where appropriate, ampicillin was added at 50 μg/ml. Solid media were supplemented with 2% Bacto agar. Yeast strains used in this study are: *C. albicans* SC5314, *S. cerevisiae* EG123 (*MAT* **a** *trp1 leu2 ura3 his4*), SF167–5a (*MAT* **a** *ste12::LEU2 trp1 ura3 his4*), DC14 (*MAT* **a** *his1*), and DC17 (*MATα his1*). Yeast strains were usually grown in YPD medium (2% Bacto peptone, 1% Bacto yeast extract, 2% glucose) or in 0.67% yeast nitrogen base without amino acids (Difco) plus appropriate supplements for selective growth. Solid media were supplemented with 2% Bacto agar for YPD plates and 1.5% agarose for selective plates. Cell density was monitored in liquid cultures by measuring optical density at 595 nm. Restriction endonucleases and DNA-modifying enzymes were obtained from New England Biolabs, Boehringer Mannheim, Pharmacia Biotech. Inc., Promega, and U. S. Biochemical Corp. [α-^{32}P]dATP was from BARC, India (specific activity, 3000 Ci/mmol). All other reagents were of the highest purity grade commercially available.

Construction of Library and Isolation of Genomic Clones—Genomic DNA from *C. albicans* strain SC5314 was isolated (Holm *et al.*, 1986) and partially digested with several dilutions of *Sau*3AI at 37 °C for 1 h. The DNA fragments were fractionated on a 10–40% continuous sucrose density gradient, and fractions in the range of 8–23 kb were pooled. One μg of this pooled DNA sample was ligated with dephosphorylated EMBL3 *Bam*HI arms (Promega) and packaged *in vitro* (Promega). Recombinant phages were amplified in *E. coli* strain LE392 resulting in 100,000 plaques corresponding to 45–50 genome equivalents of *C. albicans*. A 1.3-kb partial genomic clone, which was isolated by immunoscreening using affinity-purified anti-proteinase antibodies (Ganesan *et al.*, 1991), was used to screen 2×10^4 phages of the primary library. The 1.3-kb probe was labeled by random priming (Feinberg and Vogelstein,

* This work was supported by a Central Institute of Medicinal and Aromatic Plants (CIMAP)-sponsored grant from the Council of Scientific and Industrial Research, New Delhi. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) L16451.

‡ To whom correspondence should be addressed. Tel.: 91-11-667676 or 667557 (ext. 454); Fax: 91-11-686-5886.

[1] The abbreviations used are: Acprp, product of *ACPR* gene; *ACPR*, acid proteinase related gene; Ste12p, product of *STE12* gene; PRE, pheromone response element; kb, kilobase(s); ScSte12p, *S. cerevisiae* Ste12 protein; Stekl, *K. lactis* Ste12 protein; PCR, polymerase chain reaction; MBP, maltose-binding protein; *HY1*, hybrid-1; *HY2*, hybrid-2; bp, base pair(s).

A

```
GAATTCATTCTACAAATCAAGAGAGAGAAAGACATTGGCTAATTGGTTCTCTTTAACATTCCATTACCATAGTTTTCAGTTATATACACTACCCCTTGTATTCAATCATCTATTTATCTT   120
GGTTTCCCAACAATATAGTTTAATATCGTGTATTCACATCCCCAAGAGCAAACATTCTTTTTTTTAAATTTGCATACCAATAGTTCTTGTTGTTATTTCGGTTAAAATTATTGAAATTAA   240
TCTAGAATCATCAAATTCTAACAACAACTACATAATACCAGAAGAAAACGCAAGAGAATTTATAGTTGTATCAATTTAGATTTGGATTTACAATTTTATTATTCTTCAATTCTCGATTTC   360
CTCGTGAAGTTTTTTGTTTTAGTTGTTTTTGTTTTTGTTTTTGTTTCCCCATTGACACGAAAAAGTTTTCCATAAAAAAAGAGAAACATTTCCTAGAACCAACTAAACATTCTTGCTTTA   480
CATTATCTTCATTTTAATTTGTTGGTGTTGTTGTAAATCATAAGTTTAAATAATTGGGTTGTTTGCTTTTAAACTTTTAATACATCAAAACGATATAAAACTGTTTCGAAATCACCGATT   600
TCATATTTTTTCTTTACAAGTTTCATATATTTTATAACTAAAGAAAGAAGGTTGGAAGATCTAACAAATTTTTTTGTTCCCCGTTAACCGAAACACACATACATATACTTCTGTCCACAC   720
ACACACTACCACAAATTTTTGAGCTGATTTCAAACAAGGGTTTTAGAATTTTCCGATTTGTTTTGGTCTTAGTTTTAGTTTGACAATCAAAAAAGAATTACAAAACATCGAACAACCTTC   840
GACCTTCCCCCCACACTCGTCCTCCCTCTCCTTGCCTCGAGTACACTTCACACCTCAGATCTCATTAACATAACGTCCTATATTGCCAATCAATTTTATTACGACAAATTCAATAATTAG   960
GGCTATTTTGACACTTTGAATTACAATTTTATACCAGTGTCGCATTATTACTATTTCCATTTTCAACAAAGTATTGGATATAACACGTTTCTATCAATCATTACCAACCTGACCAACAGA  1080
GTATCCATTACTTCTTAGTTAAATATATTTCTTTTCACTTTTTTCACTTTTTTCATTTTTCAATTGTTGTGCCCCCCCCCTCGTTGCGGATCAAATAATATAACTGCCTAATACACTCTT  1200
TCGCCATGTCAATTACTAAAACATACAATGGTGATCCTACATCACTTGTACCAACTCAACTGGTAAAAGAATCATTAAGACTAATTGAAGATTTAAAATTCTTTTTAGCCACTGCTCCAG  1320
    M  S  I  T  K  T  Y  N  G  D  P  T  S  L  V  P  T  Q  L  V  K  E  S  L  R  L  I  E  D  L  K  F  F  L  A  T  A  P  A     39
CAAATTGGCAAGAGAATCAAGTGATTCGAAGATACTATTTAAACCACGATGAAGGTTTTGTCAGTTGTGTTTATTGGAATAACTTATATTTTATCACTGGTACCGATATTGTCCGTTGTA  1440
 N  W  Q  E  N  Q  V  I  R  R  Y  Y  L  N  H  D  E  G  F  V  S  C  V  Y  W  N  N  L  Y  F  I  T  G  T  D  I  V  R  C  I     79
TTGTTTATAAATTTGAACATTTTGGAAGGAAAATAATTGATCGGAAAAAGTTTGAAGAAGGTATTTTCTCCGATTTAAGAAATTTGAAATGTGGCGCCGATGCAATTTTGGAACCTCCTC  1560
 V  Y  K  F  E  H  F  G  R  K  I  I  D  R  K  K  F  E  E  G  I  F  S  D  L  R  N  L  K  C  G  A  D  A  I  L  E  P  P  R    119
GTTCAGAGTTTTTAGAATTCTTATTTAAAAATTCTTGTTTGAGAACACAAAAGAAACAAAAAGTGTTTTTTTGGTTTAATGTTCCTCATGATAAGTTAATGGCAGATGCCTTGGAACGAG  1680
 S  E  F  L  E  F  L  F  K  N  S  C  L  R  T  Q  K  K  Q  K  V  F  F  W  F  N  V  P  H  D  K  L  M  A  D  A  L  E  R  D    159
ATTTGAAAAAGGAAAAAATGGGTCAAAGACCAACAACAATGGCCCATCGAGAACCAGCATTATCATTCCATTACGACGAGTCTTCCAGTTTATACACTCAATTGGGTAAACACATGGAAA  1800
 L  K  K  E  K  M  G  Q  R  P  T  T  M  A  H  R  E  P  A  L  S  F  H  Y  D  E  S  S  S  L  Y  T  Q  L  G  K  H  M  E  T    199
CACAGAAGAGGATCAACGACGCAGCCACAAGCTCTACTTCTAATACTGCTACCACCTTGACCGACACCGGTGTATCCTCAGGATTGAATAATACCACTTCTGGTGGCGGCAGTGATAGTG  1920
 Q  K  R  I  N  D  A  A  T  S  S  T  S  N  T  A  T  T  L  T  D  T  G  V  S  S  G  L  N  N  T  T  S  G  G  G  S  D  S  A    239
CAACCTCCACACACAACAACAATGAGGCATCGACCAAACCAAGTAATGGCAGTGAAAAATCGTCACCGGAGTACACTACAACTGCCCGCGGTAGAGATGAGTTTGGATTCCTTAATGAAG  2040
 T  S  T  H  N  N  N  E  A  S  T  K  P  S  N  G  S  E  K  S  S  P  E  Y  T  T  T  A  R  G  R  D  E  F  G  F  L  N  E  A    279
CCACACCAAGTCAATACAAAGCCAATTCAGATTATGAAGACGATTTCCCATTGGATTATATCAATCAGACCACTCAAAATTCTGAAGATTATATTACTTTGGATGCAAATTATCAGGCAG  2160
 T  P  S  Q  Y  K  A  N  S  D  Y  E  D  D  F  P  L  D  Y  I  N  Q  T  T  Q  N  S  E  D  Y  I  T  L  D  A  N  Y  Q  A  G    319
GAAGTTATGCAAATATGATCGAAGACAATTACGATTCATTTTTGGATGCAACACTATTTATACCTCCAAGTCTTGGCGTACCTACAGGTACAGCTGCGACTGCAACAACATCAAACCAAG  2280
 S  Y  A  N  M  I  E  D  N  Y  D  S  F  L  D  A  T  L  F  I  P  P  S  L  G  V  P  T  G  T  A  A  T  A  T  T  S  N  Q  V    359
TTGCCTTCAACGACGAATACTTGATTGAACAAGCCCAACCAATAAGAACTCCACTACCCCCAATATCATCATCAACAATATCCGGATTGTTACAACCAAAATCAGCTGCTAAATTCTTTT  2400
 A  F  N  D  E  Y  L  I  E  Q  A  Q  P  I  R  T  P  L  P  P  I  S  S  S  T  I  S  G  L  L  Q  P  K  S  A  A  K  F  F  S    399
CACTACAGAGTGCTAATGGTGGAGAAGAATTTTTCCCTGCTTATCAAAATGACCCATCTACTGCTAATGCCGGGTTTGTGCCGCCAATTTCAGCAAAATATGCAACTCAATTTGCTACTC  2520
 L  Q  S  A  N  G  G  E  E  F  F  P  A  Y  Q  N  D  P  S  T  A  N  A  G  F  V  P  P  I  S  A  K  Y  A  T  Q  F  A  T  R    439
GACAAGTCGCCACCCCAACCTATATAAAGGCAATACCACAAACTGGTGCTGCTGCTGCCACTGGTAATGGTGGTCAACCACAACAATATTATGACCAGGCAACTGGTAATGCTTTTTACC  2640
 Q  V  A  T  P  T  Y  I  K  A  I  P  Q  T  G  A  A  A  A  T  G  N  G  G  Q  P  Q  Q  Y  Y  D  Q  A  T  G  N  A  F  Y  P    479
CAGCAGAAATACCGGTGCTGTACAATGTTGTCCATCCTGAAAGTGAATATTGGACCAATAATTCTGGTGCAGTTGCTACTACTGCTGCTGCCGCTGCTCCAATGTATGACGCTTCTGGGT  2760
 A  E  I  P  V  L  Y  N  V  V  H  P  E  S  E  Y  W  T  N  N  S  G  A  V  A  T  T  A  A  A  A  A  P  M  Y  D  A  S  G  F    519
TTCCTATACCTATTAACCAATCATATATGGTAATGAATGAACACGAAATGGTACCTTATCAATACATGAATTCCAATGGTGCTATGATTGGCATGATTCCACCTCACCAACAACAACAGC  2880
 P  I  P  I  N  Q  S  Y  M  V  M  N  E  H  E  M  V  P  Y  Q  Y  M  N  S  N  G  A  M  I  G  M  I  P  P  H  Q  Q  Q  Q  Q    559
AACAACAACAACAAATTGCCATGGGATATCAAAGCATGTTGAGACAACAACAACAACAACAGCAGCAGCAGCAGCAACAACAACCTTCTTCTACAATGACTAAAAAGAAAAAGCAAATCC  3000
 Q  Q  Q  Q  I  A  M  G  Y  Q  S  M  L  R  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  Q  P  S  S  T  M  T  K  K  K  K  Q  I  H    599
ATTCGTTCAATAATAACAAGAGTTTATCATCTAAAGGCGGTGGCATTACCAAGAAATCACACGATAATAATAATCATAGTAAAGTTAAAACTCTGTACGGATCATTAAATGATGTTGTGA  3120
 S  F  N  N  N  K  S  L  S  S  K  G  G  G  I  T  K  K  S  H  D  N  N  N  H  S  K  V  K  T  L  Y  G  S  L  N  D  V  V  N    639
ATTCCAAAGTCACCAAAGTGATCAATAAAGAAGAAGTAAAACAGTCACAAACATAGATGAATAGATACAGAGAGAAAGAGAAAGAAAGAAAAAAACTAAGTAAATGGGATTCAAATAAAT  3240
 S  K  V  T  K  V  I  N  K  E  E  V  K  Q  S  Q  T  X  M  N  R  Y  R  E  K  E  K  E  R  K  K  L  S  K  W  D  S  N  K  F    679
TTATTTTGTTTGTCGTTATCTATGCATTATGTTTATGGACTTGTTTTATTATGATTTATTTTTAAAGGAATATATACTTGGTTTGATTAATTCATTATTACTTCATACCCTTTATTATTT  3360
 I  L  F  V  V  I  Y  A  L  C  L  W  T  C  F  I  M  I  Y  F  *                                                                699
TTTTTTGGGCCTTTTCTTCCTTTTTTTCTTCCATCGTCGTCGTCTTTTTTTTTCATTTTCATTTCTTTGGACAAGATCATTAGATTAGATTTTCATTATTTTCAATGCATGTGTTATATT  3480
TTTGGTATTTTTAGTCATTTTAGTATGGTTTATTTTTCATTTTTTCCTCTTTCAATATCGATGGTTGTTACTTTTTACTAGAGGGAACATACAGTTTATTTATTTATGCCTTATAGGTCT  3600
TCTTTTTTTTAATTTCTATTCCGTTATTTACTATGCTAATGTTGCTTGTCTTTTAAGGATTTCAATTTTATATTATCAACAATTTTTAAATTACTCTTAGTATTCTATCTTTATTGAATG  3720
AACACATTCATGATATACAGTATATAAAACTT                                                                                        3752
```

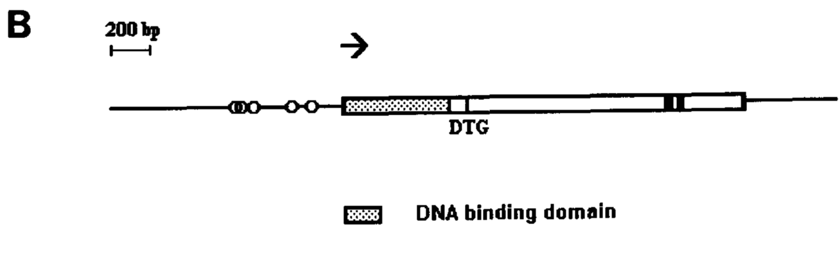

FIG. 1. *A*, nucleotide and deduced amino acid sequence of the *ACPR* gene. Sequence of the coding strand is shown with nucleotide and amino acid coordinates indicated on the *right*. Potential TATA box and CCAAT box sites are indicated. The predicted amino acid sequence is shown in standard single-letter code starting with a translation initiation codon at nucleotide position 1206 bp and extending till a termination codon at 3302 bp (indicated by a *star*). Within the coding region the potential active site for acid proteinases is indicated in *boldface print*. The amino acid *X*

1984), and hybridization was performed under conditions of high stringency according to standard protocols (Sambrook *et al.*, 1989). *Eco*RI and *Kpn*I inserts of two overlapping clones were subcloned in both orientations into pTZ18U plasmid (U. S. Biochemical Corp.).

DNA Sequence Analysis and Interpretation—The subclones in pTZ18U were trimmed to smaller sizes by unidirectional exonuclease III deletions (Henikoff, 1987). DNA sequencing of the overlapping deletion subclones was performed by the dideoxy chain termination method (Sanger *et al.*, 1977) on double-stranded DNA, using a Sequenase 2.0 kit (U. S. Biochemical Corp.) and M13 universal and reverse sequencing primers. Computer-assisted sequence analyses and comparisons were done using programs of the PC/Gene sequence analysis package (Intelligenetics).

Expression of Acprp in E. coli—The complete coding sequence (Acprp-(1–699)) and the NH_2-terminal DNA binding domain (Acprp-(1–230)) were amplified by PCR (Saiki *et al.*, 1988) and cloned into maltose binding protein (MBP) fusion expression vector, pMAL-c (Guan *et al.*, 1987), using the *Stu*I and *Stu*I/*Xba*I sites, respectively. PCR reactions were carried out using primers 1 and 2 for the complete coding region and primers 1 and 3 for the DNA binding domain (primer 1, 5′ CCTATGTCAATTACTAAAACATAC 3′; primer 2, 5′ CTAATCTAATGATCTTGTCCA 3′; and primer 3, 5′ GCTCTAGAAGTAGAGCTTGTG 3′). The *Xba*I site (underlined) in primer 3 was used for directional cloning. PCR reactions were done with *Taq* DNA polymerase under conditions recommended by the manufacturer (Perkin-Elmer Corp.), and amplification was carried out for 30 cycles (1 min at 94 °C, 1 min at 50 °C, and 1 min at 72 °C). The recombinant constructs were amplified in *E. coli* strain DH5α. All cloning procedures were carried out as described (Sambrook *et al.*, 1989).

For expression of Acprp-(1–699) and Acprp-(1–230) as fusion proteins, recombinant *E. coli* DH5α cells were grown at 37 °C to an A_{600} of 0.5. The expression was induced with 0.4 mM isopropyl-1-thio-β-D-galactopyronoside for 2 h. Cells were harvested and resuspended in lysis buffer (10 mM sodium phosphate, pH 7.0, 30 mM sodium chloride, 0.25% Tween 20, 10 mM EDTA, 10 mM EGTA, 10 mM β-mercaptoethanol). The suspension was frozen and thawed three times, sonicated for 2 min (30-s bursts with chilling on ice) to reduce viscosity, and centrifuged at 9,000 × *g* for 20 min to remove the debris. The clarified supernatant was used as an extract in mobility shift assays.

For Western blot analysis, cell pellet from 0.5 ml of induced culture was resuspended in 50 µl of 1 × SDS-polyacrylamide gel electrophoresis sample buffer; 15 µl of sample was loaded onto a 10% discontinuous SDS-polyacrylamide gel, and after electrophoresis, proteins were blotted onto nitrocellulose membrane by electrotransfer. The membranes were probed with affinity-purified secretory acid proteinase antibodies as described (Ganesan *et al.*, 1991).

Electrophoretic Mobility Shift Assays—The probes used for electrophoretic mobility shift assays were the transcriptional control elements of Ty1 and *STE2* (Errede and Ammerer, 1989). End-labeled double-stranded probes were prepared according to standard protocols (Sambrook *et al.*, 1989) and purified from unincorporated nucleotides by passing through a spun column. In the protein-DNA binding reactions, 20 µg protein from *E. coli* extract (described above) was incubated for 15 min at room temperature in a 20-µl reaction mixture containing 20 mM Tris-Cl, pH 8.0, 40 mM sodium chloride, 4 mM magnesium chloride, 1 mM dithiothreitol, 5% glycerol, 0.5 µg of poly(dI-dC), and 0.5 ng of ^{32}P-labeled probe DNA. Samples were electrophoresed at room temperature on a 5% polyacrylamide gel using 0.5 × TBE (89 mM Tris, 89 mM borate, 2.4 mM EDTA) as the buffer. Gels were dried and autoradiographed. For competition experiments the conditions were exactly as above, except that specific and nonspecific competitor DNA (pUC19) as included in the mixture (in the amounts detailed in the figure legends) prior to addition of the protein. Protein concentration in the extract was measured using the Bio-Rad assay reagent with bovine γ-globulin as the standard.

Construction of Hybrid Genes—Recombinant PCR (Higuchi *et al.*, 1988) was used to generate hybrid genes wherein either the DNA binding domain or the entire coding region beyond 21 amino acid residues of *STE12* was replaced with the corresponding region from *ACPR*. The primers used for primary PCR were: SF1, 5′ GAAGATATCCCTAACAAAGCGGAA 3′; SR2, 5′ CTTAATGATTCTTCAACTTCGCCC 3′; PF3, 5′ GTTGAAGAATCATTAAGACTAATTG 3′; PR4, 5′ AAAAAGATAATGCTGGTTCTCGAT 3′; SF5, 5′ ACCAGCATTATCTTTTTCATATGATTC 3′; and SR6, 5′ ATTGTTTAATATTAAATTCTGGCCCG 3′. The 5′ seventeen nucleotides of SR2 and PF3, PR4, and SF5 are complementary. The *STE12* fragments were amplified from plasmid pSY2 and the *ACPR* fragments from the genomic DNA of *C. albicans* strain SC5314. The fragments amplified were a 3.02-kb fragment of the entire *STE12* gene using SF1 and SR6, a 986-bp fragment of the *STE12* promoter up to the 34th amino acid using SF1 and SR2, a 1.5-kb fragment of the carboxyl-terminal 493 amino acids using SF5 and SR6, a 483-bp fragment of the DNA binding domain of *ACPR* from amino acid 21 to 182, and the entire *ACPR* gene (described earlier). Primary PCR reactions (100 µl) were done for 25 cycles comprising 94 °C for 1 min, 45 °C for 2 min, and 72 °C for 2 min for the first five cycles; the annealing temperature was increased to 50 °C for the subsequent 20 cycles. The 3.02-kb fragment was blunt ended with Klenow polymerase and cloned into the *Sma*I site of pBluescript II KS (Strategene) to give the STE12-pBSKS construct. The other products were gel purified using a Gene clean II kit (Bio 101, Inc.), and 30-ng templates were combined with primers for secondary PCR as follows: 986- and 483-bp fragments and primers SF1 and PR4 to give product 1, and 483-bp and 1.5-kb fragments and primers PF3 and SR6 to give product 2. Secondary PCR reactions (50 µl) were performed for 30 cycles under the same conditions as primary PCR but for the annealing temperature, which was 35 °C for the initial 5 cycles and 40 °C for the final 25 cycles.

The secondary PCR product 1 was digested with *Afl*II/*Eco*RI (to give a 757-bp fragment) and combined with product 2 digested with *Eco*RI/*Nco*I (994-bp fragment). The STE12-pBSKS construct was digested with *Afl*II/*Nco*I, and the vector backbone with a part of the STE12-pBSKS construct was gel purified. The three products (STE12-pBSKS *Nco*I/*Afl*II, *Afl*II/*Eco*RI, *Eco*RI/*Nco*I) were ligated and transformed into *E. coli* strain DH5α. The recombinant construct, HY1 (for hybrid-1), contained the DNA binding domain of *ACPR* gene in the *STE12* gene. For the construction of hybrid-2 (HY2), the HY1 construct was digested with *Bst*BI/*Bam*HI, and the backbone was gel-purified. A 1.96-kb *Bst*BI/*Bam*HI fragment was obtained from the PCR-amplified *ACPR* gene. The fragments were ligated and transformed into *E. coli* strain DH5α. The fidelity of the PCR reactions and continuity of the reading frames were confirmed by sequencing.

The hybrid genes *HY1*, *HY2*, and *STE12* were then cloned as *Hin*dIII/*Spe*I fragments into the *Hin*dIII/*Xba*I sites of YEplac 195 and YCplac33 plasmids to allow the expression of the genes in high and low copy number plasmids (Geitz and Sugino, 1988). The constructs were amplified in *E. coli* DH5α. The constructs were transformed into *S. cerevisiae ste12* mutants, SF167–5a and SF167–1c, by the lithium acetate procedure (Dunn *et al.*, 1984).

Mating Assays—Qualitative mating assays were performed by patch assays using a modified drop-overlay method (Spencer and Spencer, 1988). Cells were grown in SD medium with appropriate amino acid supplements to 10^7 cells/ml. After short centrifugation, the pellet from 1 ml of cell culture was resuspended in SD medium (100 µl). 10^6 cells were mixed with an equal number of cells of the tester strains (DC14 for α mating and DC17 for **a** mating), spotted on SD plates, and incubated until diploid colonies appeared.

Quantitative mating assays were performed as described (Sprague, 1991). Briefly, cultures of the transformants and the tester strains were grown to 10^7 cells in SD medium with appropriate supplements. 10^7 cells of the appropriate tester strain were mixed with 2×10^6 cells of the transformants. The cells were collected on a 0.45-µm filter and incubated on a YPD plate for 5 h at 30 °C. The cells from each of these filters were resuspended in 1 ml of SD medium. Several dilutions were plated in duplicate on SD-minimal plates to determine the number of diploids and on SD plates containing appropriate supplements to determine the total number of cells. The mating efficiency is the number of cells that mated divided by the total number of cells of the transformant in the mating reaction. The values are the mean of two independent experiments.

corresponds to the translation of stop codon TAG as a sense codon. Five short sequences homologous to the pheromone response elements (one copy of a consensus PRE and four sequences that contain 6/7 matches of the consensus) in the presumed promoter region are *underlined* or *overlined*. *B*, schematic representation of the *ACPR* gene. The *boxed* region represents the coding region, and the *thin lines* represent the 5′- and 3′-noncoding regions. The *arrow* indicates the direction of transcription. The *stippled* area represents the DNA binding domain of the encoded protein. The potential active site for acid proteinases (Asp-Thr-Gly) is marked. Two stretches of polyglutamine are also shown. *Circles* in the upstream region represent the sequences similar to pheromone response elements.

A

```
Acprp   MSITKTYNGDPTSL---------------VPTQLVKESLRLIEDLKFFLATA   37
Ste12p  MKVQITNSRTEEILKVQANNENDEVSKATPGEVEESLRLIGDLKFFLATA   50
Steklp  MAGSIVKTEDISSISGRGDTSQ------SPEEVEESLRLIEDLKFFLATR   44

Acprp   PANWQENQVIRRYYLNHDEGFVSCVYWNNLYFITGTDIVRCIVYKFEHFG   87
Ste12p  PVNWQENQIIRRYYLNSGQGFVSCVFWNNLYYITGTDIVKCCLYRMQKFG  100
Steklp  PANWQENQVIRRYYLSNDEGFVSCVFWNNLYYITGTDIVRCCAYRMQKFG   94

Acprp   RKIIDRKKFEEGIFSDLRNLKCGADAILEPPRSEFLEFLFKNSCLRTQKK  137
Ste12p  REVVQKKKFEEGIFSDLRNLKCGIDATLEQPKSEFLSFLFRNMCLKTQKK  150
Steklp  REIVERKKFEEGIFSDLRNLKCGIDATLEKPKSDLLAFLYKNMCLKTQKK  144

Acprp   QKVFFWFNVPHDKLMADALERDLKKEKMGQRPTTMAHREPALSFHYDESS  187
Ste12p  QKVFFWFSVAHDKLFADALERDLKRESLNQPSTTKPVNEPALSFSYDSSS  200
Steklp  QKVFFWFSVPHDRLFADALERDLKRSSTGSQPTTRAVSEPALSFRWEANS  194

Acprp   --SLYTQLGKHMETQKRINDAATSSTSNTATTLTDTGVSSGLNNTTSGGG  235
Ste12p  DKPLYDQLLQHLDSRRPSS--TTKSDNSPPKLESENFKDNELVTVTNQPL  248
Steklp  DVSLYDQITNHVDSQRTDSRPSTVGAQDVTQKQNNRVTDTPVSASKDVEP  244

Acprp   SDSATSTHNNNEASTK------PSNGSEKSSPEYTTTARGRDEFGFLNEA  279
Ste12p  LGVGLMDDDAPE-----------SPSQINDFIPQKLIIEPNTLELNGLTEE  288
Steklp  FESNVVEDEVQIVDNNKMCYGLPHSESNNYVPQQLIVPQSDLERNELTNE  294

Acprp   --------TPSQYKANSDYEDDFPLDYINQTTQNSEDYITLDANYQAGSY  321
Ste12p  -----TPHDLPKNTAKGRDEEDFPLDYFPVSVEY----------------  317
Steklp  FDELNADLKPSDILTSNQEDDDFPLDYFPVEIEYSQTSMDSSLHMASQGA  344

Acprp   ANMIEDNYDSFLDATLFIPPSLGVPTGTAATATTSNQVAFNDEYLIEQAQ  371
Ste12p  --PTEENAFDPFPPQAFTPAAPSMPISYDNVNERDSMPVNSLLNRYPYQL  365
Steklp  KHPSQMMFYEDMDGMMLGPKYPISAGIYEDPFFREEMAASN---------  385

Acprp   PIRTPLPPISSSTISGLLQPK---SAAKFFSLQSANGGEEFFPAYQNDPS  418
Ste12p  SVAPTFPVPPSSSRQHFMTNRDFYSSNNNKEKLVSPSDPTSYMKYDEPVM  415
Steklp  ASKYMMPPPMSATRAHFMTNREYYSKSREGKKHGKSGRQHGP--------  427

Acprp   TANAGFVPPISAKYATQFATRQVATPTYIKAIPQTGAAAATGNGGQPQQY  468
Ste12p  DFDESRPNENCTNAKSHNSGQQTKQHQLYSNNFQQSYPNGMVPGYYPKMP  465
Steklp  --DDGRRDYDNENVGSSPSENQNPNEELDATENNQDADDNRSPDIYP--P  473

Acprp   YDQATGNAFYPAEIPVLYNVVHPESEYWTNNSGAVATTAAAAAPMYDASG  518
Ste12p  YNPMGGDPLLDQAFYGADDFFFPPEGCDNNMLYPQTATSWNVLPPQAMQP  515
Steklp  TDAMANDEYIPAYNRRMQMSESLVHPYTGMMLNPYMFCNMLAVDPSLNMG  523

Acprp   FPIPINQSYMVMNEHEMVPYQYMNSNGAMIGMIPPHQQQQQQQQQIAMGY  568
Ste12p  APTYVGRPYTPNYRSTPGSAMF--PYMQSSNSMQWNTAVSPYSSRAPSTT  563
Steklp  VNPVVDPFYGQNHSMDVAHDVY--SPQEVVYPSNYRSTPKAAYFNMRSPY  571

Acprp   QSMLRQQQQQQQQQQQQQPSSTMTKKKKQIHSFNNNKSLSSKGGGITKKS  618
Ste12p  AKNYPPSTFYSQNINQYPRRRTVGMKSSQGNVPTGNKQSVGKSAKISKPL  613
Steklp  GRNFPPPSAMNPYYTPYHRRQPSSATRRYFS-------------------  602

Acprp   HDNNNHSKVKTLYGSLNDVVNSKVTKVINKEEVKQSQTXMNRYREKEKER  668
Ste12p  HIKTSAYQKQYKINLETKARPSAGDEDSAHPDKNKEISMPTPDSNTLVVQ  663
Steklp  ------------KANLINRKSKSPPRKNVVSKPSHKAPKKVNRTRHATSMH  641

Acprp   KKLSKWDSNKFILFVVIYALCLWTCFIMIYF   699
Ste12p  SEEGGAHSLE------VDTNRRSDKNLPDAT   688
Steklp  VGLGGKDSGN------TSSESRDDSTKEDSN   666
```

B

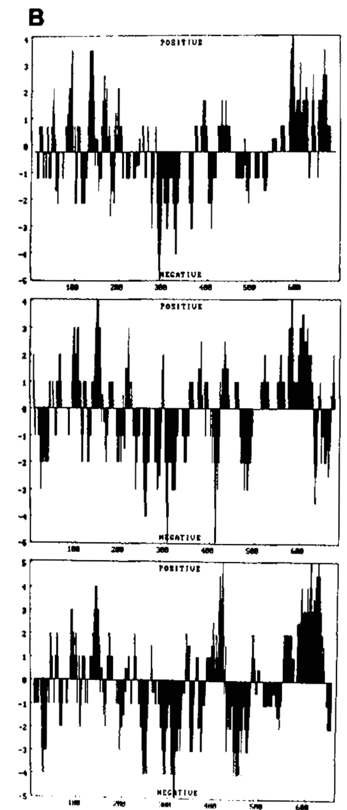

FIG. 2. *A*, the NH_2-terminal sequence of Acprp is homologous with that of *S. cerevisiae* Ste12p and *K. lactis* Ste12p (Steklp). The deduced amino acid sequences were aligned using the CLUSTAL multiple sequence alignment program of the PC/Gene sequence analysis package. Gaps were introduced for optimal alignment. Identical residues are indicated by a *star* and similar residues by a *dot*. *B*, comparison of charge distribution profiles of Acprp (*top panel*), Ste12p (*middle panel*), and Steklp (*bottom panel*). The plot was prepared using the Genepro version 5.0 sequence analysis package. The analysis of net charge is shown with a window of 10 residues. The following amino acids are assumed to be charged: Asp and Glu are -1.0, Lys and Arg are +1.0, His is +0.5.

RESULTS

Isolation of ACPR Gene—We have previously isolated a 1.3-kb genomic clone from a λ gt11 genomic expression library using affinity-purified proteinase antibodies (Ganesan *et al.*, 1991). Using this 1.3-kb DNA as probe we have screened a genomic library of *C. albicans* strain SC5314 in the phage EMBL3 and identified 11 positive clones from 2×10^4 phage isolates of the primary library. Restriction mapping of the clones showed that they could be classified into overlapping clones. Unidirectional nested deletion subclones covering the gene and its flanking regions were sequenced on both the strands. The complete nucleotide sequence of *ACPR* gene is shown in Fig. 1. The 3452-nucleotide-long sequence has a single large open reading frame beginning with an ATG at nucleotide 1206 and terminating with a TAA codon at nucleotide 3303. Within the open reading frame at amino acid position 655 is a TAG codon; the TAG codon is reported to be a sense codon in *C. albicans* (Santos *et al.*, 1990). The open reading frame could encode a protein containing 699 amino acid residues with an estimated molecular mass of 77 kDa. In the presumed promoter region, sequences with consensus to the TATA and CCAAT boxes are indicated. Since the encoded protein is antigenically related to acid proteinase, we refer to it as acid proteinase-related (*ACPR*) gene.

ACPR Gene Product Has a Modular Structure—The gene product, Acprp, showed several interesting features. Amino acid residues from 220 to 222 are Asp-Thr-Gly, which match with the consensus sequence (Asp-Thr/Ser-Gly) of active sites of acid proteinases (Blundell *et al.*, 1990). However, no other homology with sequences of acid proteinase in the data bases was demonstrable. Viral acid proteinases exist as homodimers in which one active site is present on each polypeptide (Blundell *et al.*, 1990). Close to the carboxyl terminus of Acprp, there are two stretches of polyglutamine. Such regions have been characterized as important for transcriptional activation in various factors including AP-2 (Williams and Tjian, 1991), Oct-2 (Clerc *et al.*, 1988; Muller *et al.*, 1988; Scheidereit *et al.*, 1988), and Spl (Courey and Tjian, 1988).

The amino acid sequence of Acprp was compared with the sequences in the Swiss-Prot sequence data base, using the FASTA program (Pearson and Lipman, 1988). The NH_2-terminal 200 amino acid residues showed similarity with the corresponding regions of Ste12p of *S. cerevisiae* (ScSte12) and *Kluyveromyces lactis* (Steklp) (Yuan *et al.*, 1993). Ste12p is a DNA binding protein involved in the mating. Amino acid residues 40–204 of ScSte12p represent the minimal region sufficient for DNA binding. In this region Acprp shares 73% identity with Steklp and 71% identity with ScSte12p. No other region of significant homology was found. Fig. 2*A* shows an alignment of amino acid sequences of Acprp, ScSte12p, and SteKlp. Despite the lack of substantial similarity beyond the DNA binding domain, the plot of charge distribution shows similar profiles (Fig. 2*B*). The region encompassing amino acid residues 200–458 of Acprp is highly acidic (net charge of -14) and rich in Ser, Thr, and Tyr residues (28%) as well as Pro (7%). Charged clusters are a common feature among nuclear transcription factors (Brendel and Karlin, 1989). Prosite (Bairoch, 1992) detected

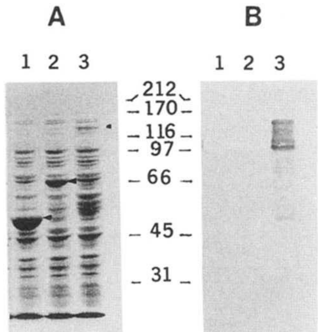

FIG. 3. ***ACPR* encodes a protein antigenically related to secretory acid proteinase.** *ACPR*-encoded products were expressed in *E. coli* strain DH5α. Lysates were prepared from cells harboring: *1*, pMAL-c plasmid; *2*, MBP-Acprp-(1–230); and *3*, MBP-Acprp-(1–699), as described under "Experimental Procedures," and resolved by discontinuous SDS, 10% polyacrylomide gel electrophoresis. The gel stained with Coomassie Brilliant Blue R-250 is shown in *panel A*. The *arrowheads* indicate that position of MBP (*lane 1*) or MBP-Acprp fusion proteins (*lanes 2* and *3*). The same samples run parallel were also analyzed by Western blot (*B*) using affinity-purified anti-proteinase antibodies. The positions and size (in kDa) of molecular mass markers are also shown.

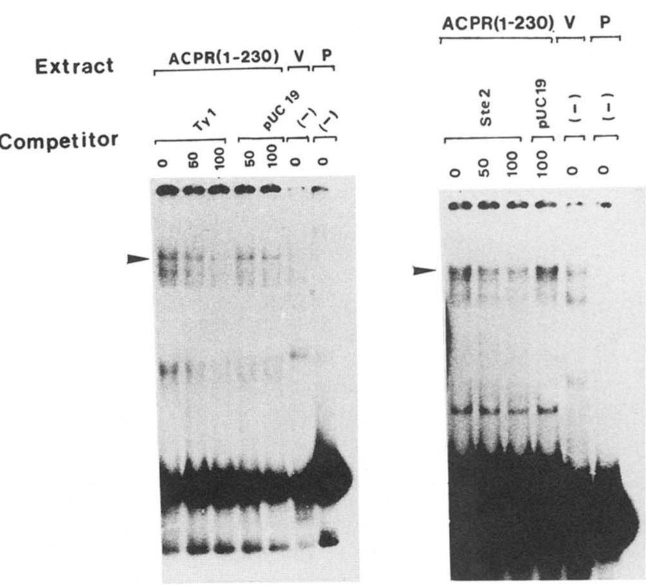

FIG. 4. **DNA binding activities of the Acprp expressed in *E. coli*.** The DNA binding activity of the MBP-Acprp-(1–230) fusion protein was determined by electrophoretic mobility shift assays. *E. coli* extracts were from cells expressing Acprp-(1–230) as a fusion with MBP and those expressing MBP only (*V*). Probe DNA in the binding reactions was either 0.5 ng of end-labeled *Ty1* (*A*) or *STE2* (*B*). Ty1, STE2, or pUC19 DNA were added as competitors at the indicated (0–100-fold) molar excess. *P* indicates a reaction that did not receive any extract. DNA-protein complexes were resolved by gel electrophoresis as described under "Experimental Procedures." *Arrowheads* indicate the complex formed.

eight potential phosphorylation sites in this region. These features suggest that Acprp could be a DNA binding protein.

ACPR Can Bind Pheromone Response Elements in Vitro—Since *C. albicans* lacks a sexual cycle, it was of interest to investigate the role of *ACPR* in light of its homology to *S. cerevisiae STE12* gene. To study the functions of the different regions of Acprp, the nucleotide sequence corresponding to either the NH_2-terminal 230 amino acid residues or the entire coding region of *ACPR* gene was amplified by PCR and cloned into pMAL-c expression vector (see "Experimental Procedures"). The sizes of the induced MBP fusion proteins were confirmed on immunoblots using anti-MBP antibodies (data not shown) and affinity-purified proteinase antibodies (Fig. 3). The MBP-Acprp (1–699 amino acid residues) fusion protein showed several immunorelated bands (Fig. 3*B*, *lane 3*), which could be degradation products. The MBP-Acprp-(1–230) did not show any reactivity to proteinase antibodies (Fig. 3*B*, *lane 2*) indicating that the antigenic epitopes are in the COOH-terminal two-thirds of Acprp. Proteinase activity could not be detected in *E. coli* extracts expressing the fusion proteins (data not shown). In another set of experiments, the Acprp-(1–230) and Acprp-(1–699) were expressed in *E. coli* by cloning into pET3c vector (Studier *et al.*, 1990) to assay for proteinase activity. But the overexpressed products were insoluble and could not be assayed. Recently, an adenovirus proteinase was shown to require DNA and a peptide as cofactors for its activity (Mangel *et al.*, 1993). It is possible that Acprp also requires DNA and other cofactors to function as a proteinase. These aspects can be studied if the *E. coli*-expressed Acprp is solubilized and purified.

The DNA binding activities of MBP-Acprp-(1–230) to the transcriptional control elements of Ty1 (Fig. 4*A*) and the α-pheromone receptor gene *STE2* (Fig. 4*B*), which contain the PREs, were checked by gel mobility shift assays (Errede and Ammerer, 1989). The specificity of the retarded complexes (indicated by *arrowheads*) was established by competition with homologous unlabeled probe (Ty1 or *STE2* UAS) and pUC19 DNA as a nonspecific competitor. In control experiments, yeast extracts from a wild type *S. cerevisiae* strain (EG123) overexpressing Ste12p and an *ste12* mutant strain (SF167–5a) were used (data not shown). These *in vitro* data imply that the NH_2-terminal 230 amino acids of Acprp can recognize and bind PRE elements like Ste12p of *S. cerevisiae*.

ACPR Can Complement STE12 in S. cerevisiae—The fact that Acprp can bind to sequences that are targets for Ste12p suggested that Acprp may complement the mating defect of an *ste12* mutant. To see if sequence homology could translate into functional complementation, domain swap experiments were done with chimeric gene constructs between *STE12* and *ACPR*. In the hybrid-1 construct (*HY1*), the DNA binding domain of *STE12* (amino acid residues 34–194) was replaced by the homologous region of *ACPR* (amino acid residues 21–182). In the hybrid-2 construct (*HY2*), the *ACPR* coding region (amino acid residues 21–699) was cloned after the 33rd amino acid residue of *STE12* gene. Both constructs were expressed from the *STE12* promoter. In addition, the *STE12* gene was expressed from its own promoter. All of these constructs were carried either on a high copy number (2-μm) plasmid or a low copy number (centromere-containing) plasmid and transformed into **a** or α *S. cerevisiae* strains having *ste12* mutation. The ability of each of these constructs to restore mating was used to study the function of *ACPR*. Mating was assessed by qualitative and quantitative assays (Fig. 5). As expected from sequence homology, *HY1* expression could correct the mating defect in the **a** cells. In fact, the mating efficiency was 3-fold higher than that conferred by the wild type Ste12p present in a high copy plasmid in the **a** cells. The α cells showed reduced mating efficiency when *HY1* was expressed from a high or low copy plasmid. In contrast, the mating efficiency of *HY2* construct was 25% of *S. cerevisiae STE12* on a high copy plasmid and 6.5% on a low copy plasmid in the **a** cells. In the α cells, compared with *STE12*, *HY2* showed 100-fold reduced mating efficiency on low copy plasmids and 75-fold reduced mating on high copy plasmids. These results suggest that the *ACPR* gene of *C. albicans* can complement the *STE12* gene of *S. cerevisiae* and function in the mating pathway. However, the activation domain of *ACPR* functions poorly, particularly in the α cells.

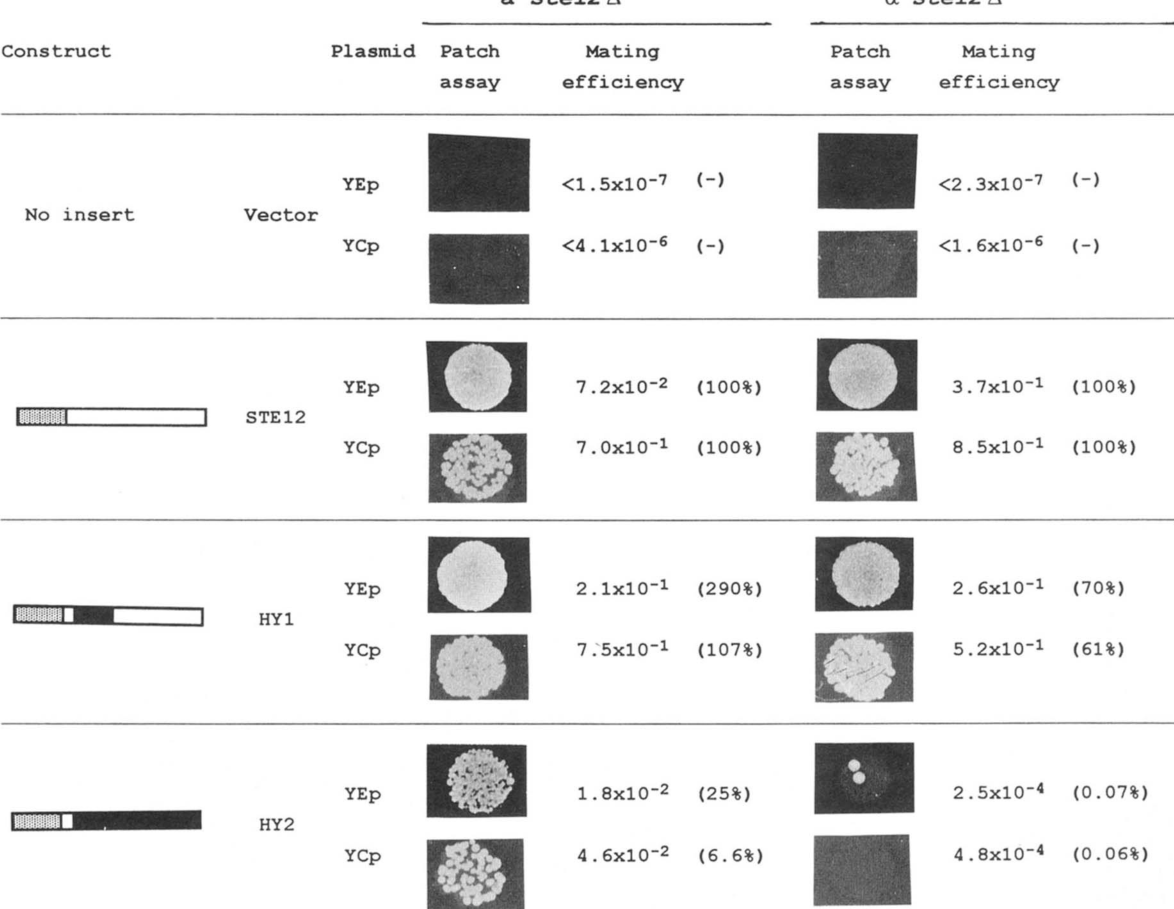

FIG. 5. **Function of the chimeric Ste12-Acpr transcription factors in *S. cerevisiae*.** The ability of the chimeric constructs in high (*YEp*) or low (*YCp*) copy number plasmids to complement the mating defect of *ste12* mutant was assessed in **a** *ste12* or α *ste12 S. cerevisiae* strain. A schematic of the chimeric constructs is shown. The *open box* indicates the *S. cerevisiae STE12* coding region, the *filled box* represents the *ACPR* coding region, and the *shaded box* indicates the *STE12* promoter. Mating assays were performed by qualitative tests (patch assay), and the mating efficiency was determined by the quantitative method. The mating efficiency of constructs calculated as a percentage of the wild type *STE12* is indicated in *parentheses*.

DISCUSSION

We describe here the cloning and initial characterization of a transcription factor from the pathogenic yeast *C. albicans* that can support mating in *S. cerevisiae*. *C. albicans* is an imperfect fungus and is not known to have any haploid phase in its life cycle. *ACPR* was identified by reaction of the encoded protein with secretory acid proteinase antibodies. The antigenic epitopes are localized in the carboxyl-terminal two-thirds of Acprp. This product has a modular structure comprising a DNA binding domain followed by a negatively charged region, which also has a potential active site for acid proteinase and a domain rich in glutamine residues. The NH_2-terminal 230-amino acid region was expressed as a fusion protein with MBP in *E. coli*. This fusion protein was used to show that Acprp, like Ste12p, can bind the pheromone response elements present in the upstream control regions of pheromone-responsive genes (*Ty1* and *STE2*). In addition, chimeric constructs of *STE12-ACPR*, when transformed into *ste12* mutants, restored the mating defect. The ability to complement *ste12* mutation has been assayed on high and low copy plasmids in either the **a** cells or the α cells. The data indicate that the *ACPR* gene can restore the mating defect of *S. cerevisiae* **a** *ste12* mutant better than an α *ste12* mutant. A similar cell-type bias and poor complementation of mating defect in *S. cerevisiae* α cells by *K. lactis STE12* gene have been demonstrated (Yuan *et al.*, 1993). This bias has been attributed to the inability of the carboxyl-terminal sequences involved in transcriptional activation and induction to interact with α1 protein (Yuan *et al.*, 1993). Since the carboxyl-terminal sequences of Acprp are significantly different from those of *S. cerevisiae* Ste12p, a similar defect is likely to cause poor complementation in α cells.

The sequence similarity of a number of genes of *C. albicans* and *S. cerevisiae* suggests that the two organisms are evolutionary related (Boone *et al.*, 1991; Sadhu *et al.*, 1992; Scherer and Magee, 1990). It is possible that the elements of mating response pathway exist in *C. albicans*. Using a dominant negative selection scheme, four *C. albicans* genes have been isolated that interfere with the mating response pathway in *S. cerevisiae* (Whiteway *et al.*, 1992). These genes have not been shown to complement the counterparts in *S. cerevisiae* mating pathway. A G-protein α-subunit, which complements the mating defect and has homology with the *S. cerevisiae* G-protein α-subunit, has been reported (Sadhu *et al.*, 1992). Using the DNA binding domain of *ACPR* as a probe, we have identified distinct bands on Southern blots of various strains of *C. albicans* and several other *Candida* species.[2] This suggests that genes related to *ACPR* are present in several *Candida* species including *C. albicans*, and it is unlikely that they do not have any biological role. We have similarly identified a putative *STE7* homologue in *C. albicans* strain SC5314.[2] These results

[2] K. Malathi, K. Ganesan, and A. Datta, unpublished results.

and the fact that the proteins involved in the first step (G-protein α-subunit) and the final step (*STE12*-like) of the mating signal transduction pathway exist in *C. albicans* strengthen the possibility that a sexual phase may exist in *C. albicans*. It is conceivable that in *C. albicans* the mating efficiency is low and has gone undetected.

Under some conditions, *S. cerevisiae* diploids show a filamentous growth (Gimeno *et al.*, 1992). Recently, *STE12* has been shown to be required for filamentous growth besides its role in pheromone-mediated signal transduction (Liu *et al.*, 1993). Dimorphism of *C. albicans* is known to play a significant role in pathogenesis (Datta *et al.*, 1989; Odds, 1988). Due to its structural and functional similarity with Ste12p, it is likely that Acprp plays a role in the developmental switch between yeast and mycelial forms of *C. albicans*. We are currently in the process of constructing mutants defective in the *ACPR* gene through directed gene disruption (Kelly *et al.*, 1987). These mutants will be used to study the role of this gene in morphogenesis and pathogenesis.

Acknowledgments—We thank D. Kirsch for the *C. albicans* strain SC5314, S. Fields for *S. cerevisiae* strains SF167–5a and SF167–1c and the plasmid pSY2, B. Errede for Ty1 and STE2 probes, J. Strathern for DC14 and DC17 tester strains, and D. R. Geitz for YEplac195 and YCplac33 plasmids. We also thank the Bioinformatics Center, Jawaharlal Nehru University, for the computer-assisted facilities. We thank K. Natarajan for valuable suggestions during the course of this work and for critical reading of the manuscript.

REFERENCES

Bairoch, A. (1992) *Nucleic Acids Res.* **20,** 2013–2018
Banerjee, A., Ganesan, K., and Datta, A. (1991) *J. Gen. Microbiol.* **137,** 2455–2461
Blundell, T. L., Lapalto, R., Wilderspin, A. F., Hemmings, A. M., Hobert, P. M., Danley, D. E., and Whittle, P. J. (1990) *Trends Biochem. Sci.* **15,** 425–430
Boone, C., Sdicu, A. M., Laroche, M., and Bussy, H. (1991) *J. Bacteriol.* **173,** 6859–6864
Brendel, V., and Karlin, S. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 5698–5702
Clerc, R. G., Corcoran, L. M., LeBowitz, J. H., Baltimore, D., and Sharp, P. (1988) *Genes & Dev.* **2,** 1570–1581
Courey, A. J., and Tjian, R. (1988) *Cell* **55,** 887–898
Datta, A., Ganesan, K., and Natarajan, K. (1989) *Adv. Microb. Physiol.* **30,** 53–88
Dolan, J. W., Kirkman, C., and Fields, S. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 5703–5707
Dunn, R., Szauter, P., Pardue, M. L., and Szostak, J. W. (1984) *Cell* **39,** 191–201
Errede, B., and Ammerer, G. (1989) *Genes & Dev.* **3,** 1349–1361
Feinberg, A. P., and Vogelstein, B. (1984) *Anal. Biochem.* **137,** 266–267
Ganesan, K., Banerjee, A., and Datta, A. (1991) *Infect. Immun.* **59,** 2972–2977
Gietz, R. D., and Sugino, A. (1988) *Gene (Amst.)* **74,** 527–534
Gimeno, C. J., Ljungdahl, P. O., Styles, C. A., and Fink, G. R. (1992) *Cell* **68,** 1077–1090
di Guan, C., Li, P., Riggs, P. D., and Inouye, H. (1987) *Gene (Amst.)* **67,** 21–30
Henikoff, S. (1987) *Methods Enzymol.* **155,** 156–165
Higuchi, R., Krummel, B., and Saiki, R. K. (1988) *Nucleic Acids Res.* **16,** 7351–7367
Holm, C., Meeks-Wagner, D. W., Fangman, W. L., and Botstein, D. (1986) *Gene (Amst.)* **42,** 169–173
Kelly, R., Miller, S. M., Kurtz, M. B., and Kirsch, D. R. (1987) *Mol. Cell. Biol.* **7,** 199–207
Liu, H., Styles, C. A., and Fink, G. R. (1993) *Science* **262,** 1741–1744
Mangel, W. F., McGrath, W. J., Toledo, P. L., and Anderson, C. W. (1993) *Nature* **361,** 274–275
McDonald, F., and Odds, F. C. (1983) *J. Gen. Microbiol.* **129,** 431–438
Muller, M. M., Ruppert, S., Schaffner, W., and Matthias, P. (1988) *Nature* **336,** 545–551
Odds, F. C. (1988) *Candida and Candidosis*, 2nd Ed., Bailliere Tindall Ltd., London
Pearson, W. R., and Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 2444–2448
Ruchel, R. (1981) *Biochim. Biophys. Acta* **659,** 99–113
Sadhu, C., Hoekstra, D., McEachern, M. J., Reed, S. I., and Hicks, J. B. (1992) *Mol. Cell. Biol.* **12,** 1977–1985
Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science* **239,** 489–491
Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
Santos, M., Colthurst, D. R., Wills, N., McLaughlin, C. S., and Tuite, M. F. (1990) *Curr. Genet.* **17,** 487–491
Scheidereit, C., Cromlish, J. A., Gerster, T., Kawakami, K., Balmaceda, C. G., Currie, R. A., and Roeder, R. G. (1988) *Nature* **336,** 551–557
Scherer, S., and Magee, P. T. (1990) *Microbiol. Rev.* **54,** 226–241
Spencer, J. F. T., and Spencer, D. M. (1988) *Yeast: A Practical Approach*, pp. 65–67, IRL Press at Oxford University Press, Oxford
Sprague, G. F. (1991) *Methods Enzymol.* **194,** 77–94
Staib, F. (1965) *Sabouradia* **4,** 187–193
Staib, F. (1969) *Mycopathol. Mycol. Appl.* **37,** 345–348
Studier, F. W., Rosenberg, A. H., Dunn, J. J., and Dubendorf, J. W. (1990) *Methods Enzymol.* **185,** 60–89
Whiteway, M., Dignard, D., and Thomas, D. Y. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 9410–9414
Williams, T., and Tjian, R. (1991) *Genes & Dev.* **5,** 670–682
Yuan, Y. O., Stroke, I. L., and Fields, S. (1993) *Genes & Dev.* **7,** 1584–1597