Isolation and sequence analysis of the gene for translation elongation factor 3 from *Candida albicans*

Kristi K. Myers, William A. Fonzi and Paul S. Sypherd*

Department of Microbiology and Molecular Genetics, College of Medicine, University of California, Irvine, CA 92717, USA

Received December 2, 1991; Revised and Accepted March 9, 1992

EMBL accession no. Z11484

ABSTRACT

Elongation factor 3 (EF-3) is a unique and essential component of the translational system in fungi. The gene, CEF-3, encoding elongation factor 3 has been isolated from the dimorphic fungus *Candida albicans*. A heterologous gene probe containing the coding region of the EF-3 gene from *Saccharomyces cerevisiae* (YEF-3) was used to screen three *Candida albicans* genomic DNA libraries. The nucleotide sequences of four partial clones were determined and combined for a full-length of 3,671 base pairs (bp). A continuous open reading frame (ORF) of 3,147 bp encoding a predicted protein of 1,049 amino acids and M_r of 116,739 daltons has been identified. A transcript of 3,400 nucleotides is seen in Northern blot hybridization of *Candida albicans* total RNA using a CEF-3 gene probe. The single locus CEF-3 gene maps to chromosome 5 in the genome. Comparison of the deduced amino acid sequences of CEF-3 and YEF-3 shows 77.6% identity. A higher degree of identity, 86.5%, is found when comparing the carboxy-terminal portions of the two proteins. At the nucleotide level, comparison of the coding regions of the two genes exhibit 79% identity while the upstream and downstream regions show 46% and 40% identity, respectively.

INTRODUCTION

The study of translation in *Saccharomyces cerevisiae* led to the discovery of elongation factor 3 (EF-3) by Skogerson and colleagues. This factor is required for *in vitro* protein synthesis by yeast ribosomes along with elongation factors 1 and 2. While EF-1 and EF-2 are functionally interchangeable with their higher eukaryotic counterparts, EF-3 is unique to fungal systems (1). EF-3 has been isolated from *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* and found to be a monomeric protein of a molecular weight of 125,000 daltons (2, 3). EF-3 is essential to the elongation step of translation (4–7) and exhibits a ribosome-dependent ATPase activity (8, 9).

Using antibodies prepared to the *Saccharomyces* EF-3 protein,

Western blot analysis has shown EF-3 to be present in many fungi such as *Neurospora*, *Aspergillus*, *Mucor*, *Candida albicans* and other *Candida* species (8, 10, 11). As is the case in *Saccharomyces*, because of the requirement for EF-3, *Candida* cell-free poly(U)-dependent elongation is an ATP-dependent process. Moreover, when purified *Candida* EF-3 is added to *Saccharomyces* ribosomes lacking the soluble factor, poly(U)-dependent elongation can proceed (11).

Candida albicans is a dimorphic fungus that can grow either as a yeast or hyphal cell. It is a common and usually harmless commensal of warm-blooded animals. In the host with a functional immune system, *C. albicans* on rare occasion causes mild, superficial infections, but in the immunocompromised host it frequently causes life-threatening systemic infections. There is a high degree of correlation between virulence and the ability to undergo morphogenesis from yeast to hyphal cell type. Although yeast cells can be found in superficial infections, hyphal cells are the most predominant form in systemic candidosis. The number of immunocompromised hosts has risen in recent years due to the increased use of immunosuppressive therapies in cancer and transplant patients as well as the increased incidence of acquired immune deficiency syndrome (AIDS). This trend has emphasized the need for non-toxic and specific drug therapies for candidal infections (12). The presence of EF-3 in *Candida albicans*, in addition to the fact that EF-3 is a fungal-specific and essential translation factor, suggests that this factor could serve as a target for anti-fungal drugs.

Here we report the isolation of and nucleotide sequence for the EF-3 gene from *Candida albicans*. Our data indicate that CEF-3 is located on chromosome 5 and, in contrast to the gene for EF-1 α (13), is a single locus gene. The deduced amino acid sequence for the *Candida* gene exhibits a high degree of identity to that of the *Saccharomyces* gene. Conservation between the two EF-3 proteins in the region containing nucleotide-binding motifs and poly-lysine blocks (7) supports the idea that these features are involved in the ribosome-dependent ATPase activity. The relative lack of identity at the nucleotide level between the non-coding regions of the two genes indicates possible differences in the control of EF-3 gene expression between the two fungal species.

* To whom correspondence should be addressed

MATERIALS & METHODS

Hybridization Analysis of Genomic DNA, RNA and Chromosomes

Genomic DNA for Southern blot hybridization was isolated from *Candida albicans* strain SC5314 as described by Boeke and co-workers (14) and digested with the indicated restriction enzyme (Bethesda Research Laboratories, Boehringer Mannheim). The digested DNA (8–10 µg) was separated by electrophoresis in 1% agarose gels in TBE buffer and transferred to nitrocellulose filters by method of Southern (15). After prehybridization for 1 hr in hybridization buffer at the desired temperature, hybridization of the filters was carried out as described by Sundstrom and Sypherd (13). The 3.5 kb XbaI fragment from pYEF-3 (16) or the 981 bp EcoRI fragment from p3R4 (*C. albicans* genomic clone isolated in this work) were labelled (17, 18) with [α^{32} P]dGTP (Dupont-NEN) using a random-primer kit (United States Biochemical) for use as hybridization probes. Hybridization was accomplished with 5×10^6 cpm of labelled probe at temperatures of 37°C and 42°C for the *S. cerevisiae* and *C. albicans* probes, respectively. Washes were carried out at hybridization temperature with a solution of $2 \times \text{SSC}/0.1\%$ SDS once, followed by $0.2 \times \text{SSC}/0.1\%$ SDS twice, each for 15 min.

Total RNA for mRNA analysis was prepared from *C. albicans* SC5314 as described by Saporito and Sypherd (19). Yeast cells were grown at 25°C for 5 hr in Lee's medium, pH 4.5, while hyphal cells were grown for 5 hr at 37°C in Lee's medium, pH 6.5. The isolated RNA was separated on a denaturing formaldehyde-agarose gel, transferred to a nitrocellulose filter and hybridized as described above using the CEF-3 981 bp EcoRI fragment as the probe.

Chromosomes for hybridization analysis were isolated from *C. albicans* SC5314 and embedded in low-melting-point agarose for electrophoretic separation using the CHEF-DRII System (Bio-Rad Laboratories)(19). The chromosomes were separated in a $5.5'' \times 5''$ 1% agarose gel in $0.5 \times \text{TBE}$ according to the electrophoresis conditions described by Magee and coworkers (20), UV-irradiated and transferred to a charged nylon membrane (Micron Separations, Inc.) in 0.4 M NaOH for 24 hrs. The membrane was neutralized in 0.5 M Tris pH 7.5, dried and UV-irradiated to cross-link the DNA to the membrane. The 981 bp EcoRI fragment from p3R4 and a 5 kb XbaI fragment from pYK5, which contains *C. albicans* DNA of unknown function that maps to chromosome 5 (B. B. Magee, personal communication), were used for hybridization of the chromosome blot. The hybridization conditions used were the same as described above for genomic Southern blots.

Library Screening

Three lambda libraries of genomic DNA from *C. albicans* SC5314 were screened with the 3.5 kb XbaI YEF-3 probe described above. Two of the libraries contained DNA completely digested with EcoRI or HindIII in the lambda vectors 509 or 607 (21), respectively (22, 23). The third library, CA5314-GEM12, consisted of genomic DNA partially digested with SAU3A and cloned in lambda GEM-12 (Promega; M. F. Ypma-Wong, personal communication). Each library was plated in triplicate on LB agar plates such that each plate contained 2–5,000 plaques. Duplicate plaque lifts were prepared according to a procedure described by Carlock (24) and hybridized as described above except that the total counts used for 6 filters was 5×10^7 cpm. A less stringent wash of 3 times for 15 min with

Figure 1. Genomic Southern hybridization analysis with YEF-3. A Southern blot of total DNA from *Candida albicans* SC5314, digested with the indicated enzyme, was hybridized using the 3.5 kb XbaI YEF-3 probe as described in Materials and Methods. The positions and sizes of the HindIII-digested lambda DNA markers are shown.

$2 \times \text{SSC}/0.1\%$ SDS at room temperature was carried out on the hybridized filter lifts. Positive plaques were cut out of the plate and plaque-purified for re-screening. Inserts from final positive lambda recombinants were cloned into plasmid vectors for subcloning and sequencing.

DNA Cloning Methods

Bacterial strains DH5 α and NM522 were the host strains for cloning vector pIBI20 (International Biotechnologies, Inc.) while XL1-Blue was used for the pBluescriptSK⁺ vector (Stratagene). *E. coli* LE392 was the host strain used for lambda phage plating. Single-stranded DNA for sequencing was purified by procedures recommended by IBI and Stratagene. Plasmid DNA for sequencing was purified by an alkaline lysis method (25) while that for subcloning was purified by a quick boiling method (26). Standard recombinant DNA procedures have been described (27, 28). Restriction enzyme digestion was carried out according to the manufacturer (BRL, BM). DNA fragments were isolated by using the Gene Clean Kit (Bio 101) to purify the desired fragments out of agarose gels.

Nucleotide Sequence Determination and Sequence Analysis

Inserts from lambda clones and subsequent subclones were cloned into either pIBI20 or pBluescriptSK⁺ for DNA sequence determination by the dideoxy chain termination method (29) using

Figure 3. Expression of CEF-3 during yeast and hyphal growth. A Northern blot of total RNA from yeast and hyphal *Candida albicans* SC5314 cultures was hybridized with the 981 bp EcoRI CEF-3 probe as described in Materials and Methods. Abbreviations: Y=yeast, H=hyphae.

A

```
-306 TGTTTAATTAAGATATGAGAAAGAGTGGCCAGAGAGAGAGAGAGAGAAAGG
-256 CCAAAGGAGGCCAAAAAATTTTTTCAGTTGTGATTCTCT
-206 TGTTCATCACTTTAATCTTTTTTTTTTGTCTTGGTGAGCACTAAC
-156 TAATTCAACCACTGTTTAAAGTAAACAAAAAATTTTTTTTCCTCT
-106 TCCTTTCATATAGAAAAATTTTTTTCCTTCTCTATTAATTTCTTTCTTTC
-56 TTTCTTCTCTCTCTCTCTCTCTCTTTTCTCTTTTATCAACAACATTT
-6 TGGAAA
```

B

3148 TAAGTTTGATCACTAAAGAGTTTTTCATTCCCTCCCCCACACCAATTAG

3198 AAGAGAGAGATAGAGAGGTATATTAGGTTTCAGGACAAGTTTATATACA

3248 TTACATTACATTACCTAACCTTCACTTAGCATAGTGTATGCTGCGCTACA

3298 TCATTGTGCTATTCTTACTAGTTTTGAATAGTATTTATATATGTTATTTCT

3348 TTGTCATTAATAAAATTTT

2.9 kb in the EcoRI lane and 1.5, 2.7 and 8.0 kb in the HindIII lane.

Two *Candida albicans* genomic DNA libraries, in lambda vectors 590 and 607 were screened under the same conditions used for the Southern hybridization described above. Three clones, 3R4, 5H3 and 6H7, were isolated and are depicted in Figure 2. The inserts from these clones, 1.0 kb EcoRI, 1.5 kb HindIII and 2.7 kb HindIII fragments, respectively, correspond to three of the bands seen in the genomic Southern hybridization analysis. A third genomic library, CA5314-GEM12, was screened as before to obtain the fourth clone, 3HB4. The 3.2 kb HindIII-BamHI fragment of 3HB4 is a subfragment of the 8.0 kb band seen in the HindIII lane. The four isolated lambda recombinants represent the entire gene.

(33, 34). No regions were found in the CEF-3 upstream sequence that showed any substantial similarity to these *S. cerevisiae* regulatory elements. Inspection of the 3' downstream region (Fig. 4) revealed a potential termination sequence proposed by Zaraf, TAG...TAGGT...TTT (35), that is also found in other *Candida* genes (32).

Deduced Amino Acid Sequence Comparison of CEF-3 and YEF-3

Comparison of the deduced amino acid sequences of the EF-3 genes of *S. cerevisiae* and *C. albicans* using the GCG GAP program reveals 77.6% identity and 86.5% similarity. The alignment of the two sequences is shown in Figure 5 along with features of a putative functional domain of EF-3. These features, which include nucleotide-binding motifs of a putative ATPase active site and poly-lysine blocks proposed to be involved in ribosome interaction (7), are conserved in the CEF-3 amino acid sequence. The only exception to the above being a single change of an alanine for a cysteine in region A of the first nucleotide-binding motif. This difference is also seen in Region A of the first and second repeated motifs in the *S. cerevisiae* amino acid sequence. The nucleotide binding motifs and poly-lysine blocks are located in a carboxy-terminal portion of the two proteins in a region spanning amino acid residues 463 to 1049 in the *C. albicans* sequence and 461 to 1044 in the *S. cerevisiae* sequence. The conservation of these functional features results in a higher percentage of identity between the carboxy-terminal portions of the two EF-3 proteins (86.5%) than that between the amino-terminal portions (66.5%).

CEF-3 Gene Structure

The assignment of the CEF-3 gene to chromosome 5 was determined by electrophoretic separation of chromosomes from *C. albicans* followed by transfer to a nylon filter and hybridization with various *C. albicans* chromosome probes (data not shown). The 981 bp EcoRI fragment from the CEF-3 gene and the 5 kb XbaI fragment from pYK5 hybridize to the same chromosomal band on the CHEF gel blot. According to the chromosome assignment of Magee and coworkers (20; B. B. Magee, personal communication) this co-hybridization identifies chromosome 5 as the location of the *C. albicans* EF-3 gene.

Nucleotide sequence, restriction map and genomic Southern hybridization data suggest that the CEF-3 gene is present at a single locus in the diploid genome of *C. albicans*. Figure 6 shows a genomic Southern blot hybridized with the 981 bp EcoRI fragment from the CEF-3 gene. From this hybridization analysis and the restriction map derived from nucleotide sequence information there is no indication of more than one locus for CEF-3 in the genome. Furthermore, for the enzymes used to construct the restriction map of a full-length clone (data not shown), there is no evidence of restriction site polymorphism between alleles at the CEF-3 locus.

DISCUSSION

Biochemical data reported by Colthurst and co-workers (11) indicate that *Candida albicans* has a soluble factor functionally analogous to the *Saccharomyces cerevisiae* EF-3. A cell-free poly(U)-dependent elongation system from *C. albicans* is ATP-dependent due to the presence of an EF3-like activity. Western blot analysis using antibody raised against *S. cerevisiae* EF-3 confirms presence of EF-3 in *C. albicans*. The work in this paper

supports the above data with the identification and isolation of the gene encoding EF-3 of *C. albicans*.

That EF-3 is essential and unique to fungi suggests that it may be a target for anti-fungal drugs. Analysis of the EF-3 amino acid sequence reveals that putative functional features are conserved between *C. albicans* and *S. cerevisiae*. A comprehensive comparison of amino acid sequences of a variety of fungal EF-3 proteins would help to determine which residues or arrangement of residues are important for activity. An understanding of the structure and function of the active site(s) as well as the site(s) of interaction with other components of the elongation cycle could lead to an informed design of specific anti-fungal drugs.

The nucleotide sequence comparison that shows the non-coding regions of the two genes to be less than 50% identical prompts further study of the expression of the *C. albicans* EF-3 gene. This divergence in the non-coding sequences, specifically the upstream region, leaves open the question of how gene expression is controlled in *C. albicans*. The isolation of many *Candida* genes via expression in *Saccharomyces* would suggest that elements involved in gene expression are conserved between the two fungi (36). On the other hand, although *Candida* genes are generally functional in *Saccharomyces*, the reverse is not true. Though sequences similar to particular *Saccharomyces* regulatory elements are found (32; this work), their roles in gene expression in *Candida albicans* cannot be ascertained until functional studies are done. Therefore, an understanding of the regulatory features in *Candida albicans* will have to go beyond a search for elements that are found in *Saccharomyces cerevisiae*.

ACKNOWLEDGEMENTS

We thank Calvin S. McLaughlin for the *Saccharomyces cerevisiae* EF-3 containing plasmid pYEF-3, Beatrice B. Magee for the chromosome 5 marker plasmid pYK5 and Mary Frances Ypma-Wong for the *Candida albicans* library CA5314-GEM12. We wish to acknowledge Michael Y. Irwin for the excellent technical assistance during various portions of this work. We also wish to thank M. Y. Irwin, S. M. Saporito and M. F. Ypma-Wong for critical reading of the manuscript and many valuable suggestions. This work was supported by a grant (GM23999) from the National Institute of General Medical Science.

REFERENCES

- Skogerson, L. and Wakatama, E. (1976) *Proc. Nat. Sci., U.S.A.* **73**, 73-76.
- Dasmahapatra, B. and Chakraburty, K. (1981) *J. Biol. Chem.* **256**, 9,999-10,004.
- Uritani, M. and Miyazaki, M. (1988) *J. Biochem.* **103**, 522-530.
- Skogerson, L. (1979) *Methods Enzymol.* **60**, 676-685.
- Kamath, A. and Chakraburty, K. (1986) *J. Biol. Chem.* **261**, 12,593-12,595.
- Kamath, A. and Chakraburty, K. (1986) *J. Biol. Chem.* **261**, 12,596-12,598.
- Qin, S., Xie, A., Bonato, M. C. and McLaughlin, C. S. (1990) *J. Biol. Chem.* **265**, 1903-1912.
- Miyazaki, M., Uritani, M. and Kagiymama, H. (1986) *Nucleic Acids Res. Symposium Series* # 17, pp.171-174.
- Kamath, A. and Chakraburty, K. (1989) *J. Biol. Chem.* **264**, 15,423-15,428.
- Chakraburty, K. and Kamath, A. (1988) *Int. J. Biochem.* **20**, 581-590.
- Colthurst, D. R., Santos, M., Grant, C. M., and Tuite, M. F. (1991) *FEMS Microbio. Letters* **80**, 45-50.
- Odds, F. C. (1988) *Candida and Candidosis: A Review and Bibliography* (Baillere Tindall, London, England).
- Sundstrom, P., Smith, D. C. and Sypherd, P. S. (1990) *J. Bact.* **172**, 2036-2045.
- Boeke, J. D., Garfinkel, D. J., Styles, C. A., and Fink, G. R. (1985) *Cell* **40**, 491-500.

15. Southern, E. (1975) *J. Mol. Biol.* **98**, 503–517.
16. Qin, S., Moldave, K. and McLaughlin, C.S. (1987) *J. Biol. Chem.* **262**, 7802–7807.
17. Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
18. Feinberg, A.P. and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
19. Saporito, S.M. and Sypherd, P.S. (1991) *Gene* **106**, 43–49.
20. Magee, B.B., Koltin, Y., Gorman, J.A., and Magee, P.T. (1988) *Mol. Cell. Biol.* **8**, 4721–4726.
21. Davis, R.W., Botstein, D. and Roth, J.R. (1980) *Advanced Bacterial Genetics: A manual for genetic engineering* (Cold Spring Harbor University Press, Cold Spring Harbor, NY).
22. Kurtz, M.B., Kirsch, D.R. and Kelly, R. (1988) *Microbio. Sci.* **5**, 58–63.
23. Whelan, W.L., Kirsch, D.R., Kwon-Chung, K.J., Wahl, S.M. and Smith, P.D. (1990) *J. Infect. Diseases* **162**, 513–518.
24. Carlock, L.R. (1986) *BRL Focus* **8**, 6–8.
25. Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.* **7**, 1513–1523.
26. Holmes, D.S. and Quigley, M. (1981) *Anal. Biochem.* **114**, 193–197.
27. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A laboratory manual* (Cold Spring Harbor University Press, Cold Spring Harbor, NY).
28. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, G., Smith, J.A. and Struhl, K. (1989) *Current Protocols in Molecular Biology* (Greene Publishing and Wiley-Interscience, New York, NY).
29. Sanger, F. and Coulson, A.R. (1975) *J. Mol. Biol.* **94**, 441–448.
30. Deveruex, J., Haerberli, P. and Smithies, O. (1984) *Nucleic Acids Res.* **12**, 387–395.
31. Struhl, K. (1987) *Cell* **49**, 295–297.
32. Kirsch, D.R., Kelly, R. and Kurtz, M.B. (1990) *The Genetics of Candida* (CRC Press, Inc., Boca Raton, FL).
33. Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M.C., Planta, R.J., Schultz, L., Friesen, J.D., Fried, H. and Rosvash, M. (1984) *Nucleic Acids Res.* **12**, 8295–8312.
34. Leer, R.J., van Raamsdonk-Duin, M.M.C., Mager, W.H. and Planta, R.J. (1985) *Current Genetics* **9**, 273–277.
35. Zarat, K.S. and Sherman, F. (1982) *Cell* **28**, 563–573.
36. Kurtz, M.B., Kirsch, D.R. and Kelly, R. (1988) *Microbio. Sci.* **5**, 58–63.