Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 2211–2215, March 1994
Chemistry

Molecular cloning, characterization, and overexpression of *ERG7*, the *Saccharomyces cerevisiae* gene encoding lanosterol synthase

(sterol biosynthesis/2,3-epoxysqualene/gene cloning)

E. J. COREY*, SEIICHI P. T. MATSUDA*, AND BONNIE BARTEL†

*Department of Chemistry, Harvard University, Cambridge, MA 02138; and †Whitehead Institute for Biomedical Research, Cambridge, MA 02142

Contributed by E. J. Corey, November 18, 1993

ABSTRACT **We report the cloning, characterization, and overexpression of *Saccharomyces cerevisiae ERG7*, which encodes lanosterol synthase [(*S*)-2,3-epoxysqualene mutase (cyclizing, lanosterol forming), EC 5.4.99.7], the enzyme responsible for the complex cyclization/rearrangement step in sterol biosynthesis. Oligonucleotide primers were designed corresponding to protein sequences conserved between *Candida albicans* ERG7 and the related *Arabidopsis thaliana* cycloartenol synthase [(*S*)-2,3-epoxysqualene mutase (cyclizing, cycloartenol forming), EC 5.4.99.8]. A PCR product was amplified from yeast genomic DNA using these primers and was used to probe yeast libraries by hybridization. Partial-length clones homologous to the two known epoxysqualene mutases were isolated, but a full-length sequence was found neither in cDNA nor genomic libraries, whether in phage or plasmids. Two overlapping clones were assembled to make a functional reconstruction of the gene, which contains a 2196-bp open reading frame capable of encoding an 83-kDa protein. The reconstruction complemented the *erg7* mutation when driven from either its native promoter or the strong *ADH1* promoter.**

Lanosterol synthase [(*S*)-2,3-epoxysqualene mutase (cyclizing, lanosterol forming), EC 5.4.99.7] (1–3) catalyzes the cyclization of (*S*)-2,3-epoxysqualene (compound **1**) to lanosterol (compound **2**) via cationic intermediates (Fig. 1) in one of the most complex and powerful one-step constructions known in either biochemistry or synthetic chemistry. Lanosterol synthase alters 20 bonds, forms 4 rings, and sets 7 stereocenters in catalyzing the cationic cyclization of epoxysqualene to lanosterol with a degree of precision that cannot yet be approached by known synthetic methodology. The understanding of this complex enzyme-mediated reaction at the atomic level requires knowledge of the active site and the precise hydrophobic interactions that control the conformation of the bound substrate during cyclization.

Lanosterol synthase also has great biological and medicinal significance. Vertebrates synthesize lanosterol as the initial sterol precursor to cholesterol, the steroid hormones, and vitamin D. Yeast and other fungi metabolize lanosterol to ergosterol, the primary fungal membrane sterol and a growth signal (4–7). Sterol biosynthetic enzymes such as lanosterol synthase are a preferred target of antifungal drugs (8).

We wished to clone and express the *Saccharomyces cerevisiae* gene encoding lanosterol synthase as part of a long-term project investigating the mechanism of this enzyme. Genes can be overexpressed or deleted easily in yeast, and the resultant strains can serve as bases for genetic studies to define the function and regulation of the sterol biosynthetic pathway. Much of the current knowledge about the detailed catalytic mechanism of lanosterol synthase was obtained from yeast using epoxysqualene analogs (9–11). The overexpression of the yeast lanosterol synthase will greatly facilitate the isolation and structural identification of products derived from such analogs, as well as the detailed mechanistic and structural analysis of the enzyme.

Most of the genes in the *S. cerevisiae* ergosterol biosynthetic pathway have now been cloned and characterized (12–24), primarily by complementation of the extensive family of mutants generated by Karst and Lacroute (25). *ERG7* (25, 26), which encodes lanosterol synthase, is a notable exception. A possible reason that the *S. cerevisiae ERG7* gene has been elusive was discovered during the course of this work. This gene is unstable in the RecA$^+$ *Escherichia coli* strain JM101 and could only be maintained in the *recA* strain DH5α. This instability was manifested by severe underrepresentation in libraries.‡

EXPERIMENTAL PROCEDURES

Strains. The *erg7* yeast strain SMY1 (*gal2 hem3-6 erg7 ura3-52*) (27) was used for complementation experiments. Lanosterol synthase activity was measured in the *ERG7* yeast strain JBY575 (*MATa ura3-52 trp1-Δ63 leu2-3,112 his3-Δ200 ade2* Gal$^+$; from J. Brill, Stanford University). Protein was purified from an *ard1-Δ pep4-Δ leu2* strain (a strain that lacks vacuoles and N-terminal acetylase, from E. Park, Alkermes, Inc., Cambridge, MA) carrying pSM61.6. *E. coli* strain DH5α (28) was used to make all constructs, and strain LE392 (29) was the host for λ derivatives.

Media, Growth, and Transformation Conditions. Yeast and *E. coli* were transformed, selected, and propagated according to published procedures (27, 30, 31). Strain SMY1 was grown in medium supplemented with ergosterol (20 μg/ml), heme (13 μg/ml), and Tween 80 (0.5%).

DNA Methods. Standard DNA analysis and cloning techniques were performed as described (31). DNA fragments were purified from agarose gels using Qiaex (Qiagen, Chatsworth, CA). Nested sets of unidirectional deletions were prepared (32) and sequenced by the dideoxynucleotide chain-termination method using the Applied Biosystems automated system. Where the resultant sequence did not overlap, the gaps were filled with custom-oligonucleotide-primed reactions. The sequence reported in Fig. 2 was determined at least once on each strand. Hybridization probes were labeled by using [α-^{32}P]dCTP and random hexamer primers (International Biotechnologies). Colony and plaque hybridizations were done as described (31), hybridizing for 12 hr at 65°C (33) and washing in 30 mM NaCl/3 mM sodium citrate/0.1% SDS, pH 7.0 at 25°C.

PCR. Two regions were identified with eight contiguous amino acids conserved between cycloartenol synthase [CAS1, (*S*)-2,3-epoxysqualene mutase (cyclizing, cy-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: CAS1, cycloartenol synthase.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. U04841).

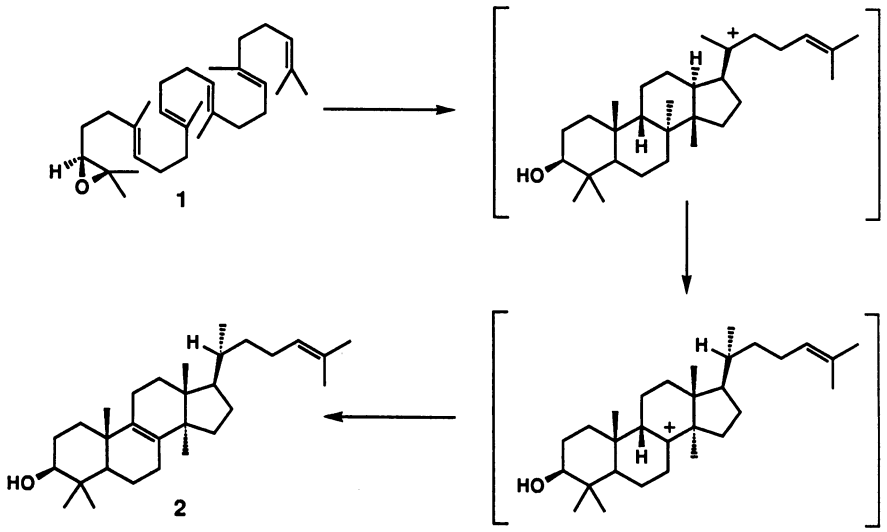

FIG. 1. Lanosterol synthase-catalyzed cyclization of (*S*)-2,3-epoxysqualene (compound **1**) to lanosterol (compound **2**) via cationic intermediates.

cloartenol forming), EC 5.4.99.8] (27) from *Arabidopsis thaliana* and ERG7 (34, 35) from *Candida albicans*: KGAW-PFST (amino acids 467–474 in CAS1) and (D/E)YPYVECT (amino acids 556–563 in CAS1). Degenerate oligonucleotides were synthesized corresponding to these regions, flanked at the 5′ ends with the sequence CGAATTC to allow subsequent digestion with the restriction enzyme *Eco*RI. These oligonucleotides were used to amplify yeast genomic DNA (36) using the Perkin–Elmer/Cetus PCR system according to the manufacturer's instructions using 30 cycles with 1 min at 94°C, 1 min at 55°C, and 3 min at 72°C.

Libraries and Plasmids. The following yeast libraries were used: a 2μ (high-copy) cDNA library (37), a 2μ genomic library (38), a *CEN* (single-copy) genomic library (39), and a

```
                                                                  gtgaaagaatc    11
acatatatagcgccaaattactgggattgctgttgtatggtgcacagttggtttgttttaacggtatgcgttttgtataccctctattatg   101
tagtgcaagaaaatctgctgctattcgtgattactgttacaacctaacggtttaaatgaaacctggttctgaagggtcattttataactt   191
caagttcccttagccttttcgattcatttttgattatgccatttctagaccgtgttatagggcgctggcgtttaatttggtgtagcttggttt   281
agtcaagagttgtattagtgttcctcgataaagtcgatgtttccggatattgtgttaaaatttcaagtatgctactaatggggtaaagtt   371
gcatgattagcagagacatatggcttgttatggttcggcttcctcatttttcatgcttagtttttgtccatctcattgtacatttctgaa   461
tcctaatgcatgactccctaacattactattaaatttctcaatagtgaagaataagcaaaatgggaaccatgataatttctagctttctct   551
ccacccctattttaatttgcaatcatatatagtactttcaatagcatcttttctagatttgatatctgcggaaacttttcgtttataatc   641
gtttaggtgaaaagttttatatcggttatttacagatatacatttttctcaaaaaaaaaaaaatataatacatgccctcagcttttaata   731
aagccattgagcacaagcctctccagtaatgtactgctgtgcccaataaccttaccaataatcgtcgcccacaaagaaagtacaaaacag   821

ATGACAGAATTTTATTCTGACACAATCGGTCTACCAAAGACAGATCCACGTCTTTGGAGACTGAGAACTGATGAGCTAGGCCGAGAAAGC   911
 M  T  E  F  Y  S  D  T  I  G  L  P  K  T  D  P  R  L  W  R  L  R  T  D  E  L  G  R  E  S

TGGGAATATTTAACCCCTCAGCAAGCCGCAAACGACCCACCATCCACTTTCACGCAGTGGCTTCTTCAAGATCCCAAATTTCCTCAACCT  1001
 W  E  Y  L  T  P  Q  Q  A  A  N  D  P  P  S  T  F  T  Q  W  L  L  Q  D  P  K  F  P  Q  P

AATCCAGAAAGAAATAAGCATTCACCAGATTTTTCAGCCTTCGATGCGTGTCATAATGGTGCATCTTTTTTCAAACTGCTTCAAGAGCCT  1091
 N  P  E  R  N  K  H  S  P  D  F  S  A  F  D  A  C  H  N  G  A  S  F  F  K  L  L  Q  E  P

GACTCAGGTATTTTTCCGTGTCAATATAAAGGACCCATGTTCATGACAATCGGTTACGTAGCCGTAAACTATATCGCCGGTATTGAAATT  1181
 D  S  G  I  F  P  C  Q  Y  K  G  P  M  F  M  T  I  G  Y  V  A  V  N  Y  I  A  G  I  E  I

CCTGAGCATGAGAGAATAGAATTAATTAGATACATCGTCAATACAGCACATCCGGTTGATGGTGGCTGGGGTCTACATTCTGTTGACAAA  1271
 P  E  H  E  R  I  E  L  I  R  Y  I  V  N  T  A  H  P  V  D  G  G  W  G  L  H  S  V  D  K

TCCACCGTGTTTGGTACAGTATTGAACTATGTAATCTTACGTTTATTGGGTCTACCCAAGGACCACCCGGTTTGCGCCAAGGCAAGAAGC  1361
 S  T  V  F  G  T  V  L  N  Y  V  I  L  R  L  L  G  L  P  K  D  H  P  V  C  A  K  A  R  S

ACATTGTTAAGGTTAGGCGGTGCTATTGGATCCCCTCACTGGGGAAAAATTTGGCTAAGTGCACTAAACTTGTATAAATGGGAAGGTGTG  1451
 T  L  L  R  L  G  G  A  I  G  S  P  H  W  G  K  I  W  L  S  A  L  N  L  Y  K  W  E  G  V

AACCCTGCCCCTCCTGAAACTTGGTTACTTCCATATTCACTGCCCATGCATCCGGGGAGATGGTGGGTTCATACTAGAGGTGTTTACATT  1541
 N  P  A  P  P  E  T  W  L  L  P  Y  S  L  P  M  H  P  G  R  W  W  V  H  T  R  G  V  Y  I

CCGGTCAGTTACCTGTCATTGGTCAAATTTTCTTGCCCAATGACTCCTCTTCTTGAAGAACTGAGGAATGAAATTTACACTAAACCGTTT  1631
 P  V  S  Y  L  S  L  V  K  F  S  C  P  M  T  P  L  L  E  E  L  R  N  E  I  Y  T  K  P  F

GACAAGATTAACTTCTCCAAGAACAGGAATACCGTATGTGGAGTAGACCTATATTACCCCCATTCTACTACTTTGAATATTGCGAACAGC  1721
 D  K  I  N  F  S  K  N  R  N  T  V  C  G  V  D  L  Y  Y  P  H  S  T  T  L  N  I  A  N  S

CTTGTAGTATTTTACGAAAAATACCTAAGAAACCGGTTCATTTACTCTCTATCCAAGAAGAAGGTTTATGATCTAATCAAAACGGAGTTA  1811
 L  V  V  F  Y  E  K  Y  L  R  N  R  F  I  Y  S  L  S  K  K  K  V  Y  D  L  I  K  T  E  L

CAGAATACTGATTCCTTGTGTATAGCACCTGTTAACCAGGCGTTTTGCGCACTTGTCACTCTTATTGAAGAAGGGGTAGACTCGGAAGCG  1901
 Q  N  T  D  S  L  C  I  A  P  V  N  Q  A  F  C  A  L  V  T  L  I  E  E  G  V  D  S  E  A

TTCCAGCGTCTCCAATATAGGTTCAAGGATGCATTGTTCCATGGTCCACAGGGTATGACCATTATGGGAACAAATGGTGTGCAAACCTGG  1991
 F  Q  R  L  Q  Y  R  F  K  D  A  L  F  H  G  P  Q  G  M  T  I  M  G  T  N  G  V  Q  T  W

GATTGTGCGTTTGCCATTCAATACTTTTTCGTCGCAGGCCTCGCAGAAAGACCTGAATTCTATAACACAATTGTCTCTGCCTATAAATTC  2081
 D  C  A  F  A  I  Q  Y  F  F  V  A  G  L  A  E  R  P  E  F  Y  N  T  I  V  S  A  Y  K  F

TTGTGTCATGCTCAATTTGACACCGAGTGCGTTCCAGGTAGTTATAGGGATAAGAGAAAGGGGGCTTGGGGCTTCTCAACAAAAACACAG  2171
 L  C  H  A  Q  F  D  T  E  C  V  P  G  S  Y  R  D  K  R  K  G  A  W  G  F  S  T  K  T  Q

GGCTATACAGTGGCAGATTGCACTGCAGAAGCAATTAAAGCCATCATCATGGTGAAAAACTCTCCCGTCTTTAGTGAAGTACACCATATG  2261
 G  Y  T  V  A  D  C  T  A  E  A  I  K  A  I  I  M  V  K  N  S  P  V  F  S  E  V  H  H  M

ATTAGCAGTGAACGTTTATTTGAAGGCATTGATGTGTTATTGAACCTACAAAACATCGGATCTTTTGAATATGGTTCCTTTGCAACCTAT  2351
 I  S  S  E  R  L  F  E  G  I  D  V  L  L  N  L  Q  N  I  G  S  F  E  Y  G  S  F  A  T  Y

GAAAAAATCAAGGCCCCACTAGCAATGGAAACCTTGAATCCTGCTGAAGTTTTTGGTAACATAATGGTAGAATACCCATACGTGGAATGT  2441
 E  K  I  K  A  P  L  A  M  E  T  L  N  P  A  E  V  F  G  N  I  M  V  E  Y  P  Y  V  E  C

ACTGATTCATCCGTTCTGGGGTTGACATATTTTCACAAGTACTTCGACTATAGGAAAGAGGAAATACGTACACGCATCAGAATCGCCATC  2531
 T  D  S  S  V  L  G  L  T  Y  F  H  K  Y  F  D  Y  R  K  E  E  I  R  T  R  I  R  I  A  I

GAATTCATAAAAAAATCTCAATTACCAGATGGAAGTTGGTATGGAAGCTGGGGTATTTGTTTTACATATGCCGGTATGTTTGCATTGGAG  2621
 E  F  I  K  K  S  Q  L  P  D  G  S  W  Y  G  S  W  G  I  C  F  T  Y  A  G  M  F  A  L  E

GCATTACACACCGTGGGGGAGACCTATGAGAATTCCTCAACGGTAAGAAAAGGTTGCGACTTCTTGGTCAGTAAACAGATGAAGGATGGC  2711
 A  L  H  T  V  G  E  T  Y  E  N  S  S  T  V  R  K  G  C  D  F  L  V  S  K  Q  M  K  D  G

GGTTGGGGGGAATCAATGAAGTCCAGTGAATTACATAGTTATGTGGATAGTGAAAAATCGCTAGTCGTTCAAACCGCATGGGCGCTAATT  2801
 G  W  G  E  S  M  K  S  S  E  L  H  S  Y  V  D  S  E  K  S  L  V  V  Q  T  A  W  A  L  I

GCACTTCTTTTCGCTGAATATCCTAATAAAGAAGTCATCGACCGCGGTATTGACCTTTTAAAAAATAGACAAGAAGAATCCGGGGAATGG  2891
 A  L  L  F  A  E  Y  P  N  K  E  V  I  D  R  G  I  D  L  L  K  N  R  Q  E  E  S  G  E  W

AAATTTGAAAGTGTAGAAGGTGTTTTCAACCACTCTTGTGCAATTGAATACCCAAGTTATCGATTCTTATTCCCTATTAAGGCATTAGGT  2981
 K  F  E  S  V  E  G  V  F  N  H  S  C  A  I  E  Y  P  S  Y  R  F  L  F  P  I  K  A  L  G

ATGTACAGCAGGGCATATGAAACACATACGCTTTAAgttttatttatataagtgttgttagaggctgcaacaattagaaactagtgcaaaac  3071
 M  Y  S  R  A  Y  E  T  H  T  L  ter

tagtatatttatttgtatatgtagtagtaaaataataggtgaaaaaaaaaaaaaaaaaaa                                  3131
```

FIG. 2. Peptide and DNA sequence of *ERG7*. Regions common to pSM61.1R and pSM61.01 are underlined.

λEMBL3a genomic library (40). Bacterial plasmids were constructed from pBluescript II (Stratagene). pRS416 was the 2μ (high copy in yeast) shuttle vector (41). The high-copy expression vector pAD4Δ (42) was used to overexpress ERG7 in yeast under the control of the *ADH1* promoter.

Cloning. The PCR product was used to screen a yeast cDNA library (37) by colony hybridization in *E. coli*. Eight out of 10^6 colonies were positive. The inserts were excised and subcloned into pBluescript II. The longest of these, named pSM61.01, had a 1.5-kb insert, which when sequenced showed homology to *C. albicans* ERG7 and *A. thaliana* CAS1 (27). The 300-bp fragment at the 5′ end was excised with *Eco*RI and *Sal* I and used to probe a λ yeast genomic library (40). One out of 50,000 plaques was positive. DNA was prepared from the phage, and the hybridizing sequence was localized to a 5-kb fragment at one end of the insert, bounded by an *Sph* I site in the insert and a *Sal* I site within a λ arm. This fragment was subcloned into pBluescript II and named pSM61.1R, and partial sequencing revealed homology to the known terpene synthases and 230 bp of overlap with the 5′ end of pSM61.01 (Fig. 2).

Plasmid Construction. The *ERG7* gene was reconstructed from a 5′ truncated cDNA and a 3′ truncated genomic fragment using the restriction enzyme *Esp*3I, which has a single recognition site in the gene, located in the overlapping region common to the two fragments. A 3.8-kb fragment of pSM61.1R containing the *ERG7* promoter and the first 1.1 kb of coding sequence was excised with *Esp*3I and *Sal* I. pSM61.01 was digested with the same two enzymes, yielding a 4.2-kb fragment that contained the last 1.2 kb of the coding sequence fused to the pBluescript II vector. The fragments containing the two halves of the gene were ligated together to make pSM61.2, which is the reconstructed *ERG7* gene on its native promoter in the pBluescript II vector. The insert of pSM61.2 was excised with *Kpn* I and *Sac* I and ligated to pRS416 cut with the same two enzymes to make pSM61.3, the reconstructed *ERG7* gene on its native promoter in a high-copy yeast vector.

Restriction sites were not located conveniently to excise the coding sequence, so it was reassembled by a three-piece construction route that removed the native promoter. pSM61.1R was digested with *Rsa* I, and the 474-bp fragment (−10 to 464 bp of the coding sequence) was ligated into pBluescript II cut with *Sma* I to make pSM61.4. The 2.0-kb *Eco*NI–*Spe* I fragment (the 2.0 kb at the 3′ end of the gene) was excised from pSM61.2 and ligated to the 3.3-kb fragment of pSM61.4 cut with the same two enzymes (which contained the 300 bp at the 5′ end of the gene, with 10 bp of 5′ nontranslated DNA in the pBluescript II vector). This construct was named pSM61.5 and consists of the reconstructed *ERG7* gene with 10 bp of 5′ nontranslated sequence in pBluescript II. The *ERG7* coding sequence was excised from pSM61.5 with *Sac* I and *Sal* I and was ligated to pAD4Δ cut with the same two enzymes to make pSM61.6. pSM61.6 contains the *ERG7* coding sequence driven from the highly expressed constitutive *ADH1* promoter (43).

Complementation of SMY1. The reconstructed *ERG7* gene was able to complement the *erg7* mutation. The *erg7*-deficient mutant SMY1 was transformed with control and *ERG7* plasmids, and the resultant strains were tested on synthetic complete plates lacking uracil (44) supplemented with ergosterol, heme, and Tween 80 (26), or heme and Tween 80 without ergosterol. Whereas SMY1 transformed with pRS416 (vector) remained an ergosterol auxotroph, SMY1 transformed with pSM61.3 (*ERG7*, 2μ) or pSM61.6 (*ADH1–ERG7*, 2μ) was able to grow without ergosterol.

Enzyme Assay. Conversion in each strain was quantitated by incubating racemic 2,3-epoxysqualene with a homogenate from that strain, recovering unreacted epoxysqualene and determining the extent that the (*S*) isomer was depleted [the (*S*) isomer is a substrate of lanosterol synthase, whereas the (*R*) isomer is not (45)] by chiral HPLC. Recombinant strains were grown in synthetic complete medium lacking either uracil or leucine, as appropriate. Yeast cells (300 mg) from each strain were suspended in 30 ml of 100 mM sodium phosphate buffer, pH 7.0, and lysed in a French press at 20,000 psi (1 psi = 6.9 kPa). Particulate cell debris was removed from the homogenates by centrifugation for 20 min at 5000 × *g*. The crude enzyme was assayed by addition of 50 μl of (±)-2,3-epoxysqualene solution (20 mg/ml) in 20% Triton X-100 (46, 47) to 1 ml of homogenate. After 1 hr, 2 ml of tetrahydrofuran was added to stop each reaction, and precipitated salts and denatured protein were removed by filtration through Celite. Tetrahydrofuran was removed by rotary evaporation, and the aqueous residue was extracted with 1 ml of ether, which was then passed through a plug of anhydrous magnesium sulfate to remove water and then through a plug of silica gel to remove Triton X-100. Epoxysqualene was purified by HPLC (DuPont Zorbax column, 4.6 mm × 25 cm, 0.1% isopropanol in hexane, 2 ml/min, 9-min retention time, UV detection at 210 nm). The recovered epoxysqualene was analyzed by chiral HPLC [Daicel Chiralpak AS column, 4.6 mm × 25 cm, 0.05% isopropanol in hexane, 1 ml/min, retention times (*R*) 6.9 min, (*S*) 8.7 min, UV detection at 210 nm]. The amount of (*S*)-epoxysqualene that was consumed was quantitated by integration, using the (*R*)-epoxysqualene as an internal standard. Protein concentrations were quantitated as described (48).

RESULTS AND DISCUSSION

It would seem straightforward to clone *ERG7* from *S. cerevisiae* by complementation of an *erg7*-deficient mutant (25, 26). This approach was used successfully to clone the *C. albicans ERG7* gene (34). However, we surveyed over a million SMY1 transformants (27) from several libraries (37–39) without finding an ergosterol prototroph. Having recently cloned the cycloartenol synthase gene (*CAS1*) from *A. thaliana* (27), we instead cloned the *S. cerevisiae ERG7* gene by homology to *CAS1* and *C. albicans ERG7*, the two known genes encoding epoxysqualene mutases.

After identifying homologous regions between the *A. thaliana* CAS1 and the *C. albicans* ERG7 proteins we designed complementary degenerate PCR primers, amplified genomic *S. cerevisiae* DNA, and probed libraries (37–40). Although each library appeared to lack the complete *ERG7* gene, partial clones were found that were reassembled to reconstruct a functional version of the gene (see *Experimental Procedures* and Fig. 3).

Yeast strains that overexpress *ERG7* have increased lanosterol synthase activity. JBY575, a strain with a wild-type *ERG7* allele, was transformed in parallel with pRS416 (vector), pSM61.3 (*ERG7*, 2μ), and pSM61.6 (*ADH1–ERG7*, 2μ). Extracts from the three resultant strains were assayed for lanosterol synthase activity (see *Experimental Procedures*). JBY575 carrying pRS416 synthesized 2.6 μg of lanosterol per mg of protein per hr, JBY575 carrying pSM61.3 synthesized 4.4 μg of lanosterol per mg of protein per hr, and JBY575 carrying pSM61.6 synthesized 21.7 μg of lanosterol per mg of protein per hr, demonstrating that overexpression of ERG7 from the *ADH1* promoter results in formation of lanosterol synthase at a level ≈8 times that in the wild type. This fact and the 62% identity between *S. cerevisiae ERG7* and the proposed *ERG7* sequence from *C. albicans* (35) provide direct evidence that the latter is, in fact, the *ERG7* gene.

Lanosterol synthase purified from an *ard1*-Δ *pep4*-Δ yeast strain carrying pSM61.6 according to a modification (to be published separately) of our published procedure (49) had a molecular mass of 80 kDa as estimated by SDS/PAGE. This molecular mass conflicts with that of a sample we reported earlier (49), which appeared to contain a single 26-kDa

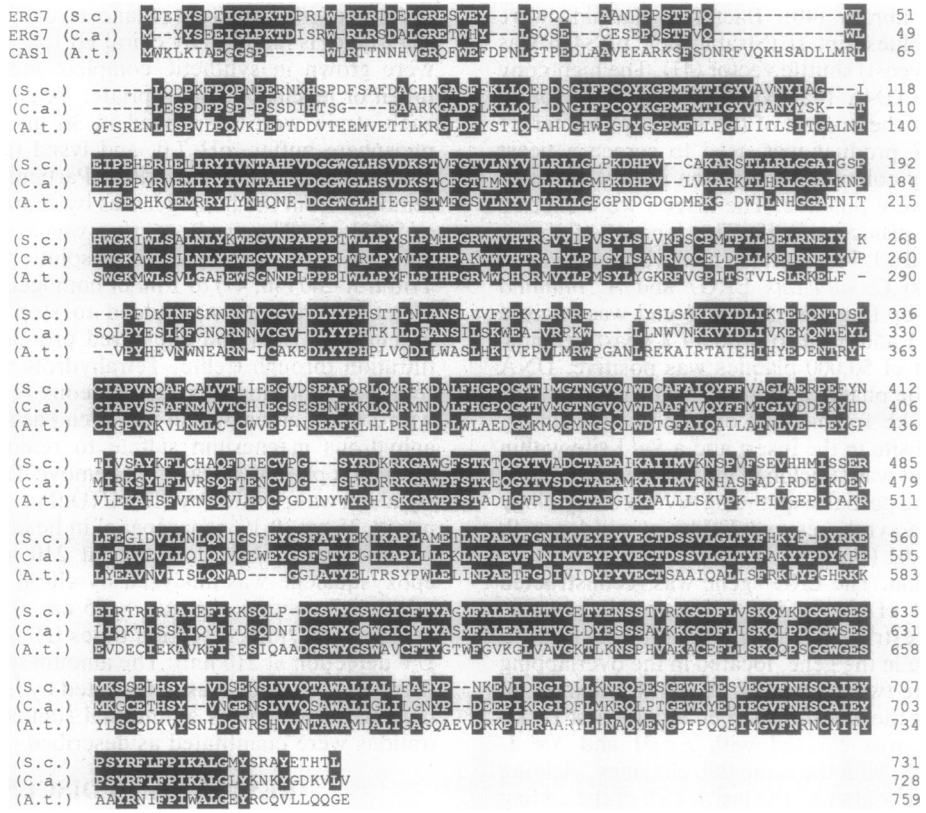

FIG. 3. Alignment of epoxysqualene mutase amino acid sequences. The sequences shown are as follows: *S. cerevisiae* ERG7 [ERG7 (S. c.)], *C. albicans* ERG7 [ERG7 (C. a.)] (34, 35), and *A. thaliana* CAS1 [CAS1 (A. t.) (27)]. Amino acid positions identical in at least two of the three sequences are boxed. Hyphens indicate gaps introduced to maximize alignment.

protein by SDS/PAGE. Purification of lanosterol synthase from the overexpressing strain proceeded similarly to the wild-type yeast, except that an 80-kDa protein that was not observed in the wild-type yeast became visible in the purest fractions from the overexpressor. Further efforts at purification succeeded in separating the 26-kDa band from the enzymatic activity. Edman degradation of the 80-kDa protein yielded the N-terminal sequence Thr-Glu-Phe-Tyr-Ser-Asp-Thr-Ile, identical to the predicted sequence lacking the initial methionine (Fig. 2), indicating that the correct molecular mass of yeast lanosterol synthase is 80 kDa.

The 2196-bp *ERG7* open reading frame (Fig. 2) predicts an 83-kDa protein 62% identical to *C. albicans* ERG7 (34, 35) and 38% identical to *A. thaliana* CAS1 (27) (Fig. 3). The homology between these enzymes is least toward the N terminus. Several regions emerge as candidates for areas of contact with the substrate 2,3-oxidosqualene, given the correspondences shown in Fig. 3 and the requirement for precise

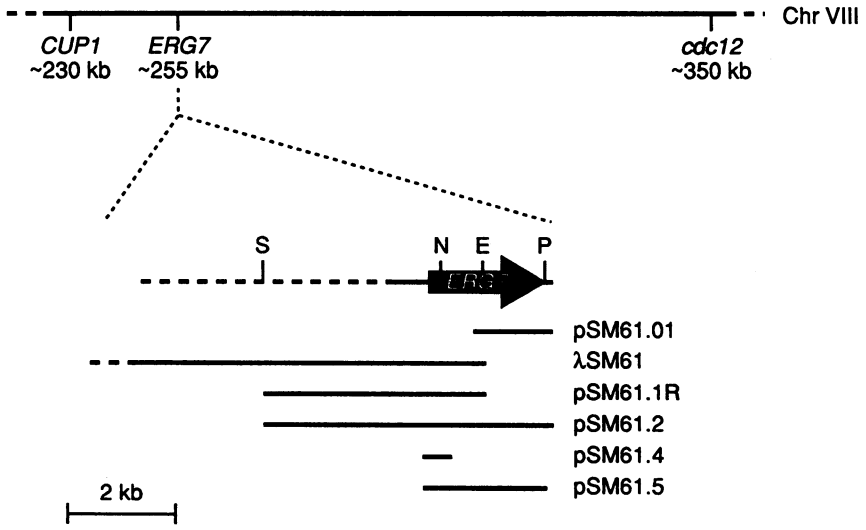

FIG. 4. Schematic diagram showing the map position of *ERG7* on yeast chromosome (Chr) VIII as determined by hybridization to an overlapping set of λ clones (50) and the *ERG7* genomic locus, with the sequenced region shown as a solid line and the *ERG7* open reading frame shown as a solid arrow (S, *Sph* I; N, *Eco*NI; E, *Esp*3I; P, *Spe* I). The inserts of relevant plasmids and λ clones are indicated as solid bars below the map.

fit between the folded substrate and hydrophobic residues on the enzyme (11).

Physical maps of several yeast chromosomes have recently been generated from sets of overlapping λ clones (50). The 2.4-kb *Xba* I–*Spe* I fragment of *ERG7* hybridized to λ clones 2335 and 5147, overlapping clones located between *CUP1* and *CDC12*, locating *ERG7* to a position between 255 and 260 kb on chromosome VIII (Fig. 4).

Further research is needed to increase the efficiency of production of recombinant lanosterol synthase to levels appropriate for detailed structural and mechanistic studies. The determination of additional sequences encoding lanosterol synthase from a variety of organisms will narrow further the set of conserved residues that may be crucial to catalysis. These residues can then be scrutinized individually with experiments in which selected amino acids are altered using site-specific mutagenesis.

We are indebted to the DNA sequencing facility of Pfizer, Inc. (Groton, CT) and, in particular, Drs. Michael Fogliano and Arthur E. Franke for the sequence shown in Fig. 2, to Dr. David P. Bartel (Whitehead Institute for Biomedical Research) for the oligonucleotides used in the PCR, to Dr. Haoping Liu (Whitehead Institute for Biomedical Research) for the *S. cerevisiae* cDNA library, to Dr. E. Park (Alkermes, Inc., Cambridge, MA) for the *ard1*-Δ *pep4*-Δ *leu2* strain, to Mr. William S. Lane (Harvard University) for performing the Edman degradation, and to Prof. William N. Lipscomb (Harvard University) for comments on the manuscript. This research was assisted financially by the National Institutes of Health.

1. Corey, E. J., Russey, W. E. & Ortiz de Montellano, P. R. (1966) *J. Am. Chem. Soc.* **88,** 4750.
2. Dean, P. D. G., Ortiz de Montellano, P. R., Bloch, K. & Corey, E. J. (1967) *J. Biol. Chem.* **242,** 3014–3015.
3. Barton, D. H. R., Gosden, A. F., Mellows, G. & Widdowson, D. A. (1968) *Chem. Commun.*, 1067.
4. Rodriguez, R. J., Taylor, F. R. & Parks, L. W. (1982) *Biochem. Biophys. Res. Commun.* **106,** 435–441.
5. Ramgopal, M. & Bloch, K. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 712–715.
6. Parks, L. W., Rodriguez, R. J. & Low, C. (1986) *Lipids* **21,** 89–91.
7. Lorenz, R. T., Casey, W. M. & Parks, L. W. (1989) *J. Bacteriol.* **171,** 6169–6173.
8. Mercer, E. I. (1990) *Biochem. Soc. Trans.* **18,** 45–65.
9. Corey, E. J. & Virgil, S. C. (1991) *J. Am. Chem. Soc.* **113,** 4025–4026.
10. Corey, E. J., Virgil, S. C. & Sarshar, S. (1991) *J. Am. Chem. Soc.* **113,** 8171–8172.
11. Corey, E. J., Virgil, S. C., Liu, D. R. & Sarshar, S. (1992) *J. Am. Chem. Soc.* **114,** 1524–1525.
12. Anderson, M. S., Yarger, J. G., Burck, C. L. & Poulter, C. D. (1989) *J. Biol. Chem.* **264,** 19176–19184.
13. Anderson, M. S., Muehlbacher, M., Street, I. P., Proffitt, J. & Poulter, C. D. (1989) *J. Biol. Chem.* **264,** 19169–19175.
14. Arthington, B. A., Hoskins, J., Skatrud, P. L. & Bard, M. (1991) *Gene* **107,** 173–174.
15. Arthington, B. A., Bennett, L. G., Skatrud, P. L., Guynn, C. J., Barbuch, R. J., Ulbright, C. E. & Bard, M. (1991) *Gene* **102,** 39–44.
16. Ashman, W. H., Barbuch, R. J., Ulbright, C. E., Jarrett, H. W. & Bard, M. (1991) *Lipids* **26,** 628–632.
17. Fegueur, M., Richard, L., Charles, A. D. & Karst, F. (1991) *Curr. Genet.* **20,** 365–372.
18. Gaber, R. F., Copple, D. M., Kennedy, B. K., Vidal, M. & Bard, M. (1989) *Mol. Cell. Biol.* **9,** 3447–3456.
19. Jandrositz, A., Turnowsky, F. & Högenauer, G. (1991) *Gene* **107,** 155–160.
20. Jennings, S. M., Tsay, Y. H., Fisch, T. M. & Robinson, G. W. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 6038–6042.
21. Lorenz, R. T. & Parks, L. W. (1992) *DNA Cell Biol.* **11,** 685–692.
22. Oulmouden, A. & Karst, F. (1990) *Gene* **88,** 253–257.
23. Oulmouden, A. & Karst, F. (1991) *Curr. Genet.* **19,** 9–14.
24. Tsay, Y. H. & Robinson, G. W. (1991) *Mol. Cell. Biol.* **11,** 620–631.
25. Karst, F. & Lacroute, F. (1977) *Mol. Gen. Genet.* **154,** 269–277.
26. Gollub, E. G., Liu, K., Dayan, J., Adlersberg, M. & Sprinson, D. B. (1977) *J. Biol. Chem.* **252,** 2846–2854.
27. Corey, E. J., Matsuda, S. P. T. & Bartel, B. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11628–11632.
28. Hanahan, D. (1983) *J. Mol. Biol.* **166,** 557–580.
29. Borck, J., Beggs, J. D., Brammer, W. J., Hopkins, A. S. & Murray, N. E. (1976) *Mol. Gen. Genet.* **146,** 199–207.
30. Schiestl, R. H. & Gietz, R. D. (1989) *Curr. Genet.* **16,** 339–346.
31. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1993) *Current Protocols in Molecular Biology* (Wiley–Interscience, New York).
32. Özkaynak, E. & Putney, S. D. (1987) *Biotechniques* **5,** 770–773.
33. Church, G. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
34. Kelly, R., Miller, S. M., Lai, M. H. & Kirsch, D. R. (1990) *Gene* **87,** 177–183.
35. Buntel, C. J. & Griffin, J. H. (1992) *J. Am. Chem. Soc.* **114,** 9711–9713.
36. Hoffman, C. S. & Winston, F. (1987) *Gene* **57,** 267–272.
37. Liu, H., Krizek, J. & Bretscher, A. (1992) *Genetics* **132,** 665–673.
38. Carlson, M. & Botstein, D. (1982) *Cell* **28,** 145–154.
39. Rose, M. D., Novick, P., Thomas, J. H., Botstein, D. & Fink, G. R. (1987) *Gene* **60,** 237–243.
40. Sweetser, D., Nonet, M. & Young, R. A. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1192–1196.
41. Sikorski, R. S. & Hieter, P. (1989) *Genetics* **122,** 19–27.
42. Ballester, R., Michaeli, T., Ferguson, K., Xu, H.-P., McCormick, F. & Wigler, M. (1989) *Cell* **59,** 681–686.
43. Ammerer, G. (1983) *Methods Enzymol.* **101,** 192–201.
44. Guthrie, C. & Fink, G. R., eds. (1991) *Guide to Yeast Genetics and Molecular Biology* (Academic, New York).
45. Shishibori, T., Fukui, T. & Suga, T. (1973) *Chem. Lett.*, 1137–1138.
46. Corey, E. J., Lin, K. & Jautelat, M. (1968) *J. Am. Chem. Soc.* **90,** 2724–2726.
47. Nadeau, R. G. & Hanzlik, R. P. (1968) *Methods Enzymol.* **15,** 346–351.
48. Bradford, M. (1976) *Anal. Biochem.* **72,** 248–254.
49. Corey, E. J. & Matsuda, S. P. T. (1991) *J. Am. Chem. Soc.* **113,** 8172–8174.
50. Riles, L., Dutchik, J. E., Baktha, A., McCauley, B. K., Thayer, E. C., Leckie, M. P., Braden, V. V., Depke, J. E. & Olson, M. V. (1993) *Genetics* **134,** 81–150.