

Molecular Cloning of a *Schizosaccharomyces pombe* cDNA Encoding Lanosterol Synthase and Investigation of Conserved Tryptophan Residues

E. J. Corey,¹ Seiichi P. T. Matsuda, C. Hunter Baker, Alice Y. Ting, and Hengmiao Cheng

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

Received January 1, 1996

A *Schizosaccharomyces pombe* cDNA encoding lanosterol synthase was cloned by complementing a *Saccharomyces cerevisiae* lanosterol synthase mutant. The predicted 83-kDa protein is 54–58% identical to other lanosterol synthases. The previously known lanosterol synthases contain 229 conserved residues, which should encompass the catalytically essential amino acids. This number is decreased dramatically by including the *Sc. pombe* lanosterol synthase in the analysis; 42 residues are no longer conserved and therefore are catalytically nonessential. We have begun mutagenic studies to identify catalytic residues from the remaining conserved residues. Mutant *Sa. cerevisiae* lanosterol synthase genes were generated in which phenylalanine was specifically substituted for conserved tryptophan residues. All of the resultant mutant enzymes retained the ability to complement the *Sc. cerevisiae* lanosterol synthase mutant, suggesting that these conserved tryptophan residues are not catalytically essential. © 1996 Academic Press, Inc.

Lanosterol synthase [(*S*)-2,3-epoxysqualene mutase (cyclizing, lanosterol forming), EC 5.4.99.7], catalyzes the cyclization of oxidosqualene to lanosterol. Although the bond changes that occur during this reaction are now established (1), the molecular details of how the enzyme promotes them have not been uncovered. Oligonucleotide-directed mutagenesis could in principle be used to investigate conserved amino acid residues and identify those that are catalytically relevant, but we found that available sequence information was insufficient to guide these experiments. Although four genes encoding lanosterol synthases have been described, they represent two pairs of closely related organisms. The rat (2) and human (3,4) predicted proteins are 83% identical, and those of the budding yeasts *Saccharomyces cerevisiae* (5) and *Candida albicans* (6) are 61% identical. The number of identical residues (229 of 732) found in alignments of these four predicted proteins precludes exhaustive mutagenesis studies. Since sequence information from a distantly related lanosterol synthase should eliminate many nonessential conserved residues, we cloned this gene from the fission yeast *Schizosaccharomyces pombe*. Fission yeast genes are nearly equidistant from mammalian and budding yeast homologs, as the ancestors of fission yeast diverged from those of budding yeast 880 million years ago, only 320 million years after the progenitor of mammals diverged from this line (7).

MATERIALS AND METHODS

Strains. Plasmids were constructed and propagated using *E. coli* strain DH5 α . The *Sa. cerevisiae* lanosterol synthase mutant SMY1 (5 *MATa gal2 hem3-6 erg7 ura3-52*), was the host for complementation experiments. *Sa. cerevisiae* strain SMY4 (*MATa hem1::TRP1 ura3-52 trpl-863 leu2-3,112 his3-D200 ade2 Gal⁺*), was constructed by integrating *TRP1* into the *HEM1* locus of JBY575 (*MATa ura3-52 trpl-863 leu2-3, 112 his3-D200 ade2 Gal⁺*, from J. Brill, Stanford University) using plasmid pSM69, described below. *Sa. cerevisiae* strain SMY8 (*MATa erg7::HIS3 hem1::TRP1 ura3-52-trpl-863 leu2-3,112 his3-D200 ade2 Gal⁺*), was constructed by integrating *HIS3* into the *ERG7* locus of SMY4 using plasmid pSM70, described below.

Libraries and plasmids. Bacterial plasmids were constructed from pBluescript II (Stratagene). The *Sc. pombe* gene was cloned from a *Sc. pombe* cDNA library (8) in which cDNAs are controlled by *Sa. cerevisiae ADHI* promoter and terminator

¹ To whom correspondence should be addressed. Fax: (617) 495-0376.

sequences (9) and which contains the *Sa. cerevisiae* 2 μ origin of replication and *URA3* selectable marker. pSM69 contains sequence designed to replace *HEM1* in a yeast genome with *TRP1*. This plasmid was constructed by cutting the *HEM1*-containing plasmid LB6 (10) with *Pst*I and *Sal*I to excise most of the coding sequence and leave the flanking sequence. The *TRP1* gene was excised from pJH-W1 (from J. Hill, Harvard Medical School) and ligated to the truncated LB6 sequence described above. pSM70 was constructed similarly by obtaining the yeast *HIS3* gene by *Cla* I-*Eco*R V digestion of pJH-W1 (from J. Hill, Harvard Medical School), then removing the coding sequence from the lanosterol synthase-containing pSM61.2 (5) with the same two enzymes, and ligating together the fragments. The integrating vector pRS305GAL was constructed by excising the promoter region from pRS316GAL with *Xho* I, and ligating that fragment to the vector pRS305 (11) cut with the same enzyme. pSM61.21 is a *LEU2*-based integrating plasmid encoding *Sc. cerevisiae* lanosterol synthase under *GAL* control, constructed by inserting the *Not*I-*Sal* I fragment of pSM61.6 (5) into pRS305GAL. pRS305GAL, pSM61.21, and pSM61.21 mutants linearized with *Bst*E II integrate into the *Sa. cerevisiae* genome at the *LEU2* locus.

Media. *E. coli* were propagated in Luria broth (LB, 11), which was supplemented with 100 μ g/mL ampicillin when the bacteria contained plasmids. SMY1 library transformants were selected on synthetic complete media plates (1.5% agar) supplemented with ergosterol (20 μ g/mL), heme (13 μ g/mL), and Tween 80 (5 μ g/mL) without uracil, and ergosterol autotrophs were identified by replica-plating onto the same medium from which ergosterol and Tween 80 had been omitted. *Sa. cerevisiae* *hem1* mutant strain SMY4 was propagated at 30°C in YPD (13) medium supplemented with heme (13 μ g/mL). SMY8 was propagated in the same medium supplemented with ergosterol (20 μ g/mL), heme (13 μ g/mL), and Tween 80 (5 μ g/mL). SMY4 transformants were selected on synthetic complete media (13) plates (1.5% agar) without tryptophan and supplemented with heme (13 μ g/mL), and SMY8 transformants were selected on synthetic complete media plates (1.5% agar) without histidine and supplemented with ergosterol (20 μ g/mL), heme (13 μ g/mL), and Tween 80 (5 μ g/mL). SMY8 strains transformed with pSM61.21 derivatives were analyzed for ergosterol autotrophy by plating on YPG (13) plates supplemented with heme (13 μ g/mL).

DNA methods. Transformation competent *E. coli* were prepared by treatment with divalent cations (12). *Sa. cerevisiae* was transformed using lithium acetate (13). Plasmid DNA was isolated from yeast using glass beads and phenol (13). Restriction enzymes were purchased from New England Biolabs and used according to the manufacturer's directions. DNA fragments for subcloning were purified from agarose gels using Qiaex (Qiagen) according to the manufacturer's instructions, and fragments were ligated using GibcoBRL T4 ligase according to the manufacturer's instructions. DNA was sequenced by the dideoxynucleotide chain termination method using Sequenase (United States Biochemical) according to the manufacturer's instructions. Nested sets of unidirectional deletions were prepared using the New England Biolabs Exo-Size Deletion Kit according to the manufacturer's instructions. Gaps in the resultant sequence were filled with custom-oligonucleotide primed reactions. The reported sequences were determined at least once on each strand. Mutations were introduced (12) using synthetic oligonucleotides that introduced in addition to the desired mutation, a mutation that did not alter the amino acid sequence, but altered the restriction pattern. Uracil-containing single-stranded DNA was prepared from phagemid pSM61.21 in *E. coli* strain RZ1032 coinfecting with M13K07 helper phage. Mutagenic oligonucleotides were used to prime second strand synthesis with T4 polymerase, and the DNA circle was closed with T4 ligase. *E. coli* strain DH5 α was transformed with the putative mutant constructs, and clones with the desired mutation were identified by their restriction pattern.

Cloning the *Sc. pombe* lanosterol synthase gene. Lanosterol synthase was cloned from *Sc. pombe* by genetic complementation of the *Sa. cerevisiae* lanosterol mutant SMY1. This strain was transformed with a yeast expression library containing *Sc. pombe* cDNAs (8) under the control of the constitutive *Sa. cerevisiae* *ADHI* promoter (9), and the resultant transformants were replica-plated onto plates lacking ergosterol. Eight out of $\sim 10^6$ transformants exhibited ergosterol autotrophy. Plasmid DNA was isolated from these strains and used to transform *E. coli*. Plasmids from the resultant *E. coli* lines were indistinguishable from one another by restriction enzyme mapping, and one plasmid was chosen for further analysis and named pSM62. The coding sequence was excised from pSM62 using *Not* I, and ligated to *Not* I-digested Bluescript II (KS+) for sequencing manipulations.

RESULTS AND DISCUSSION

The complementation experiments provided the *Sc. pombe* lanosterol synthase coding sequence (Figure 1), which contains a 2284-bp open reading frame capable of encoding an 83-kDa protein 54–55% identical to the budding yeast lanosterol synthases, and 57–58% identical to the mammalian homologs. As expected, adding the *Sc. pombe* gene to the alignment analysis decreased the number of conserved residues; 229 residues are conserved in the four previously known lanosterol synthases, and 187 remain conserved when the *Sc. pombe* sequence is included in the analysis (Figure 2). These residues are reasonable candidates for mutagenic analysis.

The activity screens we planned were not feasible with currently available yeast strains or expression plasmids, so new ones were constructed. The host strain requires a lanosterol synthase deletion to abolish background activity, and must be able to metabolize galactose, enabling the use

	ccttttttgattat	14
tttgttgcatattacaattaacatagagtaaagacgagtcctaatcaccttttctggaactgcttgatagataggat	tttttttgcgaaaaa	104
ATGGAGGCTTCAGAGTCCGCTCCTGAAITATCTAAAACGGTGAACAACATCGATAAAAAGCTTGTGGCGCTTGAATATCGAATTCAGCTGGT		194
M E A C R V R P E L S K T V E Q I D K S L W R L N I D S A G		
GGAGACATGGGAGTATGTTACAAAAGAAGAGGCTGAAAAGCGTCCACTCACCAATTGCGGAGAAGTATTTCCCTAGGTTTGTGATTTAGAT		284
G E T W E Y V V L R I L E A E K R P L T I A E I S Q F L G F P D K L D		
CTACCTAAAACGACCTCCAGCGAAAACCTCCAATTAGAGAGTCTGAGTATGGTTATGAAATTTTTTCGTCGCTTGCAGCTACCAGACGGTCAC		374
L P K R R P P A K T P L E S A E Y G Y E P F R R L Q L P D G H		
TGGGCTTCTCCTATGAAGGGCCGATGTTCTCATTTGCGGTGCCGTTTTGCTTTTTATATTTTCGCAAAACACCAATTTCCCTAAAGGTTGG		464
W A S M N Y V V L R I L E I C G A V F A F Y I S Q Y F L P F K G W		
GCTCCAGAAAATATCCAATACTTAATTAATCACACAAATGACGATGGTGGATGGGGTATTCACACAGAAGGAGTTTCTACTGTATTCGGA		554
A P E I I Q Y L I N H T N D D G G W G I H T E G V S T V F G		
ACCTCCATGAATATGTAGTTCTTCGAAATTTGGGTATGGATGCGGGACATCCAGTTGCTACACGAGCTCGTAACAGATTTACATGAACTT		644
T S M N Y V V L R I L G M D L Y F P A R H P F I A D T R L R N R I L H E L		
GGAGGAGCTATGGATGCCCGCATTTGGGAAAAATTTGGCTCGCTACACTTAACTGTACGACTGGGATGGTGTAAACCTATCCCCCA		734
G G A I G C P H W G K F W L A T L N G C Y D W D G V N P I P P		
GAACTATGGTTATACCTGACTGGATTTCACTCTCTGGAAAAATGGTGGTGCACGTTTCGTTTAGTATACITTCGCTATGGGATATATG		824
E L W L L P V N A A M N T V C V F H E G P S S K A L S D I P R L E Y M		
TATGGAGAGAGGCTGAAATGCCCTAAGGATTCGTTAATAATGCAATTTGCGAAAGGAGTTGTACGTAGAAAACTATGACAGCATTAATTTTT		914
Y G E R L K C P K D S L I M Q L R K E L Y V E N Y D S I N F		
GCGGATCATAGAAAATACGATTAGTGATGTGGACCTCTATTTTCCATACATACTCAAATATATAGATCGTTTGAATTTGGATATTTGAAAAATAT		1004
A D H R N Y V V L R I L G M D L Y F P A R H P F I A D T R L R N R I L H E L		
TTTTACTTATCTAAGACCCCTCTTGGCTTTAAAAAAGCTGGTACTCGACGGGCATACGAACTTATAAAAAATAGAAGATCGAATACCGAATTAC		1094
F T Y L R P S W L K K L G T R R R A Y E L I K I E D Q N T D Y		
TCTTTGATTTGGCCAGTAAATGACGAAATGAACACTGTTTTCGTTTTACTTTTCAATGAAGGACCTTCTAGTAAAGCTTTTCCAAAAGCACATTT		1184
S C I G P V N A A M N T V C V F H E G P S S K A L S D I P R L E Y M		
CAACGTTTGCATGATTTTCATGTGGGTCCAACCGAAGGAATGCTTATGCGTGGTACTAATGGTTTTCGAAAGTTTGGGAAACTTCTCTCAC		1274
Q R L H D F M W V Q P E G M L M R G T N G L Q V W E T S F T		
TTGCAAGCGCTCGTAGAATCAGGTTCTTTACGAAAAAGAGGCTTCAAACAGAGATATCGCCAAAGCTCTGGAGTTTTTGGATCGACAACAG		1364
L Q A L V E S G L Y E K E A F T D I A L E F L L V D I P R L R L		
ATTGCAACGCAATATGAAGGGAGTGGATATAGGTACAAATTTTGGGGCATGGCCATTTAGTAAATATAACTCAAGGTTATACGGTGTCA		1454
I R T Q Y E G S G Y R Y N S L G A W P P S N I T Q G Y T V S		
GATACTACATCTGAAGCTCTGCGTGTGATTTGCTTGTCCAATCATTTGCCAGATTTTGA AAAACTGGTTGATATACCTAGGCTAAGGTTA		1544
D T T S E A L R A V L L V L P D F I L D R L N R I L H E L		
TCGGTTGATGTGATATTTGGGAATGCAAAAATGAGAACCTTGGGTTTTGCTTCATATAGGCTCGCGAACAGGTGAATGGATGGAGTTGTTA		1634
S V D V I L G M Q N E N L G F A S Y E P A R T G E W M E L L		
AATCTGCGGAAAGTTTTTGGAAAATATTTGGTTCGAAATCTTTATCCCGAGTGCACCACATCTGTGATATTTGGCACTCCGGGCTTTTACA		1724
N P A E V F S N I M V E Y S Y P E C T T S V I L A L R A F T		
AAGTATGATCCTGGATATCGTGTGTGATAAAATGAAAAACTACTTATGAAAAATGCTTTTGGAAATACGTTGTAAAGATGCAACGACCAGACGGA		1814
K Y D P G Y R R D E I E N T I E N A L E Y V V K M Q R P D G		
AGCTGGTATGGAAGTTGGGCTATTTGTTTTACGTAATGACGACCTGTTTGGCAGAGTTCTTTGGCATCAGCCGGTGGATATTCGAAAAAC		1904
S W Y G T E T E S S L V T Q T G W A A M F A T G S L A S D I P R L E Y N		
TGCCACGCTCAGAAGAAGCGGTGCCAATTTTTACTGTGCAAGCAAAGACCAGATGGTGGATGGAGTGGAGTTACATGGCTTGTGTTACT		1994
C P V Q K K A C E F L L S K Q R P D G G W S E S Y M A C V T		
GGAGTTTTACACAGAAAACAGAATCAAGTTTGGTAACACAACTGGTTGGGCTTTGGACGCACTGATTAATGCAAAGTACCTCGATAGAAAA		2084
G V Y T E T E S S L V T Q T G W A A M F A T G S L A S D I P R L E Y N		
CCGATAGAAAAAGGAATTAATTTTTGATGGCGTCTCAAAGAGCGATGGGAGCTGGCAGCAAAAATCGATGGAAGGGATCTTTAACAAG		2174
P I E K G I K F L M A S Q K S D G S W Q Q K S M E G I F N K		
AATGTTGCCATTCATATCCCAACTACAAGTTGTACTTTTTCCATATACACATTTGGGAAAGTTTGCAGAAAGCAATACGGCAATACCTGACC		2264
N V A I A Y F N Y K L Y F S I Y T L G K F A K Q Y G N Y L T		
ATTTAGTtttagtttagtagtag		2284
I ter		

FIG. 1. Nucleotide and deduced amino acid sequences of the *Sc. pombe* lanosterol synthase cDNA (GenBank accession number U41368).

of galactose-induced expression. Lanosterol synthase mutants are only viable in a heme mutant background (14), so the expression strain must be mutant in heme biosynthesis. The expression plasmid must contain the *GAL1* promoter and should integrate into the genome to bypass the requirement for amino acid or nucleotide deficient media. Yeast strain SMY8 (*MATa erg::HIS3hemi::TRP1 ura3-52 trp1-863 leu2-3,112 his3-D200 ade2 Gal⁺*) and plasmid pRS305GAL (*amp^R, LEU2, GAL1* promoter) fulfill these requirements and were therefore constructed, as described above.

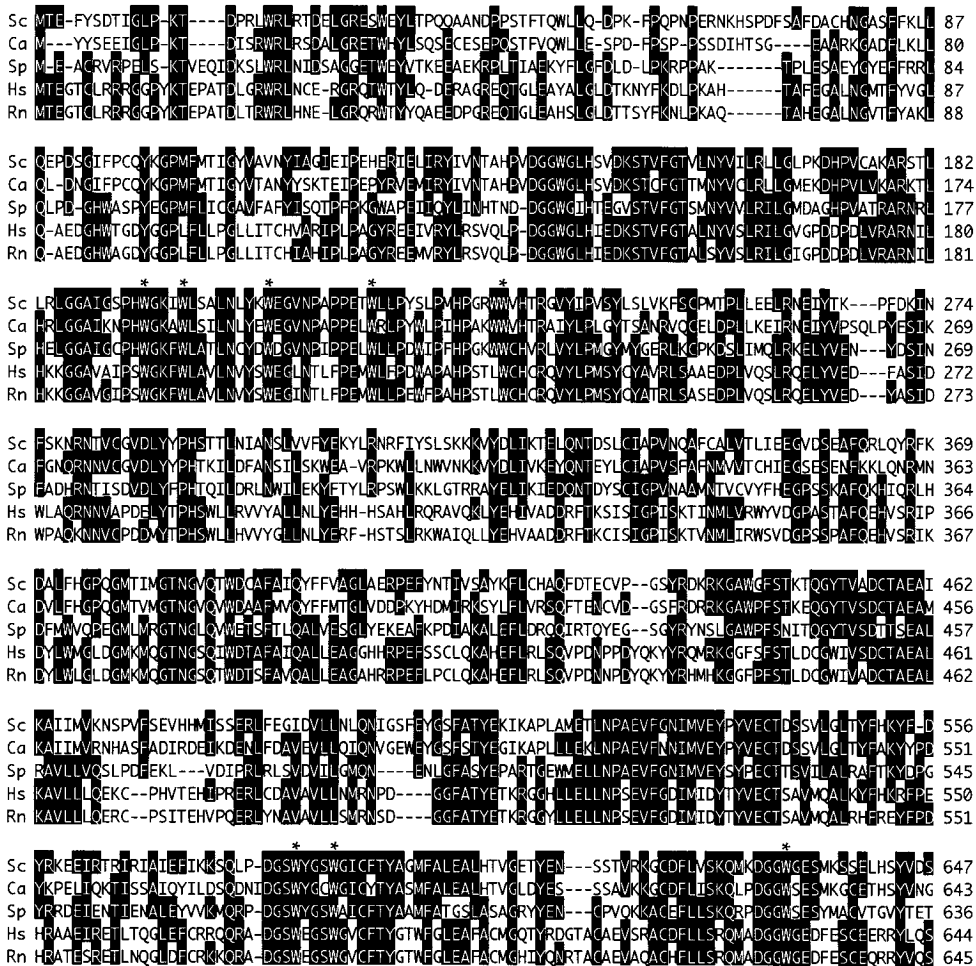


FIG. 2. Alignment of lanosterol synthase amino acid sequences from *Saccharomyces cerevisiae* (Sc), *Candida albicans* (Ca), *Schizosaccharomyces pombe* (Sp), *Homo sapiens* (Hs), and *Rattus norvegicus* (Rn). Residues conserved in at least 3 sequences are shaded, and tryptophan residues mutated to phenylalanine in this study are marked with an asterisk.

A potential complication when assaying enzymes encoded by mutant genes is that if a mutation were introduced in an essential amino acid residue, it would not be detected if that mutation should revert, i.e., spontaneously acquire a back-mutation restoring enzymatic activity. Although each copy of the gene is unlikely to revert, if lanosterol synthase mutants expressing inactive lanosterol synthase genes are allowed to divide repeatedly reversion becomes more problematic. Revertants would be able to biosynthesize sterols and consequently divide faster than the other cells, and if allowed sufficient time would comprise the majority of the population. This pitfall was avoided by maintaining the mutant lanosterol synthase genes under the control of the inducible promoter *GALI*, which is on when the yeast are grown on galactose, and off when the carbon source is glucose. Cells were maintained on glucose to prevent transcription of the mutant lanosterol synthases, so that if a reversion had occurred the progeny of that cell would have no growth advantage

and remain extremely rare. As a further precaution, at least five clones of each mutant construct were analyzed for the ability to complement a lanosterol synthase yeast mutant.

Fourteen tryptophan residues are conserved in the five known lanosterol synthases (*Sa. cerevisiae* numbering): W19, W31, W143, W194, W198, W207, W218, W232, W390, W583, W587, W632, W657, and W726. It has been speculated that the electron-rich sidechains of tryptophan residues control the carbocations in the cyclization reaction (6). If so, mutant enzymes in which a conserved tryptophan residue is replaced with phenylalanine should not be active. However, if the conserved tryptophan residues function to maintain the structural integrity of the enzyme through hydrophobic interactions or π -stacking, the conservative substitution of phenylalanine for tryptophan might be inconsequential. We constructed a series of eight lanosterol synthase mutants, in each of which either W194, W198, W207, W218, W232, W583, W587, or W632 was mutated to phenylalanine. The W194F, W198F, W207F, W218F, W232F, W583F, W587F, W632F mutants all retained catalytic activity, as evidenced by their ability to complement the lanosterol synthase mutation in SMY8. The sidechains of these eight tryptophan residues appear not to be crucial to lanosterol biosynthesis from (*S*)-2,3-epoxysqualene.

REFERENCES

1. Corey, E. J., Virgil, S. C., Cheng, H., Baker, C. H., Matsuda, S. P. T., Singh, V., and Sarshar, S. (1995) *J. Am. Chem. Soc.* in press.
2. Kusano, M., Shibuya, M., Sankawa, U., and Ebizuka, U. (1995) *Biol. Pharm. Bull.* **18**, 195–197.
3. Baker, C. H., Matsuda, S. P. T., Liu, D. R., and Corey, E. J. (1995) *Biochem. Biophys. Res. Com.* **213**, 154–160.
4. Sung, C.-K., Shibuya, M., Sankawa, U., and Ebizuka, Y. (1995) *Biol. Pharm. Bull.* **18**, 1459–1461.
5. Corey, E. J., Matsuda, S. P. T., and Bartel, B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 2211–2215.
6. Buntel, C. J., and Griffin, J. H. (1992) *J. Am. Chem.Soc.* **114**, 9711–9713.
7. Wu, C. I., Li, W. H., Limbach, K. J., Scapulla, R. C., and Shen, J. J. (1986) *J. Mol. Evol.* **23**, 61–75.
8. Fikes, J. D., Becker, D. M., Winston, F., and Guarente, L. (1990) *Nature* **346**, 291–294.
9. Ammerer, G. (1983) *Methods Enzymol.* **101**, 192–201.
10. Urban-Grimal, D., Volland, C., Garnier, T., Dehoux, P., and Labbe-Bois, R. (1986) *Eur. J. Biochem.* **156**, 511–519.
11. Sikorski, R. S., and Hieter, P. (1989) *Genetics* **122**, 19–27.
12. Current Protocols in Molecular Biology (1995) (Ausubet F. M., Brent R., Kingston R. E., Moore D. D., Seidman J. G., Smith J. A., and Struhl K., Eds.), Wiley-Interscience, New York.
13. Guide to Yeast Genetics and Molecular Biology (1991) (Guthrie C., and Fink G. R., Eds.), Academic Press, New York.
14. Lorenz, R. T., and Parks, L. W. (1991) *Lipids* **26**, 598–603.