Primary Structure of the Multifunctional α Subunit Protein of Yeast Fatty Acid Synthase Derived from FAS2 Gene Sequence*

(Received for publication, March 28, 1988)
A. Habib Mohamed, Subrahmanyam S. Chirala, Nilima H. Mody, Wei-Yong Huang, and Salih J. Wakil

From the Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

The yeast fatty acid synthase consists of two multifunctional proteins, α and β, arranged in an αβ6 complex with a molecular weight of 2.4 × 10^6. Five of the seven enzymatic activities reside in the β subunit, while the remaining two activities, β-ketoacyl synthase and β-ketoacyl reductase, and the domain of the acyl carrier protein, with its prosthetic group, 4'-phosphopantetheine, are in the α subunit. The genes FAS1 and FAS2 coding for α and β subunits, respectively, have been cloned and the sequence of FAS1 has been reported (Chirala, S. S., Kuziora, M. A., Spector, D. M., and Wakil, S. J. (1987) J. Biol. Chem. 262, 4231–4240). In this study, we present the nucleotide sequence of the FAS2 gene. The sequence has an open reading frame, coding for a protein of 1894 amino acids with a calculated molecular weight of 207,863. The location of the serine site of attachment of the prosthetic group of the acyl carrier protein domain and the active cysteine–SH site of β-ketoacyl synthase have been identified at residues 180 and 1512, respectively, in the deduced amino acid sequence. A putative NADPH binding site of β-ketoacyl reductase has been suggested at residue 1038 based on the similarities to the corresponding amino acid sequences -Gly-Ser-Ala- of the pyridine nucleotide enzymes.

We could not find any sequence homology in the 5' flanking sequence of the FAS1 and FAS2 genes that would suggest common regulatory function. However, in the sequence of these two genes there is an identical eight-base pair sequence TCATTATG at the translational initiation site suggesting that the subunit stoichiometry probably results from equal translational efficiency of the mRNAs of both FAS1 and FAS2 genes. The S1 endonuclease mapping suggests that there is a transcriptional initiation site at about 40 nucleotides upstream of the first ATG codon and a transcriptional termination site about 300 nucleotides downstream of the TAG stop codon. The gene does not contain introns as no intron consensus TACTAAC have been found in the sequence.

The fatty acid synthase complex catalyzes the synthesis of long chain saturated fatty acid from acetyl- and malonyl-CoA. In prokaryotes and plants the complex consists of an acyl carrier protein (ACP)1 and seven structurally independent monofunctional enzymes, while in animals all the component enzymatic activities and ACP are organized in one large polypeptide chain (1). In the case of lower eukaryotes, such as yeast and other lower fungi, the enzymatic activities are distributed between two large multifunctional polypeptide chains, α and β subunits (2–4). The enzymatically active synthase is an αβ6 complex having an estimated molecular weight of 2.4 × 10^6. The β subunit contains five of the activities, mapped in the following order on the polypeptide: acetyltransferase, enoyl reductase, dehydratase, and malonyl/palmitoyl transacylases. The α subunit contains the remaining two activities, β-ketoacyl reductase and β-ketoacyl synthase, and the ACP where the growing fatty acid chain is attached to the cysteamine thiol of the prosthetic group 4'-phosphopantetheine (1–4). In yeast, the α and β subunits are encoded by two unlinked genes FAS2 and FAS1, respectively (5). There are indications that these genes are coordinately expressed and that the cell synthesizes equivalent amounts of the mRNAs and the subunits to form the complex αβ6 (6, 7).

Recently, we have cloned the FAS1 and FAS2 genes using two independent methods. Immunological screening of a bank of yeast genomic DNA in the vector ColE1 yielded clones 102B5 and 33F1 (7). The other method, the genetic complementation of fatty acid requiring auxotrophs of Saccharomyces cerevisiae by plasmids selected from a bank of yeast genomic DNA sequences in the vector YEpl3, yielded clones YEpFAS1 and YEpFAS2 (7). The plasmids 33F1 and YEpFAS2 contained the entire FAS1 and FAS2 gene, respectively (7, 8). The FAS1 gene of 33F1 DNA which encodes the β subunit of fatty acid synthase was cloned in YEp33F1, and its nucleotide sequence was determined (8). In the present paper, we report the complete nucleotide sequence of FAS2 gene from the clone YEpFAS2 which encodes the α subunit. The sequence shows an open reading frame that could code for a protein of 1894 amino acids. In the nucleotide-derived amino acid sequence, the serine site of attachment of 4'-phosphopantetheinyl prosthetic group of ACP domain (9) and the active cysteine–SH site of the β-ketoacyl synthase, the enzyme that condenses the acyl and malonyl moieties (10), have been identified. A putative NADPH binding site of the β-ketoacyl reductase has been proposed.

MATERIALS AND METHODS

Yeast and Escherichia coli Strains—Yeast strain X2180-B1, GRF-18, or 6657-4D was used to prepare total RNA, and E. coli strain RR1 ("pro leu thi lacY HsdR' hsdM' endo") was used to grow plasmids from yeast transformants. JM 107 was used as the host for the isolation of single-stranded DNA on infection with M13 clones.

1. The abbreviations used are: ACP, acyl carrier protein; SDS, sodium dodecyl sulfate; bp, base pairs; kb, kilobase pairs; kbp, kilobase pairs; HPLC, high performance liquid chromatography.
YWNP2462

Yeast Fatty Acid Synthase

MATERIALS AND METHODS

Yeast Fatty Acid Synthase—Yeast was grown on YPD or SD media (11), and E. coli was grown on L-broth or M9 media supplemented with ampicillin (40 μg/ml) as required (12).

Sequence Analysis of YEPFAS2—The molecular cloning of FAS2 gene was reported earlier (7). Two genomic clones, YEpFAS2 and 102B5, were isolated by genetic complementation and immunological screening, respectively (7). YEpFAS2 has yeast genomic DNA of 11.4 kb which is capable of complementing all known fas2 mutations (8) and encodes a full-length α subunit (21). The clone 102B5 has DNA insert of 3.4 kbp which hybridizes to the YEpFAS2 yeast DNA (7). In order to determine the extent of sequence information needed from the 11.4 kb insert and to compare the sizes of the mRNAs for FAS1 and FAS2 genes, we performed Northern analyses. Yeast total RNA was separated on formaldehyde-containing agarose gels and transferred to nitrocellulose sheets. Radioactive probes specific for FAS1 and FAS2 genes were prepared using a respective gene fragment cloned in M13 and universal primer. As shown in Fig. 1, FAS1 probe hybridized to 6.5-kb mRNA and FAS2 to 6.0-kb mRNA.

RESULTS

Sequence Analysis of YEpFAS2—The molecular cloning of FAS2 gene was reported earlier (7). Two genomic clones, YEpFAS2 and 102B5, were isolated by genetic complementation and immunological screening, respectively (7). YEpFAS2 has yeast genomic DNA of 11.4 kb which is capable of complementing all known fas2 mutations (8) and encodes a full-length α subunit (21). The clone 102B5 has DNA insert of 3.4 kbp which hybridizes to the YEpFAS2 yeast DNA (7). In order to determine the extent of sequence information needed from the 11.4 kb insert and to compare the sizes of the mRNAs for FAS1 and FAS2 genes, we performed Northern analyses. Yeast total RNA was separated on formaldehyde-containing agarose gels and transferred to nitrocellulose sheets. Radioactive probes specific for FAS1 and FAS2 genes were prepared using a respective gene fragment cloned in M13 and universal primer. As shown in Fig. 1, FAS1 probe hybridized to 6.5-kb mRNA and FAS2 to 6.0-kb mRNA.

HindIII cleavage of the 11.4-kbp DNA insert of YEpFAS2 yielded five fragments with sizes of 4.0, 3.3, 0.9, 1.3, and 2.0 kb. All of these fragments hybridized with the 6.0-kb mRNA except the 4.0 kb (data not shown) suggesting that the latter fragment is not a part of FAS2 gene. The 4.0-kbp DNA segment was cleaved with PvuII endonuclease yielding 2-kbp fragments. The 3’ end of this fragment, along with the other four HindIII segments of the 11.4-kbp genomic insert, were subcloned into either pUC18 or pBR322 and M13 mp18/19, and used for DNA sequence determination. The restriction map and sequence strategy used in the structural analyses of the FAS2 gene are outlined in Fig. 2. Sequencing was accomplished either by chemical modification method of Maxam and Gilbert (17, 22) or dideoxy method of Sanger and Coulson (18, 19), and used for DNA sequence determination. The restriction map and sequence strategy used in the structural analyses of the FAS2 gene are outlined in Fig. 2. Sequencing was accomplished either by chemical modification method of Maxam and Gilbert (17, 22) or dideoxy method of Sanger and Coulson (18, 19). A total of 7.6 kb, starting about 1400 bases upstream of the translational start codon, to about 500 bases downstream of the translational stop codon, was sequenced (Fig. 3). All segments of the sequences were overlapped. The sequences of the subcloned inserts were melded by sequencing the parental plasmid YEpFAS2. More than 95% of the sequences were confirmed by sequencing both strands; wherever only one strand was sequenced, the analyses were done more than once. Fig. 3 also shows the deduced amino acid sequence. The nucleotide numbering starts with the putative translation initiation codon ATG at base 1. Starting with this Met codon,
the sequence presented here, shows an open reading frame that can code for a protein of 1894 amino acids. Table I shows the codon usage, and the predicted amino acid composition of the α subunit of yeast fatty acid synthase. The amino acid composition agreed closely to the published values (23).

The serine site (residue 180) of attachment of the prosthetic group, 3'-phosphopantetheine, of the acyl carrier protein and the cysteine–SH (residue 1312) of β-ketoacyl synthase have been located in the contiguous open reading frame presented in Fig. 3 as indicated, based on previously published amino acid sequences (9, 10).

In order to confirm the reading frame of the nucleotide sequence further, we sequenced some of the peptides derived from tryptic digest of the cyanogen bromide peptides obtained from yeast fatty acid synthase. As described under "Materials and Methods," yeast fatty acid synthase was cleaved with cyanogen bromide, and the resulting peptides were digested with trypsin. The digest was then fractionated on HPLC as shown in Fig. 4. Five fractions were arbitrarily collected and labeled T1, T2, T3, T4, and T5. Each fraction was treated with aminopeptidase M2 and rechromatographed on HPLC using various linear gradients of CH₃CN designated as methods B, C, or D (see "Materials and Methods" for details). The pure peptides obtained from each tryptic fractions were sequenced. Comparisons of the amino acid sequence of 10 peptides revealed that eight of the peptides are derived from α subunit and the remaining two peptides from β subunit. As shown in Table II, the predicted amino acid sequences between residues 87 to 100, 101 to 105, 507 to 511 near the NH₂-terminal region and between residues 1657 to 1676, 1772 to 1796, 1834 to 1842, 1852 to 1863 and 1876 to 1892 near the COOH-terminal regions of the α subunit protein matched exactly with the amino acid sequences of the peptides 1 to 8, respectively. The amino acid sequences of peptides 9 and 10 (Table II) matched with the deduced amino acid sequence of the β subunit protein.

Fig. 2. Restriction map and sequence determination strategy of yeast FAS2 DNA cloned in YEpFAS2. The FAS2 DNA shown in the top line is transcribed from left to right. It was divided into five fragments, p2.0*, p3.3, p0.9, p1.3, and p2.0 and subcloned. Each segment is sequenced to the extent of each of the arrows in the direction shown. Hatched boxes indicate the coding segment. Only the restriction sites used for sequencing is indicated. Open circles denote sequencing by chemical modification method (17, 22), and closed circles denote sequencing by dideoxy method (18, 19). Overlapping sequences were obtained by sequencing the parent plasmid YEpFAS2 by dideoxy method, using appropriate oligonucleotide primers. The DNA fragments used for 5' and 3' end SI endonuclease mapping are shown in p3.3 and p2.0, respectively.

The use of aminopeptidase was for the purpose of isolating the blocked NH₂-terminal peptides of α and β subunits (to be published). However, as a by-product of these studies, several peptides were obtained in pure state and sequenced.
Yeast Fatty Acid Synthase

The translation termination site is identified and underlined. The translation termination site is identified by a star. The underlined amino acid sequences represent the isolated tryptic peptides whose sequences were determined by automated Edman degradation (see "Materials and Methods").

Fig. 3. Nucleotide sequences of FAS2 gene, its flanking regions and the predicted amino acids of a subunit of yeast fatty acid synthase. Numbering of the nucleotides starts with the A of the first ATG Met codon. The numbers on the right represent either the last nucleotide or the amino acid as appropriate. In the 5' noncoding region the putative transcription initiation site and the possible translation initiation signal CAACACCGAA are marked and underlined. The acyl carrier protein, the putative β-ketoacyl reductase, and the β-ketoacyl synthase synthase sites are so indicated and underlined based on the published protein sequences (9, 10). In the 3' noncoding region the transcription termination site, and the TAGT, TAGTG, and TGTGAA sequences (27) are identified and underlined. The translation termination site is identified by a star. The underlined amino acid sequences represent the isolated tryptic peptides whose sequences were determined by automated Edman degradation (see "Materials and Methods").
Yeast Fatty Acid Synthase

Table I

| Amino acids (Codon) | Codon usage and frequency used |
|---------------------|-------------------------------|
| Ala (GCT) 100       | (GGA)*  92 (GGG)*  92           |
| Arg (CGT) 18        | (CGA)*  22 (CCG)*  22           |
| Lys (AAA)* 55       | (CTT)*  166 (CTC)*  166          |
| Met (ATG) 47        | (CTG)*  145 (CA)  145            |
| Pro (CCT) 47        | (CAC)*  74 (CCA)*  74            |
| Thr (ACT) 54        | (ACG)*  145 (ACG)*  145          |
| Val (GTT) 67        | (GTA)*  122 (GTA)*  122          |
| END (TAA) 1         | (TAA)  1 (TGA)  1                |

* Values reported (23).
* Indicates rarely used codons in highly expressed yeast genes (28, 29).

Fig. 4. HPLC separation of trypsin digest of cyanogen bromide peptides of yeast fatty acid synthase. The column was developed with a gradient of CH3CN in method A as described under “Materials and Methods.” Fractions T1, T2, T3, T4, and T5 were arbitrarily collected at the times indicated and used for further purification and sequencing as described under “Materials and Methods.”

between residues 40 to 63 and 971 to 982 as reported earlier (8).

Transcription Initiation and Termination Sites of FAS1 Gene—Since the yeast genomic insert in YEpFAS2 is about 11.4 kbp and the size of the mRNA coding for the α subunit is about 6.0 kb (Fig. 1), it was necessary to determine the beginning and the end of the transcription unit. This information would also be useful in establishing the extent of the coding region in the clone YEpFAS2.

In order to determine the transcription initiation site, we labeled the proximal AvaII site in p3.3 at the 5' end (see Fig. 2), and the AvaII-HindIII fragment was then separated and hybridized with yeast total RNA in 80% formamide at 37 or 47 °C (for details, see legend to Fig. 5). After S, endonuclease digestion, the resistant DNA hybrid was analyzed on denaturing polyacrylamide gels. On autoradiography, two fragments with the estimated sizes of 170 and 180 bases were found to be protected; however, these DNA fragments were lost when the hybridization was performed at 47 °C (Fig. 5A) presumably because these fragments are AT-rich (64%) resulting in inefficient hybridizations. These results indicate that the initiation site for transcription maps at about 40 bases upstream of the first ATG codon in the sequence shown in Fig. 3.

The mapping of the 3' end of the transcript was done by labeling the distal AvaII site in the plasmid p2.0 (see Fig. 2) at the 3' end. The AvaII-HindIII fragment was excised and hybridized with yeast total RNA and digested with S, endonuclease as described above. The S, resistant hybrids were analyzed by electrophoresis on denaturing polyacrylamide gel. After autoradiography a band with an estimated length of 800 nucleotides was obtained at hybridization temperatures of 37 and 47 °C (Fig. 5B). These results indicate that the transcription termination occurs at about 300 bases downstream of the TAG stop codon in the sequence presented in Fig. 3. This was confirmed when MluI site at the 3' end of p2.0 (Fig. 2) was labeled and the MluI-HindIII fragment was isolated and hybridized to the yeast RNA. A DNA fragment of about 90 nucleotides was protected by the mRNA (as shown in Fig. 5C).

Discussion

The plasmid YEpFAS2 contains a yeast genomic fragment that codes for the entire α subunit of yeast fatty acid synthase. The nucleotide sequence presented here shows an unique open reading frame that could code for a protein of 1894 amino acids starting with Met at position 1. There is no striking homology in the 5' and 3' flanking sequences of FAS1 and FAS2 genes that could suggest common regulatory functions.
TABLE II

Amino acid sequence of the purified peptides of yeast fatty acid synthase

Purified cyanogen bromide-trypic peptides of yeast fatty acid synthase were sequenced by gas phase Protein Sequencer (Applied Biosystem) as described under "Materials and Methods." Amino acid sequences of peptides 1–8 and 9–10 are matched with the deduced amino acid sequence of FAS2 and FAS3 genes (8), respectively.

| Amino acid sequence and no. of residues | Position matched in deduced amino acid sequence | Source of peptide (trypic pool) elution method and time (min) |
|---------------------------------------|-----------------------------------------------|---------------------------------------------------------------|
| 1. EIYYTPDPSELAAK 14                 | 87–100                                        | (T3) D 35                                                   |
| 2. EEPAK 5                           | 101–105                                       | (T5) C 25                                                  |
| 3. TAIKD 5                           | 507–511                                       | (T1) B 27                                                   |
| 4. GGGQAVHPDYLYGAITEDR 20            | 1857–1676                                     | (T1) C 23                                                  |
| 5. SNGVGYVDELSINVENDTFIER 25        | 1772–1796                                     | (T1) B 27                                                   |
| 6. SLGGGAALK 9                       | 1834–1842                                     | (T1) B 25                                                   |
| 7. NAPAVELHGNK 12                    | 1852–1863                                     | (T1) B 26                                                   |
| 8. VSISHDLLQAVAVAVST 17              | 1876–1892                                     | (T5) B 38                                                   |
| 9. ILPEPTEGFAADEPTTPAEVLGK 24       | 40–63                                         | (T5) C 23                                                   |
| 10. SLLDKPDEAIEK 12                  | 971–982                                       | (T2) D 30                                                   |

Note: The SI endonuclease mapping of the 5' and 3' ends of FAS2 DNA insert in p2.0 at the site of translational initiation (8) including the Met codon. This sequence may serve as a signal that would influence translation efficiency of the mRNA for both α and β subunits leading to stoichiometric accumulation of the subunits and subsequent association into αβ complex.

The SI endonuclease mapping of FAS2 gene at the 5' end indicated that there is a transcriptional initiation site at −26 and −36 bp (Fig. 5A). There is a TATAATTA sequence at −51 bp that could serve as TATA box. The next upstream TATA box is at −118 bp. Polypyrimidine tracts are present between −500 and −210 bp in the sequence presented; however, sequences are found in the 5'-untranslated region of most yeast polII genes (26). The 3' end of the transcriptional unit appears to be about 300 bp downstream of the translational termination site (Fig. 5, B and C). The eukaryotic polyadenylation signal AATAAA has not been found in the downstream sequence. However, signals for transcription termination and polyadenylation TACTA and TATGT (26) are found in the upstream of the transcription termination site (Fig. 3, underlined).

In the entire sequence presented in Fig. 3, there is no intron-specific sequence TACTAAC (27) indicating that the open reading frame of FAS2 represents the only exon which codes respectively. C, MluI site of the DNA insert in the same p.2.0 fragment used in B was labeled at the 3' end and the 600-bp Mlu-HindIII fragment was isolated and hybridized with yeast total RNA (30 μg) at 37 °C. S, endonuclease treatment and analysis of the protected DNA fragments were carried out as in B, lanes 1 and 6, X174 DNA cut with HaeIII and used as standard; lanes 2, denatured DNA; lane 3, DNA hybridized with 30 μg of yeast tRNA at 37 °C; lanes 4 and 5, DNA hybridized with yeast RNA at 37 and 47 °C, respectively.

TABLE III

Similarity of amino acid sequence of the putative β-ketoacyl reductase domain of the α subunit of yeast fatty acid synthase and other NADPH enzymes

| Putative β-ketoacyl reductase of α subunit of yeast fatty acid synthase residues (1093–1043) | E | V | G | P | W | G | S | A | R | T | K |
| Enoyl reductase of β subunit of yeast fatty acid synthase | F | G | S | G | F | S | A | D | T |
| Enoyl reductase of goose uropygial gland fatty acid synthase | V | F | T | T | V | G | S | A | E | K |
| Horse alcohol dehydrogenase | F | S | T | G | V | G | S | A | V | K |

Note: a See Ref. 8. b See Ref. 24. c See Ref. 25.
for the \( \alpha \) subunit protein of 1894 amino acids and that the adjoining sequences are not part of FAS2 gene. In the deduced protein sequence we could identify the serine site of attachment of the prosthetic group, 4'-phosphopantetheine, of the ACP domain and the active cysteine–SH of the \( \beta \)-ketoacyl synthase, the condensing enzyme, by comparison with the published amino acid sequences (9, 10). The amino acid sequence between residues 176–192 matches exactly with the sequence reported for the serine site of ACP domain; however, leucine, residue 175, follows aspartic acid, which was missing in the sequence reported (9). The amino acid sequence between residues 1307–1312 matches exactly with the peptide sequence published for \( \beta \)-ketoacyl synthase (10).

The amino acid sequence of \( \beta \)-ketoacyl reductase is not known. However, by comparison with the consensus sequences (Gly–Ser–Ala) for NADPH binding site for the enoyl reductase of the yeast \( \beta \) subunit (8), goose uropygial enoyl reductase (24) and horse alcohol dehydrogenase (25), a site between residues 1038 and 1040 has been located (see Fig. 3 and Table III). The -Gly-Ser-Ala- sequence occurs only once in the entire amino acid sequence of \( \alpha \) subunit and could probably be the site of \( \beta \)-ketoacyl reductase.

The amino acid composition of the deduced protein matches closely with that reported for yeast fatty acid synthase \( \alpha \) subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23). The percent fraction of each amino acid in the deduced sequence is comparable to an average low, as in the case of the subunit (Table I) (23).
the observed stoichiometry of α and β subunits in yeast could result from the equal translational efficiency of the mRNAs of both the subunits, as mentioned above.

Our calculated molecular weight for the β subunit of yeast synthase from 1980 amino acids, deduced from the nucleotide sequence, is 220,077 (8), while that of the α subunit is 207,863, as stated earlier. These values are the reverse of what had been estimated for the molecular weights of α subunit (213,000) and β subunits (203,000), based on their mobilities in SDS-polyacrylamide gel electrophoresis where the β subunit always moves faster than the α subunit (2). The reason for this discrepancy is not clear at this time. It is possible, however, that the slow mobility of the α subunit in SDS gel may be due to the presence of phosphopantetheine prosthetic group, unique structural features and/or to post-translational modifications such as glycosylation. Though there are six potential N-glycosylation sites (Asn-X-Ser/Thr) in the deduced amino acid sequence, our tests for the presence of carbohydrates in the protein were negative.

It is noteworthy that E. coli ACP migrates anomalously in SDS-polyacrylamide gels (32). Therefore, the anomalous behavior of the α subunit in SDS gels could be due to its ACP domain. We find, for instance, that E. coli ACP with a calculated molecular weight of 7985 (31) migrated in SDS gel in a distinctly anomalous fashion with an apparent molecular weight of 16,500, in agreement with Rock and Cronon (32). Performic acid oxidation of the E. coli ACP did not change its position on the gel, suggesting that the slower mobility was not due to dimerization through disulfide bond formation (data not shown). It is conceivable, therefore, that the anomalous behavior of the α subunit in SDS gel electrophoresis may be due to its ACP moiety in a manner similar to that of E. coli ACP. Neither performic acid oxidation of the α subunit protein nor β-elimination of the 4'-phosphopantetheine (33, 34) alter its mobility in SDS gel. Moreover, in SDS-urea gel electrophoresis, in which E. coli ACP migrates according to its expected molecular weight of about 8,000 (32), the α subunit shows slower mobility than that of β subunit. The reason for this behavior is not apparent at this time. However, it is of interest to note that upstream of this site for attachment of 4'-phosphopantetheine (Ser-180) in the deduced amino acid sequence (residues 108–138), there is a cluster of prolines and alanines; out of the 31 amino acids in this stretch there are 20 alanines and 8 prolines. It is not known whether such a sequence motif could result in abnormal folding of the protein that would hinder its complete unfolding in SDS, hence its anomalous mobility in SDS-polyacrylamide gels. Moreover, anomalous behavior of proteins is not uncommon in SDS gels (35, 36).

Examination of the amino acid sequence at the ACP domain of the α subunit shows no similarity to that of E. coli. However, when the predicted secondary structure of the ACP region of α subunit, according to Finkelstein (37), was compared to the proposed secondary and tertiary structural model for E. coli ACP (32, 38, 39), a remarkable similarity was noted. In these studies, the amino acid sequence of the ACP domain of α subunit (residues 140–226) was analyzed. The results showed that this region is composed of four α-helical structures interrupted by three β turns. Ser-180 is located in a β turn similar to that of E. coli ACP (32). The four α-helical regions can be folded in such a way as to mimic the proposed structure of E. coli ACP based on NMR studies (39) as shown in Fig. 7. The most striking feature of the predicted model is the conservation of Phe-194 and its location at a distance from the prosthetic group comparable to that of Phe-50 in the E. coli ACP. The amino acid sequence in this region -Glu-Phe-Gly-Thr- of the yeast ACP domain and -Glu-Phe-Asp-Thr- of the E. coli ACP (31) is highly conserved and may signify the suggested role of this Phe residue in the acylation of the prosthetic group in early stages of acyl chain elongation (38, 39). The Ile-172 and Phe-217 may play similar roles as Ile-69 and Phe-28 of E. coli ACP by interacting with the prosthetic group (39). Even though the model proposed in Fig. 7 is very speculative, its overall similarity to the structure of E. coli ACP is striking despite the considerable variations in the primary structures of the two proteins (31).

The discrepancy in molecular weight and mobility in SDS gels does not reflect on the correctness of the sequence presented here for the following reasons. 1) In the entire sequence there are no intron-specific sequences, as stated above. The sequence TACTAC found at −375 bases might not serve as processing signal because of the stringent requirement for TACTAAC (27). 2) An Aa2I subclone of the PASF2 insert in YEpFAS2 containing the entire sequence starting from −700 bases complements all known fas2 mutants and codes for full-length α subunit (data not shown). 3) The insert of the upstream p2.0* (see Fig. 2) hybridizes to about 1.5-kb mRNA, and not to the 6.0-kb fatty acid synthase mRNA, suggesting that p2.0* is a different gene. S1 endonuclease mapping of this p2.0* gene shows that the 1.5-kb mRNA starts at about −2000 bp upstream of the HindIII and probably terminates at about −500 bp in the sequence as shown in Fig. 3. Both genes are transcribed with the same polarity. 4) Finally, starting from the beginning of the sequence presented here (−1350 bases), there is an open reading frame terminating at −500 bases, coding for a protein of approximately 30 kDa. However, the nucleotide sequence between −500 and 1 does not appear to code for a protein of any consequence. Therefore, the data presented here in Fig. 3 represents a unique sequence coding for the full-length yeast fatty acid synthase α subunit.

Another possibility we have considered is that the β subunit undergoes post-translational proteolysis, resulting in loss of a
peptide of about 150 amino acids. However, there is no way that such a peptide would have been cleaved from the NH2-terminal end of the molecule and lost on purification, since we have isolated and sequenced a peptide that has an amino acid sequence that matches the predicted sequence near the NH2-terminal end (residues 40-63) (Table II). Moreover, we have cloned FAS1 gene in an expression vector which, when expressed in E. coli, yields β subunit protein having the same mobility as the wild type β subunit (data not shown). Although possible, it would be hard to believe that the products produced in both E. coli and yeast would undergo the same modification.

Recently, Schweizer and co-workers (40) sequenced the FAS1 gene and reported a reading frame of 5535 bp corresponding to a β subunit protein of 1845 amino acids. This is shorter than the 5940 bp that we reported for the FAS1 gene or about 135 amino acids less than what we predicted for the β subunit protein (8). In comparing the two sequences, we have reconfirmed our sequence of FAS1 gene and found that the discrepancy lies on an additional G in the sequence CCAA-GAGA G TGAGTTG reported by Schweizer et al. (40) at position 7002. This additional nucleotide in their sequence resulted in premature termination of the open reading frame as reported by them. Also, there are many differences between their nucleotide sequence and ours at the 3' end of the gene. Our data on the S1 mapping of the 3' end and the nucleotide sequence show correctly that the FAS1 gene codes for a protein of 1980 amino acids (8). Moreover, as shown in Fig. 1, the mRNA for the β subunit is larger than that for the α subunit, an observation which reaffirms the differences in the calculated molecular weights for the two subunits.

The clone 102B5, the truncated FAS2 gene isolated by immunological screening method that we reported earlier (7), must have the 3' end of the gene because the homologous region between YEpFAS2 and 102B5 shows that the truncated gene could code for the β-ketoacyl synthase and the putative β-ketoacyl reductase, but not for the acyl carrier protein domain, which lies at the 5' end of the FAS2 gene. The transcription of this clone in E. coli might have resulted from a promoter-like sequence either within the insert or in the ColE1 plasmid, resulting in the production of protein antigenically reactive to the anti-yeast fatty acid synthase antibodies.

Finally, our nucleotide-derived amino acid sequence helped establish the order of the domains of ACP and the other two partial activities as 5'-ACP/β-ketoacyl reductase/β-ketoacyl synthase-3' which is different from the order determined by genetic mapping by Schweizer and co-workers (6). Our identification of a β-ketoacyl reductase, however, is tentative and is based on the consensus -Gly-Ser-Ala- sequence. Our sequence shows that the ACP domain is at the NH2-terminal end of the multifunctional molecule with an interdomain of about 950 amino acids between it and the putative β-ketoacyl reductase region. On the other hand, a stretch of about 300 amino acids lie between the reductase and the β-ketoacyl synthase which is about 500 amino acids from the COOH terminus of the molecule. The very long interdomainal separation of the ACP and the synthase may explain why the bifunctional reagent dibromopropionane did not result in cross-linking within α subunit of the pantetheine-SH and the active cysteine-SH of the β-ketoacyl synthase (41). However, it did link all α subunits by cross-linking the cysteine-SH of one subunit to the cysteine-SH of the adjacent subunit. This long stretch of amino acids, therefore, would have facilitated this cross-linking and also help arrange the α subunit (plate-like) with the β subunit (arch-like) in the structures depicted by our model of active site of yeast fatty acid synthase (2). This model was based on electron microscopic studies of negatively stained yeast synthase and depicts the synthase as an ovate structure containing on its short axis, plate-like proteins (α subunits) around which six arch-like proteins (β subunits) are distributed, three on either side. The plate-like proteins are organized in such a way that the cysteine-SH of the β-ketoacyl synthase is juxtaposed within 2 Å from the pantetheine-SH of the adjacent α subunit. The location, therefore, of the pantetheine prosthetic group on Ser-179 and the active cysteine-SH of β-ketoacyl synthase on Cys-1312 would make it possible to construct the site for carbon-carbon bond formation, hence, chain elongation to produce long chain fatty acyl groups.

Acknowledgment—We would like to thank Professor Florante Quiocio for his help in the calculation and prediction of the secondary structure of the α subunit protein.

REFERENCES

1. Wakil, S. J., Stoops, J. K., and Joshi, V. C. (1983) Annu. Rev. Biochem. 52, 537-579
2. Stoops, J. K., Awad, E. S., Arslanian, M. J., Gunberg, S., Wakil, S. J., and Oliver, R. M. (1978) J. Biol. Chem. 253, 4464-4475
3. Wyman, J. (1980) Eur. J. Biochem. 112, 431-442
4. Schweizer, E. (1984) in Fatty Acid Metabolism and Its Regulation (Numa, S., ed) pp. 59-83, Elsevier Scientific Publishing Co., Amsterdam
5. Burkhl, C., Castroph, H., and Schweizer, E. (1972) Mol. Gen. Genet. 119, 315-322
6. Schweizer, M., Roberts, L. M., Holtkie, J., Takabayashi, K., Muller, G., Hojo, U., Hollerer, E., Schuh, B., and Schweizer, E. (1986) Biochem. Soc. Trans. 14, 572-574
7. Kuziora, M. A., Chalmers, J. H., Jr., Douglas, M. G., Hitzeman, R. A., Matteick, J. S., and Wakil, S. J. (1983) J. Biol. Chem. 258, 11648-11653
8. Chirala, S. S., Kuziora, M. A., Spector, D. M., and Wakil, S. J. (1987) J. Biol. Chem. 262, 4231-4240
9. Schreckenbach, T., Wobser, H., and Lynen, F. (1977) Eur. J. Biochem. 80, 13-23
10. Kreese, G.-B., Steber, L., Oesterhelt, D., and Lynen, F. (1977) Eur. J. Biochem. 79, 181-190
11. Sherman, F., Fink, G., and Lee, E. (1975) Methods in Yeast Genetics, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
12. Miller, J. H. (1972) Experiments in Molecular Genetics, p. 433, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
13. Thomas, P. S. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 5201-5205
14. Birnboim, H. C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523
15. Maniatis, T., Fritsch, E. F., and Sambrook, J. (eds) (1982) Molecular Cloning: A Laboratory Manual, pp. 292-293, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
16. Maxam, A. M., and Sharp, P. A. (1977) Cell 12, 712-723
17. Maxam, A. M., and Gilbert, W. (1980) Methods Enzymol. 65, 499-560
18. Sanger, F., and Coulson, A. R. (1975) J. Mol. Biol. 94, 441-448
19. Messing, J. (1983) Methods Enzymol. 101 C, 20-78
20. Schwindinger, W. F., and Warner, J. R. (1984) Nucleic Acids Res. 12, 601-604
21. Kuziora, M. A. (1984) Molecular Cloning and Characterization of the Yeast Fatty Acid Synthase Genes. Ph.D. thesis, Baylor College of Medicine, Houston, TX
22. Chirala, S. S., and Wakil, S. J. (1986) Gene (Amst.) 47, 297-306
23. Wieland, F., Renner, L., Verfurth, C., and Lynen, F. (1977) Eur. J. Biochem. 49, 159-197
24. Poulose, A. J., and Kolattukudy, P. E. (1983) Arch. Biochem. Biophys. 220, 655-656
25. Branden, C., and Kibbland, H. (1980) in Dehydrogenases Requiring Nicotinamide Coenzymes (Jeffrey, J., ed) pp. 40-84, Birkhaeuser Verlag, Boston
26. Zaret, K. S., and Sherman, F. (1982) Cell 28, 563-573
27. Langford, C. J., and Gallwitz, D. (1983) Cell 33, 519-527
28. Guthrie, C., and Abelson, J. (1982) in Molecular Biology of Yeast
Yeast Fatty Acid Synthase

29. Bennetzen, L. J., and Hall, B. D. (1982) J. Biol. Chem. 257, 3026–3031
30. Kyte, J., and Doolittle, R. F. (1982) J. Mol. Biol. 157, 105–132
31. Vanaman, T. C., Wakil, S. J., and Hill, R. L. (1968) J. Biol. Chem. 243, 6420–6431
32. Rock, C. O., and Cronan, J. E., Jr. (1979) J. Biol. Chem. 254, 9778–9785
33. Stoops, J. K., Arslanian, M. J., Yang, H. O., Aune, K. C., Vanaman, T. C., and Wakil, S. J. (1975) Proc. Natl. Acad. Sci. U. S. A. 72, 1940–1944
34. Pugh, E. L., and Wakil, S. J. (1965) J. Biol. Chem. 240, 4727–4733
35. Dunker, A. K., and Rueckert, R. R. (1969) J. Biol. Chem. 244, 5074–5080
36. Kulomaa, M. S., Weigel, N. L., Kleinsek, D. A., Beattie, W. G., Conneely, O. M., Marc, C., Zarucki-Schulz, T., Schrader, W. T., and O’Malley, B. W. (1986) Biochemistry 25, 6244–6251
37. Finkelstein, A. V. (1975) Dokl. Akad. Nauk SSSR 223, 744–747
38. Mayo, K. H., Tyrell, P. M., and Prestegard, J. H. (1983) Biochemistry 22, 4485–4493
39. Mayo, K. H., and Prestegard, J. H. (1985) Biochemistry 24, 7834–7838
40. Schweizer, M., Roberts, L. N., Holte, H.-J., Takabayashi, K., Hollerer, E., Hoffmann, B., Muller, G., Kottig, H., and Schweizer, E. (1986) Mol. Gen. Genet. 203, 479–486
41. Stoops, J. K., and Wakil, S. J. (1981) J. Biol. Chem. 256, 8364–8370