JOURNAL OF BACTERIOLOGY, July 1991, p. 4379–4385
0021-9193/91/144379-07$02.00/0
Copyright © 1991, American Society for Microbiology
Vol. 173, No. 14

Cloning, Characterization, and High-Level Expression in *Escherichia coli* of the *Saccharopolyspora erythraea* Gene Encoding an Acyl Carrier Protein Potentially Involved in Fatty Acid Biosynthesis

W. PETER REVILL AND PETER F. LEADLAY*

Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, United Kingdom

Received 26 December 1990/Accepted 29 April 1991

The erythromycin A-producing polyketide synthase from the gram-positive bacterium *Saccharopolyspora erythraea* (formerly *Streptomyces erythraeus*) has evident structural similarity to fatty acid synthases, particularly to the multifunctional fatty acid synthases found in eukaryotic cells. Fatty acid synthesis in *S. erythraea* has previously been proposed to involve a discrete acyl carrier protein (ACP), as in most prokaryotic fatty acid synthases. We have cloned and sequenced the structural gene for this ACP and find that it does encode a discrete small protein. The gene lies immediately adjacent to an open reading frame whose gene product shows sequence homology to known β-ketoacyl-ACP synthases. A convenient expression system for the *S. erythraea* ACP was obtained by placing the gene in the expression vector pT7-7 in *Escherichia coli*. In this system the ACP was efficiently expressed at levels 10 to 20% of total cell protein. The recombinant ACP was active in promoting the synthesis of branched-chain acyl-ACP species by extracts of *S. erythraea*. Electrospray mass spectrometry is shown to be an excellent method for monitoring the efficiency of in vivo posttranslational modification of ACPs.

The macrolide antibiotic erythromycin A is produced by the gram-positive filamentous bacterium *Saccharopolyspora erythraea* (formerly *Streptomyces erythraeus* [32]), starting from simple fatty acyl-coenzyme A (CoA) precursors. The biosynthesis of erythromycin and other polyketides shows important similarities (25) to the more familiar process of fatty acid biosynthesis, and this has led to renewed interest in the structure, organization, and mechanism of action of fatty acid synthase in *S. erythraea* and in *Streptomyces* spp. Initial reports indicated that *S. erythraea* (39) and *Streptomyces coelicolor* (17) might contain a type I fatty acid synthase complex like that of *Mycobacterium smegmatis* (9), *Brevibacterium ammoniagenes* (30), or *Saccharomyces cerevisiae* (43, 48), in which multifunctional polypeptides are tightly associated in a complex of high molecular weight. More recently, a small discrete acyl carrier protein (ACP) has been identified and purified from *S. erythraea* (21) on the basis of its ability to stimulate the incorporation of malonyl-CoA into acyl-ACP in a cell-free system. This implies that the fatty acid synthase of *S. erythraea* may be patterned on the type II system found generally in prokaryotes, consisting of freely dissociable, monofunctional enzyme components. The best-studied system of this kind is that of *Escherichia coli* (see reference 52 for a recent review), but a discrete ACP is also apparently involved in fatty acid synthase from gram-positive bacteria more closely related to *S. erythraea* (1, 27) and which, like *S. erythraea*, characteristically produce iso- and anteiso-terminally branched fatty acids. Alternatively, there may exist more than one type of fatty acid synthase activity, as in *Euglena gracilis* (13). The recent (11) finding that the erythromycin-producing polyketide synthase of *S. erythraea* is a multifunctional complex resembling a type I fatty acid synthase has also prompted further examination of the subunit structure of the *S. erythraea* fatty acid synthase.

In this paper, we report the cloning and characterization of the structural gene for the ACP purified previously from *S. erythraea*, which has given further insight into the organization of the putative synthase. Electrospray mass spectrometry (16, 22) is shown to provide a powerful and convenient means of monitoring the correct posttranslational modification of the *S. erythraea* ACP when expressed in a heterologous bacterial host.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. *S. erythraea* NRRL 2338, obtained from the Northern Regional Research Laboratories, Peoria, Ill., was maintained in M1-102 agar medium and grown in M1-102 medium at 30°C (28). *E. coli* TG1 [K-12 Δ(*lac-pro*) *supE thi hsdD5*/F' *traD36 proA*$^+$*B*$^+$ *lacI*^{q} *lacZ* ΔM15] (18a), the gift of E. S. Ward, MRC Laboratory of Molecular Biology, Cambridge, United Kingdom, was maintained on M9 minimal medium supplemented with 0.1 mM thiamine hydrochloride (34) to maintain selection for the F' episome. *E. coli* NM538 (18) was the gift of M. Carrington, Department of Biochemistry, University of Cambridge. *E. coli* K38 (described in reference 40) and the expression vector pT7-7 were the gifts of S. Tabor, Harvard Medical School. Plasmid pT7-7 is a derivative of pT7-5 (49, 50) and has the strong φ10 T7 promoter (15) and the translation initiation site for the T7 gene *10* protein (15) inserted into pBR322, upstream of a polylinker derived from pUC12 (48a). Plasmid pGP1-2 carries the gene for T7 RNA polymerase, under the control of the p_L promoter (50). The genomic library of *S. erythraea* DNA in λEMBL3 was constructed by J. Cortes (11). Strains of *E. coli* were grown in 2× TY broth (1% yeast extract, 1% tryptone, 0.5% NaCl) at 37°C unless otherwise specified.

S. erythraea total DNA was prepared essentially as de-

* Corresponding author.

scribed by Maniatis et al. (34). DNA manipulations were generally performed by standard procedures for both *Streptomyces* spp. (24) and *E. coli* (34). Plasmid DNA was prepared by the method of Kieser (31). Restriction maps of inserts in clones of bacteriophage lambda were obtained as described by Rackwitz et al. (38).

Cloning of the *S. erythraea* ACP gene. A 38-mer oligonucleotide was designed, on the assumption of typical *Streptomyces* codon bias (6), which would encode amino acids 1 to 12 of the amino-terminal sequence of the ACP, isolated from *S. erythraea* as described previously (21). This oligonucleotide, 5′-ATGGACCG[C/G]AAGGAGATCTTCGAG CG[C/G]ATCGAGCAGGT-3′, was synthesized on a Biosearch Cyclone automated synthesizer and was shown to hybridize to a specific target DNA in restriction digests of *S. erythraea* genomic DNA (data not shown). It was used as a hybridization probe to screen a genomic library of *S. erythraea* DNA in λEMBL3 (18), and one of the positively hybridizing clones (λEMBL-PR.1) was selected for further study. A restriction map of the 17.8-kbp insert in λEMBL-PR.1 is shown in Fig. 1.

DNA sequencing and sequence analysis. DNA was subcloned into M13mp18 (54) and sequenced by using the dideoxy chain termination method (42), adapted for use with modified T7 DNA polymerase (Sequenase) (56). Computer-assisted analysis of DNA sequences was performed by using the DBUTIL and DBAUTO programs of Staden (47) and the programs of the Genetics Computer Group (14). The MULT ALIGN (3) program was used for the simultaneous alignment of more than two protein sequences.

Construction and use of *E. coli* expression plasmid pFEX-1. An *Nde*I restriction site spanning the ATG start codon was introduced into the *S. erythraea* ACP gene, using the polymerase chain reaction (PCR) (41). The template for PCR was a clone in pUC19 (54) with a 560-bp *Bam*HI-*Sma*I insert containing the ACP gene. Universal sequencing primer and the mutagenic oligonucleotide 5′-TGAAGTCGACATATGG ACCGCAAGGAGAT-3′ were used as primers in a protocol based on that of Saiki et al. (41). Reaction mixtures contained 100 ng of template, 100 pmol of each primer, and 20 pmol of each deoxynucleoside triphosphate in 100 μl of 10 mM Tris HCl buffer, pH 8.3, containing 50 mM KCl, 2.5 mM $MgCl_2$, and 0.1 mg of gelatin per ml. An overlay of 200 μl of liquid paraffin was used to prevent evaporation. The template was initially denatured at 95°C for 6 min and then held at 55°C for 5 min. *Taq* DNA polymerase (2 U) was added, and the sample was subjected to 30 cycles of extension (1 min at 72°C), denaturation (1 min at 95°C), and annealing (1 min at 55°C) in a programmable heating block (Cambio, Cambridge, United Kingdom). Reaction products were analyzed on an 8% polyacrylamide gel. DNA of the correct size (451 bp) was eluted by diffusion into 1.0 ml of 10 mM Tris HCl buffer (pH 8.0)–0.1 mM EDTA and subcloned into M13mp18, and the inserts in several clones were sequenced. The ACP gene was excised from one of these clones as a 414-bp *Nde*I-*Eco*RI fragment and ligated into pT7-7 to give the expression plasmid pFEX-1 (see Fig. 3).

For expression of ACP, *E. coli* K38 pGP1-2/pFEX-1 was grown at 30°C on 2× TY medium containing 50 μg of kanamycin and 100 μg of ampicillin per ml. When the cultures reached an A_{600} of approximately 1.0, they were heated to 42°C for 30 min and then incubated at 37°C for a further 2 h to induce ACP production. Cell extracts were analyzed by using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (33). The solubility of overexpressed proteins was assessed as follows: induced cells were disrupted by sonication and the lysate was centrifuged at 4°C, first at 40,000 × *g* for 10 min and then at 160,000 × *g* for 2 h. Pellets and supernatant were analyzed by gel electrophoresis as described above.

Purification of ACP from recombinant *E. coli*. The procedure for purification of *S. erythraea* ACP was similar to that of Hale et al. (21). The culture (15 liters) was concentrated by using a Pellicon (Millipore) filtration unit and then centrifuged. The cells were suspended in 100 mM potassium phosphate buffer (pH 7.0)–1 mM EDTA–0.1 mM phenylmethanesulfonyl fluoride–1 mM dithiothreitol–10% (vol/vol) glycerol and disrupted by sonication (six 15-s pulses, with cooling intervals of 1 min; Ultrasonic Processor model W-225; Heat Systems Inc., Farmingdale, N.Y.). Purification of the ACP by fractionation with streptomycin sulfate and then ammonium sulfate precipitation at pH 4.0, followed by gel filtration on a Sephadex G-75 column (21), was monitored by using SDS-PAGE to detect a protein of M_r 6,000. As a final purification step, ACP (up to 40 mg of protein) was loaded onto a fast protein liquid chromatography (Pharmacia) MonoQ HR10/10 anion-exchange column (8-ml bed volume) in 50 mM Tris HCl buffer (pH 7.0)–0.2 mM EDTA and eluted by using a linear salt concentration gradient (0 to 1.0 M KCl; total volume, 120 ml) in the same buffer at a flow rate of 1 ml/min.

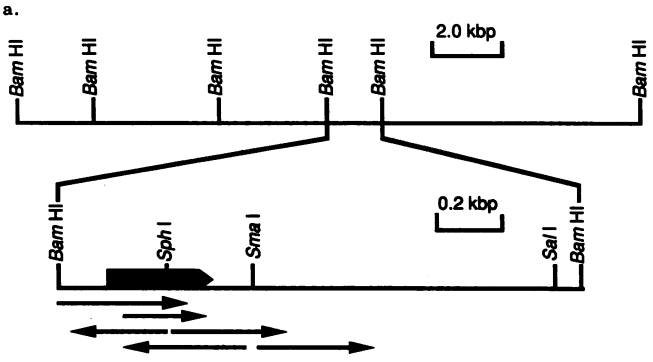
Enzyme and protein assays. *S. erythraea* ACP was assayed as described by Hale et al. (21). Protein concentrations were determined by the method of Bradford (8), using bovine serum albumin as a standard. For N-terminal analysis of proteins separated by SDS-PAGE, individual proteins were electroblotted onto a polyvinylidene difluoride membrane (35). The protein bands were visualized by Coomassie blue staining and excised. The proteins were sequenced with an Applied Biosystems model 470A pulsed-liquid protein sequencer fitted with an on-line model 120A analyzer for the detection of phenylthiohydantoin amino acids.

Electrospray mass spectrometry of *S. erythraea* ACP. Samples of *S. erythraea* ACP were kindly analyzed by Brian Green of VG Masslabs Ltd., Manchester, United Kingdom, on a VG quadrupole instrument fitted with an electrospray source and operated in positive ion mode. The sample (25 μl) was introduced as a solution (60 pmol/ml) in 50% aqueous methanol, acidified with 5% formic acid.

Nucleotide sequence accession number. The nucleotide sequence reported here has been submitted to the Genbank/EMBL Data Bank under accession number M64477.

RESULTS

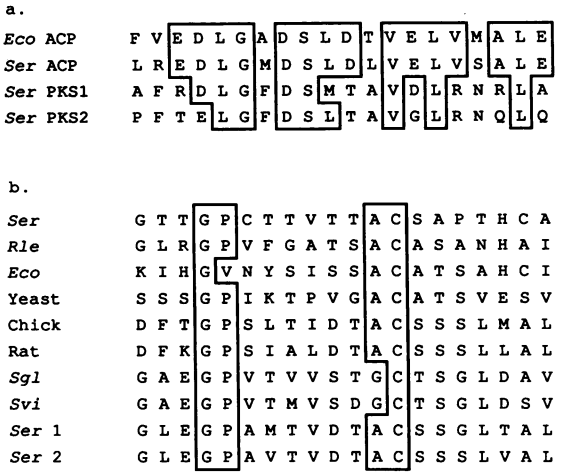
Cloning and characterization of *S. erythraea* DNA encoding the ACP gene. The genome of typical *Streptomyces* spp., and of related organisms such as *S. erythraea*, contains 69 to 74 mol% G+C, and this results in a very strong bias (up to 91% [6]) towards the use of G or C in the third position of codons. A synthetic 38-mer oligonucleotide was designed, taking advantage of this biased codon usage, on the basis of the known N-terminal amino acid sequence of the *S. erythraea* ACP isolated by Hale et al. (21). This oligonucleotide was used as a hybridization probe to screen a genomic library of *S. erythraea* DNA in λEMBL3. The restriction map of one positively hybridizing clone (λEMBL-PR.1) and the sequencing strategy used are shown in Fig. 1, together with the deduced amino acid sequence of the gene product. The inferred amino acid sequence exactly matches the N-terminal amino acid sequence determined previously (21), if the reading frame starts with the ATG at nucleotide 145. The 100



b.

```
CACCTTCGTCACCGCCAACCGCAACCAGGTCGTGCGCCTCCAGCCGCCGCTGGTGGTCAC
        10        20        30        40        50        60

CGAGGCCGAGATCGACCGGTTCGCCGAATCCGCCCCGCGCGGTCTGCGAAACCGCAACGGG
        70        80        90       100       110       120

                                M  D  R  K  E  I  F  E  R  I  E  Q
CATCCACAACAACTGGGAGTGATCATGGACCGCAAGGAGATCTTCGAGCGCATCGAGCAG
       130       140       150       160       170       180

 V  L  A  E  Q  L  G  I  P  A  E  Q  I  T  E  E  A  D  L  R
GTGCTGGCCGAGCAGCTGGGCATCCCCCGCCGAGCAGATCACCGAGGAGGCCGACCTGCGC
       190       200       210       220       230       240

 E  D  L  G  M  D  S  L  D  L  V  E  L  V  S  A  L  E  D  E
GAGGACCTCGGCATGGACTCGCTGGACCTGGTCGAGCTGGTCAGCGCGCTGGAGGACGAG
       250       260       270       280       290       300

 V  G  M  R  V  E  Q  S  Q  L  E  G  I  E  T  V  G  H  V  M
GTCGGCATGCGCGTGGAGCAGAGCCAGCTGGAGGGCATCGAGACCGTCGGGCACGTCATG
       310       320       330       340       350       360

 E  L  T  L  D  L  V  A  R  L  A  T  A  S  A  A  D  K  P  E
GAGCTGACGCTGGACCTCGTCGCCCGGCTGGCCACCGCCTCGGCCGCCGACAAGCCGGAA
       370       380       390       400       410       420

 A  A  S  *
GCGGCGTCGTGACCG
       430  435
```

FIG. 1. Nucleotide and deduced amino acid sequence of the *S. erythraea* ACP. (a) Restriction map of the 17.8-kbp insert in λEMBL-PR.1 and strategy used to sequence the ACP gene. The arrows indicate the length and direction of the sequence determined from individual M13 clones. The partial sequence data obtained for the gene 3′ to the ACP are not shown. (b) DNA sequence of the ACP structural gene and the inferred amino acid sequence. The putative ribosome binding site is underlined. Ser-39 is the inferred attachment site for the 4′-phosphopantetheinyl prosthetic group (Fig. 2).



a.

Eco ACP	F	V	E	D	L	G	A	D	S	L	D	T	V	E	L	V	M	A	L	E
Ser ACP	L	R	E	D	L	G	M	D	S	L	D	L	V	E	L	V	S	A	L	E
Ser PKS1	A	F	R	D	L	G	F	D	S	M	T	A	V	D	L	R	N	R	L	A
Ser PKS2	P	F	T	E	L	G	F	D	S	L	T	A	V	G	L	R	N	Q	L	Q

b.

Ser	G	T	T	G	P	C	T	T	V	T	T	A	C	S	A	P	T	H	C	A
Rle	G	L	R	G	P	V	F	G	A	T	S	A	C	A	S	A	N	H	A	I
Eco	K	I	H	G	V	N	Y	S	I	S	S	A	C	A	T	S	A	H	C	I
Yeast	S	S	S	G	P	I	K	T	P	V	G	A	C	A	T	S	V	E	S	V
Chick	D	F	T	G	P	S	L	T	I	D	T	A	C	S	S	S	L	M	A	L
Rat	D	F	K	G	P	S	I	A	L	D	T	A	C	S	S	S	L	L	A	L
Sgl	G	A	E	G	P	V	T	V	V	S	T	G	C	T	S	G	L	D	A	V
Svi	G	A	E	G	P	V	T	M	V	S	D	G	C	T	S	G	L	D	S	V
Ser 1	G	L	E	G	P	A	M	T	V	D	T	A	C	S	S	G	L	T	A	L
Ser 2	G	L	E	G	P	A	V	T	V	D	T	A	C	S	S	S	L	V	A	L

FIG. 2. Sequence comparison of *S. erythraea* putative fatty acid synthase ACP and β-ketoacyl-ACP synthase active-site residues with known fatty acid and polyketide synthases. (a) Alignment of the deduced amino acid sequence of the *S. erythraea* ACP (*Ser* ACP) with the *E. coli* fatty acid synthase ACP (*Eco* ACP; residues 28 to 47 [51]) and two of the ACP domains of the erythromycin synthase complex (residues 1428 to 1447 and 2858 to 2874) from *S. erythraea* (*Ser* PKS1 and PKS2 [11]). (b) Deduced amino acid sequence obtained from partial DNA sequence data of the gene 3′ to the ACP (*Ser*), aligned with active-site residues from known β-ketoacyl-ACP synthases (KAS). *Rle*, nodulation protein E from *Rhizobium leguminosarum* (46); *Eco*, fatty acid synthase KAS I from *E. coli* (29); yeast, fatty acid synthase KAS domain from *Saccharomyces cerevisiae* FAS A gene product (36); chick, fatty acid synthase KAS domain from chicken (55); rat, fatty acid synthase KAS domain from rat (44); *Sgl*, polyketide TCM I KAS from the tetracenomycin producer *Streptomyces glaucescens* (7); *Svi*, polyketide GRA I KAS from the dihydrogranaticin producer *Streptomyces violaceoruber* (45); *Ser* 1 and *Ser* 2, erythromycin KAS domains (residues 190 to 209 and 1658 to 1677) from *S. erythraea* open reading frame A (11). Alignments were done by using the program MULTALIGN (3).

nucleotides upstream of this potential start codon could encode part of a potential open reading frame transcribed in the same direction as the ACP gene with a stop codon at nucleotide 140. There is a potential ribosome binding site 10 nucleotides upstream of the ACP start codon, with reasonable complementarity to the 3′ end of 16S rRNA from *Streptomyces coelicolor* (26). A second potential open reading frame begins at nucleotide 429, overlapping the stop codon for the ACP, and continues downstream. The complete sequence of the latter open reading frame remains to be determined, but one portion of the amino acid sequence of the inferred gene product bears a very strong resemblance to the active site of known β-ketoacyl-ACP synthases (Fig. 2b).

The ACP gene encodes a protein of 95 amino acids, with a predicted molecular mass of 10,422.7 Da. A comparison of the amino acid sequence of the *S. erythraea* ACP with that of *E. coli* fatty acid synthase ACP (51) and with the ACP domains in one of the *S. erythraea* erythronolide synthase subunits (11) is shown in Fig. 2a. The site of attachment of the 4′-phosphopantetheine prosthetic group is predicted to be serine 39 (corresponding to serine 36 in *E. coli* ACP [51]), which lies in a region in which 17 of 21 consecutive residues are identical between *E. coli* and *S. erythraea*. With no gaps required for their alignment, the two sequences show an overall degree of homology of 40%. If similar amino acids are considered, the degree of homology increases to 64%. The *S. erythraea* ACP shows about 25% homology to each of two ACP domains from the *S. erythraea* polyketide synthase.

Expression of the *S. erythraea* ACP gene in *E. coli*. Expression of the *S. erythraea* ACP gene in *E. coli* was accomplished by using the T7 dual plasmid system developed by Tabor and coworkers (49, 50). An expression plasmid was constructed for the *S. erythraea* ACP gene by the strategy described in Materials and Methods and outlined in Fig. 3. A plasmid containing the ACP gene, on a 560-bp *Bam*HI-*Sma*I fragment of *S. erythraea* DNA cloned into pUC19, was used as the template for specific amplification and concomitant mutagenesis of the ACP gene by means of the PCR. The primers for PCR were universal sequencing primer and a mutagenic primer designed to introduce an *Nde*I site spanning the start codon for the ACP gene (41). The PCR-amplified products were cloned into M13, and 9 of 10 clones sequenced were found to possess the expected altered sequence, the other containing an additional single base substitution. The ACP gene was subcloned from M13 into

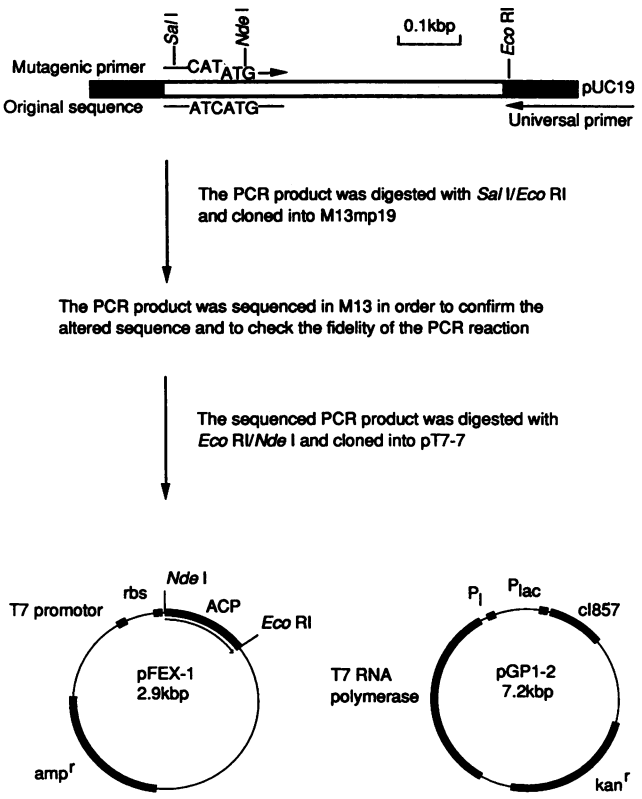


FIG. 3. Construction of expression plasmid pFEX-1. A 560-bp *Bam*HI-*Sma*I fragment containing the ACP gene, subcloned into pUC19, served as a template for PCR. The mutagenic primer oligonucleotide contained *Sal*I and *Nde*I restriction sites, which allowed the product to be cloned into M13mp19 for sequencing, prior to cloning into the *Nde*I site of pT7-7. Plasmid pGP1-2 contains the gene for T7 RNA polymerase under the control of a heat-sensitive repressor (49, 50). rbs, ribosome binding site.

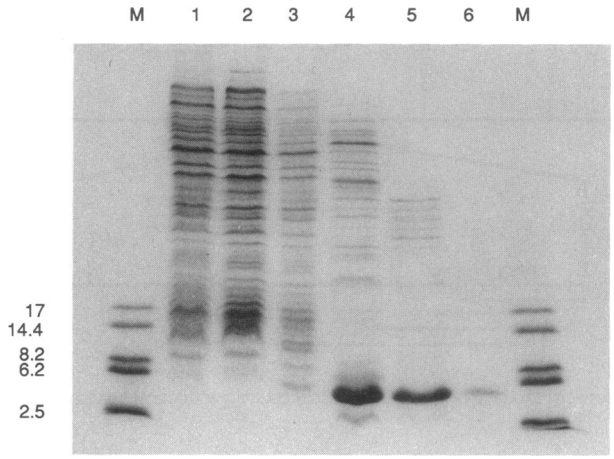


FIG. 4. Expression of *S. erythraea* ACP in *E. coli* K38 examined by 24% SDS-PAGE. Lane M contains low-molecular-weight markers (10^3). Lane 1 contains whole-cell lysate of *E. coli* K38(pT7-7, pGP1-2) grown at 30°C and induced at 42°C for 30 min. Lane 2 contains whole-cell lysate of *E. coli* K38(pFEX-1, pGP1-2) grown at 30°C but uninduced. Lane 3 contains whole-cell lysate of *E. coli* K38(pFEX-1, pGP1-2) grown at 30°C and induced. Lane 4 contains ACP purified by fractionation with ammonium sulfate. Lane 5 contains ACP fractions after gel filtration on Sephadex G-75. Lane 6 contains ACP after anion-exchange chromatography on Mono-Q HR 10/10. Details of the purification protocol (21) are described in the text.

the *Nde*I and *Eco*RI sites of pT7-7, which placed the ACP gene downstream of the T7 promoter, at an optimal spacing from an efficient ribosome binding site. The resulting expression plasmid (pFEX-1) was used to transform *E. coli* K38 containing plasmid pGP1-2 (50), which carries the structural gene for T7 RNA polymerase under the control of a temperature-sensitive repressor. SDS-PAGE was used to monitor the expression of *S. erythraea* ACP in *E. coli* K38 pGP1-2/pFEX-1 after heat induction. Typical results are shown in Fig. 4. Under these conditions, the *S. erythraea* ACP was obtained as a wholly soluble protein. The N-terminal amino acid sequence of the ACP, which represented at least 10% of total cell protein after induction, was determined after SDS-PAGE (33) by transfer to a polyvinylidene difluoride membrane and direct automated sequencing of the stained blot (35). The single sequence obtained matches the expected sequence for the first 38 residues, with no evidence for any posttranslational removal of the N-terminal methionine residue (23). At position 39, the expected site of attachment of the prosthetic group, an unidentified residue was found, removal of which caused a very large drop in signal.

Purification and characterization of *S. erythraea* ACP expressed in *E. coli*. The purification procedure developed previously by Hale et al. (21) was readily adapted for the purification of the *S. erythraea* ACP from *E. coli*. Typically, over 100 mg of ACP was obtained starting from 40 g (wet weight) of cells. The purified ACP was active in catalyzing the isobutyryl-CoA-dependent incorporation of radioactivity from ^{14}C-malonyl-CoA into acyl-ACP species, in the presence of extracts of *S. erythraea* depleted of ACP previously. Under the conditions used by Hale et al. (21), the addition of 30 μg of ACP catalyzed the incorporation of 1.5 nmol of malonyl-CoA into protein-bound radioactivity in 30 min. To determine the extent to which the *S. erythraea* ACP was modified during expression in *E. coli*, by the specific attachment of a 4′-phosphopantetheine prosthetic group, a sample of the purified ACP was analyzed by electrospray mass spectrometry (16, 22), which allows the exact mass of each ACP species to be separately determined. The mass of the apo-ACP was measured as 10,422.9 ± 0.6 Da (three determinations), which is in excellent agreement with the mass value (10,422.7 Da) calculated from the sequence of the structural gene. Similar agreement between observed (10,761.0 ± 1.6 Da; three determinations) and calculated (10,761.7 Da) mass values was also seen for the holo-ACP. The relative peak sizes indicated that about 70% of the total ACP was in the holo form, a ratio independently confirmed by high-pressure liquid chromatography (HPLC) analysis (38a).

DISCUSSION

We have reported the cloning, characterization, and sequencing of a potential fatty acid synthase ACP from the erythromycin-producing organism *S. erythraea*. The data from nucleotide sequence analysis are consistent with our previous proposal (21) that this ACP is a discrete protein and was not generated by adventitious proteolysis of a larger polypeptide. The region immediately 5′ to the ACP start codon, in the same reading frame, is strongly predicted to be noncoding. A potential open reading frame, transcribed in

the same sense as the ACP gene, terminates at a stop codon 5 nucleotides upstream. Further sequence analysis should determine the size and nature of this potential reading frame. The overall degree of sequence identity between *E. coli* and *S. erythraea* fatty acid synthase ACPs, over 40%, is markedly higher than the degree of sequence identity (about 25%) between the potential *S. erythraea* fatty acid synthase ACP and any of the polyketide synthase ACPs (7, 45) or ACP domains (11) that were tested. *E. coli* ACP can stimulate fatty acyl-ACP synthesis by cell extracts of *S. erythraea* that have been previously depleted of endogenous ACP (21), so the close similarity in primary structure evidently gives rise to a very similar three-dimensional structure. Expression of the *S. erythraea* ACP gene in *E. coli* gave an enzymatically active protein, indicating that the *S. erythraea* ACP is a substrate for the *E. coli* (holo)-ACP synthetase, which transfers the 4′-phosphopantetheine prosthetic group from CoA to the ACP. This is further evidence of the relatively broad substrate specificity of the synthetase (4, 5, 37).

We have measured the proportion of ACP present in the holo form by using the new technique of electrospray mass spectrometry (16, 22) to determine the exact mass of each component. The electrospray technique is direct, simple, and extremely sensitive; it has already been used to monitor the in vivo posttranslational modification of isopenicillin N synthetase (cyclase) (2) and of lipoyl domains of *Bacillus stearothermophilus* pyruvate dehydrogenase (12) and to study the thioesterase activity associated with both fatty acid synthase (54) and the erythromycin-producing polyketide synthase (10). Encouragingly, the experimentally determined values for the ACP were within one mass unit of those predicted from the sequence of the gene. The retention of the N-terminal methionine residue is consistent with the findings of Hirel et al. (23). The relative peak sizes in the mass spectrum indicated that approximately 70% of the total ACP was present as holo-ACP, the rest being apo-enzyme. The response factors of the two forms of ACP are expected to be identical, given their structural similarity and the fact that under the conditions used for the analysis they contain identical numbers of chargeable groups. In agreement with this conclusion, direct measurement of the relative concentrations of apo- and holo-ACPs after HPLC separation also showed that about 70% of the ACP was present as holo-ACP in these preparations. In one preparation of ACP, in which 2-mercaptoethanol had been used during the purification, evidence was obtained that some of the holo-ACP had formed a disulfide link with the 2-mercaptoethanol. We have also used electrospray mass spectrometry to characterize, in some detail, the acyl-ACP products of the potential *S. erythraea* fatty acid synthase. The results of these experiments will be presented elsewhere (8a). Spinach ACP-I, when expressed in *E. coli*, accumulates to a large extent (50%) as the 18:1Δ11(*cis*) thioester (19). The possibility cannot be excluded that a small proportion of the *S. erythraea* ACP is similarly modified but is separated from unacylated ACP during the purification procedure.

The nucleotide sequence analysis also provided evidence that the ACP gene is clustered together with at least one other gene that would be required for fatty acid biosynthesis, that for a β-ketoacyl-ACP synthase, although the exact size of this gene has not yet been established. The start codon for the latter gene overlaps the stop codon for the ACP, an arrangement that would permit translational coupling between the two genes (20), but detailed transcriptional analysis will be required to confirm this possibility. A portion of the β-ketoacyl-ACP synthase sequence containing the active-site cysteine is shown in Fig. 2b, together with the corresponding active-site sequences from known fatty acid and polyketide synthases. Experiments are in progress to define the exact extent of the putative fatty acid synthase gene cluster in *S. erythraea*. It will be of considerable interest to compare the *S. erythraea* gene arrangement with that of a recently characterized gene cluster for *E. coli* fatty acid synthase (25). Although the *S. erythraea* ACP has the catalytic activity expected of a component of the fatty acid synthase, it remains unclear whether or not this is its physiological role. However, it will now be possible to carry out specific disruption of the genes identified here to provide the answer to that question.

ACKNOWLEDGMENTS

We thank S. Tabor (Department of Biological Chemistry, Harvard Medical School, Cambridge, Mass.) for the provision of strains and plasmids, L. C. Packman of the Department of Biochemistry Protein Sequencing Facility (funded jointly by the Science and Engineering Council [S.E.R.C.] and the Wellcome Trust) for the automated determination of peptide sequences, and Mike Weldon of the oligonucleotide synthesis service of the Departments of Pathology and Biochemistry, University of Cambridge, funded by the Wellcome Trust, for the provision of synthetic oligonucleotides. We particularly thank Brian Green of VG Masslabs Ltd. for advice and for mass determinations using electrospray mass spectrometry and J. Staunton for helpful discussions.

This work was supported by the award of a studentship to W.P.R. from the S.E.R.C. (U.K.) Biotechnology Directorate. P.F.L is a member of the S.E.R.C. Cambridge Centre for Molecular Recognition.

REFERENCES

1. **Ahmad, P. M., L. A. Stirling, and F. Ahmad.** 1981. Partial characterization of fatty acid synthase of *Propionibacterium shermanii*. J. Gen. Microbiol. **127**:121–129.
2. **Aplin, R. T., J. E. Baldwin, Y. Fujishima, C. J. Schofield, B. N. Green, and S. A. Jarvis.** 1990. Molecular weight analysis of isopenicillin N synthetase by electrospray mass spectrometry. FEBS Lett. **264**:215–217.
3. **Barton, G. J., and M. J. E. Sternberg.** 1987. A strategy for the rapid multiple alignment of protein sequences. Confidence levels from tertiary structure comparisons. J. Mol. Biol. **198**:327–337.
4. **Beremand, P. D., D. D. Elmore, K. Dziewanowska, and D. Guerra.** 1989. Expression of an active spinach acyl carrier protein-I/protein A gene fusion. Plant Mol. Biol. **12**:95–104.
5. **Beremand, P. D., D. J. Hannapel, D. J. Guerra, D. N. Kuhn, and J. B. Ohlrogge.** 1987. Synthesis, cloning and expression in *Escherichia coli* of a spinach acyl carrier protein-I gene. Arch. Biochem. Biophys. **256**:90–100.
6. **Bibb, M. J., S. Biro, H. Motamedi, J. F. Collins, and C. R. Hutchinson.** 1989. Analysis of the nucleotide sequence of the *Streptomyces glaucescens tcmI* genes provides key information about the enzymology of polyketide antibiotic biosynthesis. EMBO J. **8**:2727–2736.
7. **Bibb, M. J., P. R. Findlay, and M. W. Johnson.** 1984. The relationship between base composition and codon usage in bacterial genes and its use for the simple and reliable identification of protein coding sequences. Gene **30**:157–166.
8. **Bradford, M. M.** 1978. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.

8a. **Bridges, A., P. F. Leadlay, W. P. Revill, and J. Staunton.** J. Chem. Soc. Chem. Commun., in press.

9. **Brindley, D. N., S. Matsumara, and K. Bloch.** 1969. *Mycobacterium phlei* fatty acid synthetase- a bacterial multienzyme complex. Nature (London) **224**:666–669.
10. **Caffrey, P., B. Green, L. C. Packman, B. J. Rawlings, J. Staunton, and P. F. Leadlay.** 1991. An acyl-carrier protein-thioesterase domain from the 6-deoxyerythronolide B synthase

of *Saccharopolyspora erythraea*: high level production in *Escherichia coli*, purification and characterisation. Eur. J. Biochem. **195**:823–830.
11. **Cortes, J., S. F. Haydock, G. A. Roberts, D. J. Bevitt, and P. F. Leadlay.** 1990. An unusually large multifunctional polypeptide in the erythromycin-producing polyketide synthase of *Saccharopolyspora erythraea*. Nature (London) **348**:176–178.
12. **Dardel, F., L. C. Packman, and R. N. Perham.** 1990. Expression in *Escherichia coli* of a sub-gene encoding the lipoyl domain of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. FEBS Lett. **264**:215–217.
13. **Delo, J., M. L. Ernst-Fonberg, and K. Bloch.** 1971. Fatty acid synthase from *Euglena gracilis*. Arch. Biochem. Biophys. **143**: 384–391.
14. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
15. **Dunn, J. J., and F. W. Studier.** 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. J. Mol. Biol. **166**:477–535.
16. **Fenn, J. B., M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse.** 1989. Electrospray ionization for mass spectrometry of large biomolecules. Science **246**:64–71.
17. **Flatman, S., and N. M. Packter.** 1983. Partial purification of fatty acid synthase from *Streptomyces coelicolor*. Biochem. Soc. Trans. **11**:597.
18. **Frischauf, A. M., H. Lehrach, A. Poutska, and N. Murray.** 1983. Lambda replacement vectors carrying polylinker sequences. J. Mol. Biol. **170**:827–842.
18a. **Gibson, T. J.** 1984. Ph.D. thesis. University of Cambridge, Cambridge.
19. **Gold, L.** 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. **57**:199–233.
20. **Guerra, D. J., and J. Browse.** 1989. The recombinant spinach acyl-acyl carrier protein-I expressed in *Escherichia coli* is the 18:1Δ11(cis) thioester. Arch. Biochem. Biophys. **271**:246–253.
21. **Hale, R. S., K. N. Jordan, and P. F. Leadlay.** 1987. A small, discrete acyl carrier protein is involved in *de novo* fatty acid biosynthesis in *Streptomyces erythraeus*. FEBS Lett. **224**:133–136.
22. **Henry, K. D., E. R. Williams, B. H. Wang, F. W. McLafferty, J. Shabanowitz, and D. F. Hunt.** 1989. Fourier-transform mass spectrometry of large molecules by electrospray ionization. Proc. Natl. Acad. Sci. USA **86**:9075–9078.
23. **Hirel, P.-H., J.-M. Schmitter, P. Dessen, G. Fayat, and S. Blanquet.** 1989. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side chain length of the second amino acid. Proc. Natl. Acad. Sci. USA **86**:8247–8251.
24. **Hopwood, D. A., M. J. Bibb, K. F. Chater, G. R. Janssen, F. Malpartida, and C. P. Smith.** 1986. Regulation of gene expression in antibiotic-producing *Streptomyces*, p. 251–274. *In* I. R. Booth and C. F. Higgins (ed.), Regulation of gene expression. 25 years on. The 39th Symposium of the Society for General Microbiology. Cambridge University Press, Cambridge.
25. **Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, and H. S. Schrempf.** 1984. Genetic manipulation of *Streptomyces*, a laboratory manual. John Innes Institute, Norwich, United Kingdom.
26. **Hopwood, D. A., and D. H. Sherman.** 1990. Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu. Rev. Genet. **24**:37–66.
27. **Kaneda, T.** 1973. Biosynthesis of branched long-chain fatty acids from the related short chain α-ketoacid substrates by a cell-free system of *Bacillus subtilis*. Can. J. Microbiol. **19**:87–96.
28. **Kaneda, T., J. C. Butte, B. Taubman, and J. W. Corcoran.** 1962. Actinomycete antibiotics. III. The biogenesis of erythronolide, the C_{21} branched chain lactone in erythromycin. J. Biol. Chem. **237**:322–327.
29. **Kauppinen, S., M. Siggard-Andersen, and P. von Wettstein-Knowles.** 1988. β-Ketoacyl-ACP synthase I of *Escherichia coli*: nucleotide sequence of the *fabB* gene and identification of the cerulenin binding residue. Carlsberg Res. Commun. **53**:357–370.
30. **Kawaguchi, A., and S. Okuda.** 1977. Fatty acid synthetase from *Brevibacterium ammoniagenes*: formation of monounsaturated fatty acids by a multienzyme complex. Proc. Natl. Acad. Sci. USA **74**:3180–3183.
31. **Kieser, T.** 1984. Factors affecting the isolation of cccDNA from *S. lividans* and *E. coli*. Plasmid **12**:19–36.
32. **Labeda, D. P.** 1987. Transfer of the type strain of *Streptomyces erythraeus* (Waksman 1923) Waksman and Henrici 1948 to the genus *Saccharopolyspora* Lacey and Goodfellow 1975 as *Saccharopolyspora erythraea* sp. nov., and designation of a neotype strain for *Streptomyces erythraeus*. Int. J. Syst. Bacteriol. **37**:19–22.
33. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
34. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
35. **Matsudaira, P.** 1987. Sequence from picomole quantities of proteins electroblotted onto polyvinylidene difluoride membranes. J. Biol. Chem. **262**:10035–10038.
36. **Mohamed, A. H., S. S. Chirala, N. H. Mody, W. Y. Huang, and S. J. Wakil.** 1988. Primary structure of the multifunctional a subunit protein of fatty acid synthase derived from FAS 2 gene sequence. J. Biol. Chem. **263**:12315–12325.
37. **Pazirandeh, M., S. S. Chirala, W.-Y. Huang, and S. J. Wakil.** 1989. Characterization of recombinant thioesterase and acyl carrier protein domains of chicken fatty acid synthase expressed in *Escherichia coli*. J. Biol. Chem. **264**:18195–18201.
38. **Rackwitz, H. R., G. Zehetner, A. M. Frischauf, and H. Lehrach.** 1984. Rapid restriction mapping of DNA cloned in lambda phage vectors. Gene **30**:195–200.
38a. **Revill, W. P.** Unpublished data.
39. **Rossi, A., and J. W. Corcoran.** 1973. Identification of a multienzyme complex synthesizing fatty acids in the actinomycete *Streptomyces erythraeus*. Biochem. Biophys. Res. Commun. **50**:597–602.
40. **Russel, M., and P. Model.** 1984. Replacement of the *fip* gene of *Escherichia coli* by an inactive gene cloned on a plasmid. J. Bacteriol. **159**:1034–1039.
41. **Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Ehrlich.** 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.
42. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
43. **Schweizer, E., K. Werkmeister, and J. Jain.** 1978. Fatty acid biosynthesis in yeast. Mol. Cell. Biochem. **21**:95–107.
44. **Schweizer, M., K. Takabayashi, T. Laux, K. F. Beck, and R. Schreglmann.** 1988. Rat fatty acid synthase: localisation of the constituent domains and two functional poyadenylation/termination signals in the cDNA. Nucleic Acids Res. **17**:567–586.
45. **Sherman, D. H., F. Malpartida, M. J. Bibb, H. M. Kieser, M. J. Bibb, and D. A. Hopwood.** 1989. Structure and deduced function of the granaticin-producing polyketide synthase gene cluster of *Streptomyces violaceoruber* TU22. EMBO J. **8**:2717–2725.
46. **Spaink, H. P., J. Weinman, M. A. Djordjevic, C. A. Wijffelman, R. J. H. Okker, and B. J. J. Lugtenburg.** 1989. Genetic analysis and cellular localization of the *Rhizobium* host specificity-determining NodE protein. EMBO J. **8**:2811–2818.
47. **Staden, R.** 1982. Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. **10**:4731–4751.
48. **Stoops, J. K., E. S. Awad, M. J. Arslanian, S. Gunsberg, S. J. Wakil, and R. M. Oliver.** 1978. Studies on the yeast fatty acid synthetase subunit composition and structural organization of a large multifunctional enzyme complex. J. Biol. Chem. **253**: 4464–4475.
48a. **Tabor, S.** Unpublished data.
49. **Tabor, S., H. E. Huber, and C. C. Richardson.** 1987. *Escherichia coli* thioredoxin confers processivity on the DNA poly-

merase activity of the gene 5 protein of bacteriophage T7. J. Biol. Chem. **262**:16212–16223.
50. **Tabor, S., and C. C. Richardson.** 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA **82**:1074–1078.
51. **Vanaman, T. C., S. J. Wakil, and R. L. Hill.** 1968. The complete amino acid sequence of the acyl carrier protein of *Escherichia coli*. J. Biol. Chem. **243**:6420–6431.
52. **Van den Boom, T., and J. E. Cronan, Jr.** 1989. Genetics and regulation of bacterial lipid metabolism. Annu. Rev. Microbiol. **43**:317–343.
53. **Witkowska, H. E., B. N. Green, and S. Smith.** 1990. The carboxyl-terminal region of thioesterase II participates in the interaction with fatty acid synthase. J. Biol. Chem. **265**:5662–5665.
54. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
55. **Yuan, Z., W. Liu, and G. G. Hammes.** 1988. Molecular cloning and sequencing of DNA complementary to chicken liver fatty acid synthase mRNA. Proc. Natl. Acad. Sci. USA **85**:6328–6331.
56. **Zhang, H., R. Scholl, J. Browse, and C. Somerville.** 1988. Double stranded DNA sequencing as a choice for DNA sequencing. Nucleic Acids. Res. **16**:1220.