The Gene Encoding *Escherichia coli* Acyl Carrier Protein Lies within a Cluster of Fatty Acid Biosynthetic Genes*

(Received for publication, December 23, 1991)

Merriann Rawlings‡ and John E. Cronan, Jr.
From the Departments of Microbiology and Biochemistry, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

The gene encoding *Escherichia coli* acyl carrier protein (ACP) has been isolated and sequenced. The ACP gene (called acpP) was located on the genetic map between fabF and fabD which encode two fatty acid biosynthetic enzymes, 3-ketoacyl-ACP synthase II and malonyl CoA-ACP transacylase, respectively. An open reading frame between acpP and fabD encodes a 26.5-kDa protein that has significant sequence identity (>40%) with two acetoacetyl-CoA reductases and thus is believed to encode a 3-ketoacyl-ACP reductase. This gene (called fabG) is cotranscribed with acpP. Thus, the gene encoding ACP, the key carrier protein of fatty acid synthesis, is located within a cluster of fatty acid biosynthetic genes.

Acyl carrier protein (ACP)\(^1\) plays a key role in lipid biosynthesis in bacteria (1) and plants (2). ACP carries the nascent fatty acid chain esterified to the thiol group of the 4'-phosphopantetheine prosthetic group and delivers the finished acyl chain to the acyltransferases catalyzing complex lipid synthesis (phospholipids and lipid A) (1, 2). ACP has also been reported as an acyl donor in protein acylation (3). *Escherichia coli* ACP and its acyl forms thus interact with at least 12 different *E. coli* enzymes. The ACPs of other bacteria and plants are very similar to that of *E. coli*; all are small (<90 residues) acidic proteins modified with 4-phosphopantetheine with strong similarities of the sequences neighboring the modification site (1, 2). Indeed, several of these proteins are known to function with various of the ACP-dependent enzymes of *E. coli* in vitro (2) and in vivo (2, 4), suggesting similar solution structures. The large polynucleotides that catalyze fatty acid synthesis in mammals (5) and fungi (6) contain 4'-phosphopantetheine-modified domains with strong sequence similarity to *E. coli* ACP. ACP-like proteins also function as acyl group carriers in the biosynthesis of polyketide (7) and polyamino acid antibiotics (8).

ACP has also been shown to function in three unexpected areas of metabolism: (i) as a cofactor in the synthesis of the membrane-derived oligosaccharides found in the periplasm of *E. coli* (9); (ii) as an essential component in the induction of nitrogen-fixing nodules by Rhizobia (where ACP appears involved in the synthesis of acylated oligosaccharides (10); and (iii) most recently as a subunit of mitochondrial NADH-ubiquinone oxidoreductases (11, 12).

*E. coli* ACP is the paradigm of this class of proteins. *E. coli* ACP was the first such protein isolated (13), the first in which the primary sequence was determined (14), and is the only ACP of known solution structure (15, 16). Despite the continuing interest in *E. coli* ACP, the gene encoding this protein had not been isolated, and no mutants are available. We report the isolation of the ACP-encoding gene and its location within a cluster of genes encoding known enzymes of fatty acid synthesis.

**EXPERIMENTAL PROCEDURES**

All bacterial strains used in this study were derivatives of *E. coli* K-12. Strains JM103 (17) and M13 mp18 (18) were described elsewhere. Strain DB4640 (F' argE lac-pro rif, sal) was used as a source of chromosomal DNA. Strains L48 (fabD88), DM57 (fabB20 zfa:Tn10), and DM83 (fabF3 fabB20) used to map the kanamycin resistance (Kan\(^{R}\)) determinant were described previously (19). The growth media and genetic methods were as described (19).

Plasmid pMR23 was constructed by ligation of a 0.9-kbp PstI-ScaI fragment from Tn9 into pACYC177 (20) digested with the same two enzymes. Plasmid pMR24 was constructed by ligating the 2.6-kbp *EcoRI-PstI* chromosomal fragment (Fig. 1) from the M13 mp18 clone containing the *acpP* region to pMR23 digested with the same enzymes. Plasmid pMR39 was constructed by inserting the Kan\(^{R}\) gene excised from plasmid pUC4K (Pharmacia LKB Biotechnology Inc.) with HincII into the NruI site of pMR35 and was used to introduce the Kan\(^{R}\) determinant into the chromosome by homologous recombination (19). Plasmid pMR48 is derived from the 1.1-kbp *PstI-PvuII* fragment of pMR23 into pTZ19R (21) digested with *EcoRI* and *HindIII*. Plasmid pMR36 is a derivative of pMR24 having the 2.1-kbp *KpnI* fragment in the opposite orientation. Plasmid pMR33 was constructed by inserting the 1.5-kbp *EcoRI-PvuII* fragment of pMR24 into pTZ19R (21) digested with *EcoRI* and *HindIII*. Plasmid pMR62 was constructed by digestion of pMR24 with *EcoRI* and *SalI*, filling of the recessed ends by DNA polymerase I, and followed by religation.

The *acpP* gene was isolated from a library of 2.3-kbp *SalI-BglII* fragments of strain DB4640 genomic DNA ligated into M13 mp18 RF digested with *SalI* and *BamHI*. Strain JM103 was transformed with the recombinant DNA and the resultant plaques transferred in *situ* to nitrocellulose membranes such that single-stranded DNA was selectively retained (22). The plaques were screened (17) with the [α-\(32P\)]ATP-labeled synthetic *ACP* gene (18) that encodes the entire protein sequence and has 88% DNA sequence identity with the *acpP* gene.

**RESULTS AND DISCUSSION**

Our previous attempts to isolate the ACP gene were unsuccessful despite application of several different cloning and detection strategies. It, therefore, seemed possible that DNA segments encoding ACP were somehow toxic to *E. coli*. To investigate this possibility, we assembled a synthetic gene encoding ACP (18) and, indeed, found that high level production of ACP was lethal to *E. coli* (39). In light of this finding,

---

*This project was supported by National Institutes of Health Grant AI 15650. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C.

‡ Present address: Dept. of Microbiology, University of Iowa, Iowa City, Iowa 52242.

§ The abbreviations used are: ACP, acyl carrier protein; kbp, kilobase pair; ORF, open reading frame;
we sought the ACP gene using the synthetic sequence as a hybridization probe and maintained cloned DNA segments in a low copy number vector to limit ACP expression. The synthetic ACP gene proved a stringent hybridization probe in Southern blot analysis of E. coli chromosomal DNA fragments (data not shown). A size-selected mini-library of genomic DNA was constructed in a phage M13 vector, and recombinant phage plaques were screened with the 32P-labeled synthetic gene. Screening was done under conditions allowing hybridization only to single-stranded DNA bound to the nitrocellulose filters, thus avoiding the background due to homologous sequences present in the chromosomal DNA (22).

Consistent with the toxicity of the synthetic gene, the high (albeit variable) copy number of M13 clones carrying the natural gene (called acp) gave spontaneously deleted variants at very high frequency, thus necessitating transfer of acp to a low copy number vector to give plasmid pMR24. Even in such a plasmid, the presence of the acp gene resulted in a decreased cellular growth rate.

The nucleotide sequence of the acp gene region (Fig. 1) showed an open reading frame (ORF) that agreed with the ACP amino acid sequence. Note that the two published ACP amino acid sequences conflict at two positions. Vanaman et al. (14) reported residues 24 and 43 as Asp and Val, respectively. We find residue 24 to be Met, whereas Jackowski and Rock (23) reported residues 24 and 43 as Asn and Ile, respectively. We find residue 24 to be Asn and residue 43 to be Val. Our results are consistent with assignments from nuclear magnetic resonance analysis (15).

The observed post-translational removal of the N-terminal Met is consistent with the known specificity of the aminopeptidase (24). The codon preference (25) of 1.34 (versus 0.48 for the randomized sequence) is consistent with this known high expression of acp (about 5 x 10^6 molecules/cell) (1). Despite the high level of expression no sequences matching the promoter and ribosome binding sites consensus sequences are obvious.

The acp gene was localized on the genetic map of E. coli by inserting a kanamycin resistance (KanR) determinant into pMR35 within a DNA sequence downstream of acp. The KanR sequence was then transferred into the E. coli chromosome by homologous recombination (26) to give strain MR52. Conjugational mapping located the acp gene between min 13 and 30, whereas P1-mediated transduction localized the gene to min 24 (98% linkage of the KanR insertion with the zce-726::Tn10 insertion (27)). This location is very close to those we previously assigned to genes encoding two other fatty acid biosynthetic proteins, fabD and fabF, which encode malonyl-CoA ACP transacylase and 3-ketoacyl-ACP synthase II, respectively (1, 19). Therefore, it seemed probable that acp was linked to these genes. Indeed, when phage P1 grown on the KanR insertion strain was used to transduce a fabD strain to KanR, 98% of the transductants were fab+. The KanR insertion of strain MR52 could not be mapped in relation to fabF because the insertion of the KanR element resulted in a strain having a fabF phenotype.

FabF mutants have no growth phenotype unless the strain also has a lesion in the fabB gene that encodes 3-ketoacyl-ACP synthase I (1, 19). Mutants with a temperature-sensitive lesion in fabB fail to grow at 42 °C on the usual media but grow well if the medium is supplemented with oleate (or other appropriate unsaturated fatty acids). FabFB double mutants fail to grow at 42 °C even when supplemented with oleate (due to defective synthesis of saturated fatty acids). These mutants have no growth phenotype unless the strain is grown in the presence of oleate (due to defective synthesis of saturated fatty acids). FabFB double mutants fail to grow at 42 °C even when supplemented with oleate (due to defective synthesis of saturated fatty acids). We found that P1 cotransduction of a fabB− lesion into strain MR52 gave a fabB−/fabF+ phenotype. Strain MR52 also showed other aspects of the fabF phenotype (19): (i) an increased level of palmitoleic acid and a decreased level of cis-vaccenic acid; (ii) defective thermal regulation of fatty acid composition; (iii) lack of the 3-ketoacyl-ACP synthase II-ACP mixed disulfide in cell extracts. Thus, the KanR insertion of strain MR52 is either in fabF or is polar on fabF expression. We favor the former explanation since strain MR52 was unable to donate a functional fabF gene to a fabF+ strain via P1 transduction.

The segment of DNA containing acp was also located on the physical map of E. coli. The acp DNA segment hybridized to phases 235 and 236 of the ordered miniset bank of Kohara and co-workers (28). Comparison of our restriction map of the physical map (28) placed the acp gene at 1170 kbp, a site fully consistent with the genetic map location.

Fig. 1. Physical map and sequence of the fabG−/acp region. Top panel, a restriction map of the 2.6-kbp Sal1-BglII genomic fragment is given on the top line. The outermost PstI, KpnI, and EcoRI sites are vector sites. The strategy used for DNA sequencing is shown underneath the map. Arrows indicate the extent and direction of sequencing by base pairs. Bottom panel, nucleotide assignments deduced amino acid sequence of the fabG− and acp genes. A putative transcriptional terminator is indicated by arrows. The first 63 nucleotides encode the carboxyl terminus of malonyl transacylase (fabD)
Fig. 2. Maxicell analysis of plasmid-encoded proteins. Lane 1, 14C-labeled protein standards; lane 2, vector pTZ18R (21); lane 3, pMR62, a derivative of vector pACYC177 (20); lane 4, pMR24, which carries the entire 2.6-kbp chromosomal insert in pMR62 (Fig. 1); lane 5, pMR48, which carries the left-hand 1.1-kbp PstI-PvuII fragment (Fig. 1); and lane 6, pMR33 which contains the right-hand EcoRI-PvuII fragment (Fig. 1). pMR48 and pMR33 are derived from vector pTZ19R (20). The 20- and 26.5-kDa proteins were the products of the acpP and fabG genes, respectively, kan, aminoglycoside phosphotransferase; pBlu, the precursor of β-lactamase; Blu, β-lactamase.

Fig. 3. Comparison of the amino acid sequence of the fabG gene product with similar sequences. Shown are the complete sequences for E. coli fabG, R. meliloti nodG (38), aceatoacetyl-CoA reductases from A. eutrophus (AEUT (34)) and Z. ramigera (ZRAM (35)), S. violaceoruber ORFS (7) and S. coelicolor actIII (7). Residues identical in all six sequences are boxed. Also shown is a portion of the ketoreductase domains from the rat (RAT) (5) and chicken (CHIC (35)) multifunctional fatty acid synthases. Amino acids which comprise a putative NADPH-dimucleotide fold are indicated by asterisks.

Fig. 4. Northern blot analysis of acpP and fabG transcripts. Whole cell lysates were electrophoresed on a 1% agarose/formaldehyde gel, blotted, and probed (38). Lanes of strain F- M15 (lane 1) or F-M18A carrying either pMR24 (lane 2) or pMR36 (lane 5) were probed with the synthetic ACP gene. Plasmid pMR24 contains the intact acpP-fabG region whereas in pMR36 the acpP gene is inverted. Lanes 3 and 4 are the same RNA samples of lanes 1 and 2, respectively, but the probe was the 0.4-kbp PmlI-Stul fragment (Fig. 1) specific for the fabG gene. Plasmid DNA (which is isolated with the RNA in the procedure used (38)), the cross-hybridizing rRNAs, and the size (in kb) of the two major transcripts are indicated at the left margin.

much larger protein) due to lower sodium dodecyl sulfate binding than the marker proteins (30). Deletion of the ACP sequence resulted in loss of the 20-kDa protein. Maxicell analysis of various subclones (Fig. 2) showed that the 43-kDa protein was encoded by a DNA segment located downstream of acpP, and hence, this protein may well represent a slightly truncated fabG gene product, 3-ketoacyl-ACP synthase II, a protein of 44 kDa (31). (The Kanβ insertion of strain MR52 would interrupt the synthesis of this protein.) The 26.5-kDa protein is encoded upstream of acpP but is not the fabD gene product since malonyl CoA-ACP transacylase has a molecular mass of 36.5 kDa (32). We have sequenced the DNA segment upstream of acpP and find an ORF that encodes a protein of 244 residues having a calculated molecular weight of 25,549 in good agreement with the maxicell results (Fig. 2). Comparison of the derived amino acid sequence of this ORF with those of GenBank showed a number of proteins with strong similarity to the ORF. The most definitive of these similarities was with two enzymes involved in poly-3-hydroxybutyrate synthesis in bacteria (33, 34). These are aceatoacetyl-CoA reductases which reduce aceatoacetyl-CoA (formed by condensation of two acetyl-CoA units) to the 3-hydroxybutyl-CoA used in polymer synthesis (33, 34). The ORF upstream of acpP showed 43 and 41% amino acid identity with the NADPH-specific aceatoacetyl-CoA reductases of Zoogloea ramigera (33) and Alcaligenes eutrophus (34), respectively (Fig. 3). We also found significant similarities to a segment of the large polyfunctional fatty acid synthase proteins of rat (5) and chicken (35). Strong similarities (40–53% amino acid identities) were also found to genes involved in polyketal synthesis in various Streptomyces (7) and to the nodG protein of Rhizobium meliloti (36) which may be involved in synthesis of acylated polysaccharides (10, 37). These relationships (Fig. 3) together with the presence of a plausible NADPH binding site in the upstream ORF and cotranscription of the ORF with acpP (see below) lead us to believe this ORF encodes a 3-ketoacyl-ACP reductase of fatty acid biosynthesis (13), a gene we term fabG.

The close juxtaposition of these coding sequences suggested...
possible cotranscription of these genes. The maxicell results suggested promoters were present just upstream of both the acpP and fabG coding sequences, and this was confirmed by Northern blot analyses. Two chromosomal transcripts of about 0.3 and 1.1 kb were detected using the synthetic ACP gene as a probe (Fig. 4). A strain carrying pMR24 showed increased levels of both mRNA species. The 0.3-kb transcript has the capacity to encode ACP whereas the 1.1-kb transcript could encode both acpP and fabG. Only the 1.1-kb transcript was observed when a probe containing fabG sequences alone was used (Fig. 4). The relationship between the 0.3- and 1.1-kb transcripts was examined by inverting the acpP coding sequence within pMR24. When probed with the synthetic ACP gene, the 0.3-kb mRNA was the dominant transcript (Fig. 4). Therefore, the 0.3-kb mRNA seems a primary transcript and not a degradation product of the 1.1-kb mRNA. Thus, acpP is transcribed from two promoters. A strong promoter is located just upstream of the coding sequence, and a second is located upstream of the fabG sequence.

CONCLUSIONS

ACP is encoded by the acpP gene. Genes encoding other fatty acid biosynthetic genes lie both upstream and downstream of acpP. One downstream gene is fabF, encoding 3-ketoacyl-ACP synthase II, and several genes are located upstream. We have shown that fabG (which almost certainly encodes a 3-ketoacyl-ACP reductase) lies just upstream of acpP and is cotranscribed with acpP. In other work (41) we find that fabD encoding malonyl CoA:ACP transacylase is transcribed from two promoters. A strong downstream gene is fabF, and one is located upstream of the fabG gene. Indeed the first 63 bp of Fig. 1 encode the last 21 amino acids of fabD (41). Upstream of fabD lies another ORF (called fabH) that encodes 3-ketoacyl-ACP synthase III (40). Thus, the genes encoding several enzymes of fatty acid synthesis are clustered around the acpP gene and the toxicity of increased expression of acpP probably explains prior difficulties in cloning these genes. Our current map of this region (clockwise on the physical/genetic map) is fabH-fabD-fabG-acpP-fabF. The gene order has no obvious relationship to the order or protein domains in the multifunctional fatty acid synthases of mammals or fungi.

REFERENCES

1. Cronan, J. E., Jr., and Rock, C. O. (1987) in Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology (Neidhardt, F. C., Ingraham, J. L., Low, K. B., Magasanik, B., Schaechter, M., and Umbarger, H., eds) Vol. 1, pp. 474–497, American Society for Microbiology, Washington, D. C.
2. Ohlrogge, J. B., Browse, J., and Somerville, C. R. (1983) Biochim. Biophys. Acta 1022, 1-26
3. Issartel, J-P., Koronakis, V., and Hughes, C. (1991) Nature 351, 759-761
4. Revill, W. P., and Leadley, P. F. (1991) J. Bacteriol. 173, 4379–4385
5. Amy, C. M., Witkowski, A., Naggert, J., William, B., Randhawa, Z., and Smith, S. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 3114–3118
6. Pazirendih, M., Chirala, S. S., Huang, W-Y., and Wakil, S. J. (1989) J. Biol. Chem. 264, 18195–18201
7. Hopwood, D. A., and Sherman, D. H. (1990) Annu. Rev. Genet. 24, 37–66
8. Lipman, F. (1980) Adv. Microbiol. Physiol. 21, 227–266
9. Theriot, H., and Kennedy, E. P. (1987) Proc. Natl. Acad. Sci. U.S.A. 84, 8225–8228
10. Spanik, H. P., Shelly, D. M., van Brussel, A. A. N., Glushka, J., York, W. S., Tak, T., Gerger, O., Kennedy, E. P., Reinhold, V. N., and Lugtenberg, B. J. J. (1991) Nature 354, 125–130
11. Szacka, V., Zensen, R., Rohlen, D., Jahneke, V., and Weiss, H. (1991) Eur. J. Biochem. 200, 463–469
12. Runswick, M. J., Fearney, I. M., Shelly, J. M., and Walker, J. E. (1991) FEBS Lett. 286, 121–124
13. Prescott, D. J., and Vagelos, P. R. (1972) Adv. Enzymol. Relat. Areas Mol. Biol. 36, 288–311
14. Vannamee, T. C., Elledge, S. J., and Hill, R. L. (1968) J. Biol. Chem. 243, 6420–6436
15. Holak, T. A., and Prestegard, J. H. (1989) Biochemistry 28, 5766–5774
16. Kim, Y., and Prestegard, J. H. (1986) Biochemistry 25, 8792–8796
17. Maniatis, T., Fritsche, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
18. Cronan, J. E., Jr., Narasimhan, M. L., and Rawlings, M. (1988) Gene (Amst.) 70, 161–169
19. Ulrich, A. K., de Mendoza, D., Warner, J. L., and Cronan, J. E., Jr. (1983) J. Bacteriol. 154, 221–230
20. Chang, A. C. Y., and Cohen, S. N. (1978) J. Bacteriol. 134, 141–156
21. Mead, D. A., Szczesa-Skorupa, E., and Kemper, B. (1986) Protein Eng. 1, 67–74
22. Wei, G. G., and Surzynski, S. (1986) Gene (Amst.) 48, 251–256
23. Jackowski, S., and Rock, C. O. (1987) J. Bacteriol. 169, 1469–1472
24. Hirel, P.-H., Schmitter, J. M., Dessen, P., Fayat, K., and Banquet, S. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 8247–8251
25. Grisbok, M., Devereaux, J., and Burgess, R. R. (1984) Nucleic Acids Res. 12, 537–549
26. Williams, S. C., Elledge, S. J., Krueger, J. H., and Walker, G. C. (1985) J. Bacteriol. 161, 1219–1221
27. Singer, M., Baker, T. A., Schnitter, G., Deischel, S. M., Goel, M., Dover, W., Janks, K. J., Grossman, A. D., Erickson, J. W., and Gross, C. A. (1989) Microbiol. Rev. 53, 1–24
28. Kohara, Y., Akiyama, K., and Isono, K. (1987) Cell 50, 495–506
29. Deveauux, L. C., Clevenston, D. S., Bradbeer, C., and Kadner, R. L. (1986) J. Bacteriol. 167, 920–927
30. Rock, C. O., and Cronan, J. E., Jr. (1979) J. Biol. Chem. 254, 9779–9785
31. Garwin, J. L., Klages, A. L., and Cronan, J. E., Jr. (1980) J. Biol. Chem. 255, 11949–11956
32. Ruch, F. E., and Vagelos, P. R. (1973) J. Biol. Chem. 248, 8086–8094
33. Peoples, O. P., and Sinskey, A. J. (1989) Mol. Microbiol. 3, 349–357
34. Peoples, O. P., and Sinskey, A. J. (1989) J. Biol. Chem. 264, 15293–15297
35. Chirala, S. S., Kuziora, M. A., Spector, D. M., and Wakil, S. T. (1987) J. Biol. Chem. 262, 4291–4294
36. Debellé, F., and Sharma, S. B. (1986) Nucleic Acids Res. 14, 7453–7472
37. Lerouge, P., Roche, P., Faucher, C., Marillet, F., Truchet, G., Prone, J. C., and Denarie, J. (1990) Nature 344, 781–784
38. Kornblum, J. S., Projan, S. J., Maghazek, S. L., and Novik, R. P. (1988) Gene (Amst.) 63, 75–85
39. Rawlings, M., and Cronan, J. E., Jr. (1988) FASEB J. 2, A1559
40. Tsay, T.-T., Oh, W., Larson, T. J., Jackowski, S., and Rock, C. O. (1992) J. Biol. Chem. 267, 6807–6814
41. Magnuson, K., Oh, W., Larson, T. J., and Cronan, J. E., Jr. (1992) FEMS Lett., in press