Acyl carrier protein of Azospirillum brasilense: properties of the purified protein and sequencing of the corresponding gene, acpP

Mrinal Kumar Maiti† and Sudhamoy Ghosh

Acyl carrier protein (ACP) plays a crucial role in bacterial fatty acid synthesis. Cloning genes encoding ACPS from Gram-negative bacteria in Escherichia coli is difficult due to adverse effects of the cloned gene on host cell viability, and we were unsuccessful in cloning the full length ACP gene (acpP) from Azospirillum brasilense using conventional methods. Therefore, ACP from A. brasilense was purified to homogeneity and a part of the acpP gene was cloned using the polymerase chain reaction (PCR) technique with two primers, one designed from the N-terminal amino acid sequence of the purified ACP and the other from the highly conserved amino acid sequence of bacterial ACPS. The nucleotide sequence of the gene was obtained by cloning and sequencing inverse PCR products containing the acpP region generated by two oppositely oriented internal primers designed from the partial acpP gene sequence using restriction-enzyme-digested, self-circularized chromosomal DNA fragments as templates. Characterization of the purified ACP and analysis of the derived amino acid sequence of the acpP gene of A. brasilense revealed that: (a) the mature ACP, composed of 78 amino acids, is a highly expressed protein (about 2.0-3.0 x 10^4 molecules per cell), (b) compared to E. coli ACP, it has a more compact structure and contains significantly more hydrophobic amino acid residues and (c) the potential mRNA sequence of the ACP gene has some structural features typical of a stable mRNA.

Keywords: Azospirillum brasilense, acyl carrier protein, reverse genetics, inverse PCR cloning, acpP

INTRODUCTION

Acyl carrier protein (ACP) plays a pivotal role as a cofactor in de novo fatty acid synthesis in bacteria and plant plastids. The Escherichia coli ACP (Magnuson et al., 1993) is one of the most abundant proteins in the cells. In spite of the fact that ACP is a small protein (about 9 kDa) and the amino acid sequence of the ACP from E. coli was determined 28 years ago (Vanaman et al., 1968), cloning of the gene encoding ACP from Gram-negative bacteria has proved to be very difficult. Only recently was the acpP gene from E. coli cloned (Rawlings & Cronan, 1992). ACPs have been purified from a number of bacteria including two Gram-negative aerobic diazotrophs, Rhodobacter sphaeroides (Cooper et al., 1987) and Rhizobium meliloti (Platt et al., 1990), and two cyanobacteria (Froehlich et al., 1990); amino acid sequences of these proteins were also determined. However, the only constitutively expressed ACP gene cloned from Gram-negative bacteria is that from E. coli (Rawlings & Cronan, 1992). We wanted to clone and sequence the gene encoding the ACP of Azospirillum brasilense to analyse the gene structure, which was expected to throw new light on ACP gene expression in a Gram-negative aerobic bacterium.

In this report we describe an approach based on reverse genetics that yielded the full-length sequence of acpP, the gene encoding ACP of A. brasilense, involving purification of ACP, determination of the N-terminal amino acid sequence of the protein to construct authentic, ACP-specific degenerate primers, partial cloning of the gene employing the polymerase chain reaction (PCR), and

†Present address: Department of Chemistry, Miami University, Oxford, OH 45056, USA.

Abbreviation: ACP, acyl carrier protein.

The EMBL accession number for the nucleotide sequence reported in this paper is X82399.
METHODS

Bacterial strains and growth conditions. Growth and maintenance of *A. brasilense* RG, the strain used in this study, were described previously (Mukherjee & Ghosh, 1987). For purification of ACP from *A. brasilense* RG, the bacteria were grown at 32 °C in an 8 l fermentor with vigorous aeration in fructose-salt minimal medium nearly to stationary phase (OD_{600} 4-1). The cells were collected by centrifugation after washing with 0.02 M Tris/HCl (pH 8-0) and the packed cells (77 g) stored frozen at −20 °C until used. *E. coli* cells were grown at 37 °C in LB liquid medium (Sambrook et al., 1989) in a rotary shaker, or on LB agar plates containing appropriate antibiotics when necessary at 37 °C for 30 min and the mixture homogenized by sonication (Braunsonic Sonicator, model 1510) to prepare the crude extracts. An equal volume of 2-propanol was added to the sonicated crude extracts and the 2-propanol supernatant fraction (150 ml) was purified by DEAE-Sepharose (Pharmacia) ion-exchange chromatography using increasing concentrations of NaCl in 20 mM Tris/HCl buffer (pH 7-5) containing 14 mM 2-mercaptoethanol, first collecting the active fraction by batchwise elution and then subjecting it to gradient elution. The combined peak fractions eluting at about 0.26 M NaCl by gradient chromatography were dialysed against 1 mM KH_{2}PO_{4}/KH_{2}PO_{4} (pH 7-0), 14 mM 2-mercaptoethanol, freeze-dried, redissolved in dialysis buffer and then purified by a preparative native-PAGE treatment. Only the fastest running protein band in the native gel was associated with ACP activity. The unstained gel band containing ACP activity (after identification by parallel staining) was extracted with 0.6 M NaCl in 20 mM Tris/HCl (pH 7-5), the extract dialysed (preparative native-PAGE), freeze-dried and stored in a refrigerator. This freeze-dried preparation served as the source of pure ACP of *A. brasilense* (Fig. 1).

Determination of N-terminal amino acid sequence. The pure ACP preparation after SDS-PAGE was blot-transferred to a PVDF membrane following the method of Matsudaira (1987). N-terminal amino acid sequencing was performed with an automated microsequencer (Model 477A, Applied Biosystems).

Recombinant DNA techniques. Procedures for DNA manipulations, such as plasmid DNA purification, *in situ* colony hybridization, Southern hybridization, DNA band elution from low-melting-point agarose, cloning, subcloning, construction of a genomic library and recombinant plasmid DNA transformation into *E. coli*, were as described by Sambrook et al. (1989) with minor modifications. For the construction of a cosmid genomic library, pLAFR3 was used as vector and *E. coli* S17.1 was used as host (Chattopadhyay et al., 1994). Total chromosomal DNA from *A. brasilense* was isolated following the technique of Marmur (1961); the isolated DNA was purified by RNase treatment and a phenol/chloroform extraction procedure. The method yielded high-molecular-mass DNA which was easily digested by all restriction enzymes tested.

### Table 1. Purification of ACP from *A. brasilense*

<table>
<thead>
<tr>
<th>Purification step</th>
<th>Total volume (ml)</th>
<th>Protein concentration (mg ml⁻¹)</th>
<th>Total ACP (μg)</th>
<th>Purification (-fold)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Crude extracts</td>
<td>100</td>
<td>37.70</td>
<td>2070*</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>2. 2-Propanol supernatant</td>
<td>200</td>
<td>2.89</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3. DEAE-Sepharose-I (batch elution)</td>
<td>12</td>
<td>0.56</td>
<td>642</td>
<td>1.73</td>
<td>31</td>
</tr>
<tr>
<td>4. DEAE-Sepharose-II (gradient elution)</td>
<td>9</td>
<td>0.39</td>
<td>625</td>
<td>3.19</td>
<td>30</td>
</tr>
<tr>
<td>5. Preparative native PAGE</td>
<td>4-35</td>
<td>0.10†</td>
<td>435</td>
<td>1823</td>
<td>21</td>
</tr>
</tbody>
</table>

*Not determined.

* Determination of ACP in the crude extracts was performed with the heat-treated soluble supernatant fraction.

† In the step 5 fraction, ovalbumin was used as a standard in protein determination instead of BSA.
Design of degenerate primers, amplification and cloning of a part of the ACP gene by PCR and cloning of the PCR product. Based on the available codon usage pattern of *A. brasilense* (Zimmer *et al.*, 1991), two degenerate oligonucleotides were chemically synthesized (Fig. 2a, b). One (named dgn) was deduced from the N-terminal amino acid sequence of the purified ACP of *A. brasilense* and the other (named dgc) was designed from the conserved amino acid sequence of ACPs of a few Gram-negative bacteria. The primer dgn is a 30-mer having 16-fold degeneracy with a BamHI site at the 5' end, and the primer dgc is a 31-mer having 64-fold degeneracy with an EcoRI site at its 5' end. PCR was carried out in a DNA thermal cycler (Perkin-Elmer Cetus) using reaction conditions optimized according to Innis *et al.* (1990). The PCR products obtained were purified and digested with BamHI and EcoRI. The digested products were ligated to the compatible restriction sites of a pUC18 vector and transformed into *E. coli* strain DH5α (BRL). The desired transformants (pMG plasmids) with a 163 bp insert derived from the ACP gene (see Fig. 4) were identified by restriction enzyme analysis followed by sequencing.

Construction of a subgenomic library and screening for ACP gene clones. Total genomic DNA of *A. brasilense* was completely digested with different restriction enzymes: EcoRI, PstI, SaI, NcoI, ApaI, BstHI and XmaI. The digested DNA fragments were subjected to gel electrophoresis, capillary-transferred onto a nylon membrane (Sigma) and UV-fixed. The partial acpP DNA (163 bp EcoRI-BamHI DNA fragment from pMG) was radiolabelled with [α-32P]dATP by nick translation. Southern blot experiments were carried out following the procedures of Singh & Jones (1984) hybridizing with the denatured radiolabelled probe at high stringency (in 4xSET buffer with heparin at 72°C for 18-20 h) followed by washing (1st wash in 2xSSC+0.2% SDS at 55°C for 10 min; 2nd wash in 0.1xSSC+0.2% SDS at 50°C for 10 min; 3rd wash in 0.1x SSC at 45°C for 10 min).

Based on Southern analysis data (Fig. 3a), a SaI-digested subgenomic library was constructed with a pUC18 vector (Yanisch-Perron *et al.*, 1985) in *E. coli* DH5α. The library was screened by in situ colony hybridization with the labelled 163 bp DNA fragment and the positive clones were further confirmed by purifying the recombinant plasmid DNA followed by hybridizing with the labelled probe.

Amplification and cloning of inverse PCR products to identify the full-length coding sequence of ACP. The inverse PCR method followed was as described by Triglia *et al.* (1988). A pair of divergent oligonucleotide primers, IP-21 and IP-22, were designed from the internal sequenced region (102 bp) of the acpP gene (shown in Fig. 4) in such a way that they annealed to the target DNA template 26 bp apart, and their 3'-termini were in opposite directions. *A. brasilense* genomic DNA was completely digested with individual restriction enzymes (ApaI, BstHI, EcoRI, SauI, SmaI and XmaI). After enzyme inactivation (65°C for 5 min), the DNA fragments were precipitated from the digests and washed twice with 70% ethanol, dried under vacuum, resuspended in ligation buffer at a low concentration of DNA [about 500 ng DNA fragments (ml buffer)] and self-circularized by ligase treatment. After a pre-heating step (95°C, 5 min), inverse PCR reactions were performed with the self-circularized DNA samples using the standard reaction mixture containing dNTPs, primers (IP-21 and IP-22), buffer and Vent DNA polymerase (New England Biolabs) in the thermocycler. Three inverse PCR products from *ApaI*, *BstHI* and *XmaI* digests showing unique DNA bands in agarose gel (Fig. 3b) after purification were ligated to the SmaI site of the pUC18 vector and transformed into the *E. coli* host SURE (Stratagene) yielding plasmids pIGM-A, pIGM-B and pIGM-X, respectively. The desired recombinant clones were checked by sequencing of the target DNA region.

Sequencing of DNA. Both strands of the target DNAs of PCR-, subgenomic- and inverse PCR clones were sequenced by the dideoxy chain-termination method (Sanger *et al.*, 1977) using dT-deaza-dGTP to reduce the G/C band compressions arising from the high G+C content of the *A. brasilense* DNA. Sequencing reactions were performed with Sequenase version 2.0 (US Biochemicals) following the manufacturer's instructions except that the labelling and termination reactions were performed at 20°C for 5 min and 45°C for 15 min, respectively.

Computer analysis. DNA and protein sequence analyses were performed using selected programs of the University of Wisconsin Genetic Computer Group and PC/Gene (Intelligenetics) software packages.

RESULTS

Purification of ACP from *A. brasilense*, and determination of its N-terminal amino acid sequence and its expression level in cells

As shown in Table 1, the five-step procedure starting from 38.5 g frozen *A. brasilense* cells yielded a pure preparation of ACP after 1800-fold purification with about 20% yield. Unlike ACP from *E. coli*, *A. brasilense* ACP did not show anomalous migration in SDS-PAGE (Fig. 1b, lanes 1 and 2) indicating its more compact structure. Although the most purified ACP preparation (preparative native-PAGE fraction) showed a single band in native PAGE (Fig. 1a, lane 1), in SDS-PAGE it clearly

![Fig. 1. PAGE analysis of purified ACP from *E. coli* and *A. brasilense*. (a) Native-PAGE (15 %, w/v, acrylamide), performed according to Jacobowski & Rock (1983). Lanes: 1, *A. brasilense* ACP (preparative native-PAGE fraction); 2, *E. coli* ACP and BSA. (b) SDS-PAGE (15% w/v, acrylamide), performed according to Laemmli (1970). Lanes: 1, *A. brasilense* ACP (preparative native-PAGE fraction); 2, *E. coli* ACP; 3, molecular mass markers.](https://example.com/image)
resolved into two closely spaced bands—a faster migrating major band and a slower migrating minor band (Fig. 1b, lane 1). The N-terminal amino acid sequences of the two polypeptides (Fig. 2a) were identical except for the N-terminal residues, which were Ser and Ala for the minor and the major polypeptides, respectively.

In the quantitative assay of ACP using E. coli acyl-ACP synthetase and radiolabelled [1-14C]palmitate, a linear response of [1-14C]palmitoyl-ACP synthesis was obtained in the range 50–200 pmol pure A. brasilense ACP (results not shown). Under our assay conditions, approximately 0.20 pmol [1-14C]palmitoyl-ACP was formed per pmol pure A. brasilense ACP compared to 0.25 pmol [1-14C]palmitoyl-ACP in the case of E. coli ACP within the same range. Since ACP is heat-stable, the concentration of ACP in the crude extracts was determined using heat-treated supernatant of the extracts. The number of ACP molecules per cell of A. brasilense was calculated to be approximately 2.0–3.0 ∗ 10⁴ compared to 5.6 ∗ 10⁴ for E. coli ACP (Magnuson et al., 1993). Any co-precipitation of the soluble ACP with heat-coagulated cellular protein was not tested for but would cause an underestimation of the calculated value. These results show that ACP is a highly expressed protein in A. brasilense.

Cloning and sequencing of an internal part of the acpP gene using PCR

Analysis of three separate PCR-derived pMG plasmid clones (see Methods) showed that they all contained a 163 bp insert having an identical sequence in the 102 bp internal region flanked by 30-mer and 31-mer PCR primer sequences (Figs 2, 4). Since the two primers were degenerate, the nucleotide sequences of the three insert DNAs of the three clones were slightly different (G replaced by C, C by G, and C by T, without altering the coded amino acids) in the region where the primers

Fig. 2. Design of two degenerate oligonucleotide primers (dgN and dgC) for cloning and sequencing of a part of the ACP gene of A. brasilense. (a) Nucleotide sequence of the N-terminal degenerate primer dgN with a BamHl restriction site. (b) Nucleotide sequence of the C-terminal degenerate primer dgC with an EcoRl restriction site.
Acyl carrier protein of *Apospirillzlm* *brasilense*

GCTGGTCAACGCTGTAAMATGTCTTATCGCAGGGGGCTTTTTGTGGCAGGAMGCACGC
60
AGCCGCGTGTTGAGCCGCCCGMTTCGCCCGAGCGCTCTC~GGATTGGMGGT~~ 120
AAATGAGCGATGTCGCCGAGCGTTAAGAAGATCGTTGTTGTGGACCACCTCGGGGTCGAGG 180
MSDVAERVKKIVVVDHLLGVEE
IP-22
AGTCGAAGGTGACGGAGMCGCCTC~CATCGACGATCTCGGCGCCGACAGCCTCGACA 240
SKVTENASFIDDLGADSLDT
IP-21
CGGTGAGCTGTTATGGCTTTCAAGGAAGATCGGTGCTGCCAGATTCGGGACCGACGCCGCCT 300
VELVMFAEEEPFGCEIPDDA

---

CCGAAGATTTTGCAGGTAGAGACGCCATCGACCTCATCAAGGGAGACGCCGACGCCGCT 360
EKILTVDKDAIFDIKANA

---

GATCAGGCCGTGCGCTTGAAGGACCGCGAGCTTTCCAAAGTGGCCTGCCGCTGGTGA
420

TTTCCGAGGAAAAGTTCAAGGAAT
---

**Fig. 4.** Nucleotide and deduced amino acid sequence of *A. brasilense* acpP. The nucleotide residues are numbered on the right. The potential ribosome-binding site is double-underlined. An asterisk denotes the stop codon. Two potential hairpin structures, one around the stop codon and the other in the 3' flanking sequence, are indicated by inverted-arrow-headed dashed overlines. The *Narl* restriction site within the coding sequence is underlined. The T-rich sequence downstream of the putative transcription terminator is indicated by a dashed underline. An A-T rich region between the putative ribosome-binding site and the ATG start codon is shown in bold. Overhead vertical arrows at positions 153 and 256 delimit the core 102 bp region within the 163 bp fragment of acpP cloned in pMG. Overhead arrowed lines indicate the sequences used for constructing IP-21 and IP-22 primers.

Annealed. The deduced 33 amino acid sequence corresponding to the 99 bp core region matched well with the known amino acid sequences in the same region of three bacterial ACPs, *R.* *meliloti*, *R.* *spheroides* and *E.* *coli* (28 out of 33, 29 out of 33 and 23 out of 33 identical amino acid residues, respectively) (alignment not shown).

**Attempts to isolate a full-length ACP gene clone from a subgenomic library of *A. brasilense***

As shown in Fig. 3(a), each of the restriction enzyme digests of *A. brasilense* DNA hybridized with a unique acpP-specific hybridizing band (except the *Narl* digest which has a restriction site within acpP). This suggests that the acpP gene may be present as a single copy in the *A. brasilense* genome. In order to obtain a full-length ACP gene clone, a number of putative acpP-positive clones were selected from the *SalI*-digested subgenomic library (see Methods) for identification of the ACP gene by sequencing. It was observed, however, that none of these clones contained the full-length sequence of the ACP gene (all had a terminally deleted acpP). This was not unexpected as the full-length ACP gene, cloned in a high-copy plasmid, could be unstable in the *E. coli* host (Magnuson *et al*., 1993).

**Sequencing the acpP gene locus using inverse PCR**

Inverse PCR with the two gene-specific primers, IP-21 and IP-22, yielded no amplified product in some cases (*SalI-, SmaI- and EcoRI-digested fragments*), but in other cases (*Apal-, BstHII- and XmaI-digested fragments*) unique inverse PCR products with discrete sizes were obtained (Fig. 3b), that matched with the results of Southern experiments shown in Fig. 3(a) (about 1.6 kb for *Apal*, 1.3 kb for *BsaHII* and 3 kb for *XmaI*). Partial sequencing of both strands in the target region of the cloned recombinant plasmids (pIGM-A, pIGM-B and pIGM-X), derived independently from inverse PCR products, showed identical nucleotide sequences flanking the 102 bp core acpP region (Fig. 4). This proved the reliability of the inverse PCR strategy for cloning and sequencing of the full-length acpP gene.

**Analysis of the nucleotide sequence of acpP and the deduced primary structure of *A. brasilense* ACP**

The complete nucleotide sequence (Fig. 4) of the *A. brasilense* acpP gene was compiled from the overlapping sequences of the PCR and inverse PCR products, revealing an ORF encoding a polypeptide of 79 amino acid

---
residues. The G+C content in the third base of the codons of this ORF is 88.6 mol% which is in agreement with the previous report (Zimmer et al., 1991) of high G+C content of other genes of this bacterium. It has the putative ribosome-binding site GAAGG at 10–14 nt upstream (Fig. 4, positions 109–113) from the ATG start codon. A probable transcription terminator region, 58 nt downstream from the stop codon, has a T-rich sequence preceded by a G+C-rich sequence with dyad symmetry (position 381–410 in Fig. 4). This could form a stable stem–loop structure once transcribed, and could be a rho-independent transcription terminator. No consensus sequences (−10, −35) for σ70 recognition sites could be identified upstream of the ATG start codon.

The deduced amino acid sequence (Fig. 4) reveals that the ACP has 79 amino acid residues with a high sequence homology with other bacterial ACPs (e.g. 81% and 68% identity with ACPs of R. meliloti and E. coli, respectively). The Ser residue predicted at position 2 differs from the Ala residue detected in the protein sequence of the major polypeptide. This suggests post-translational modification of the resulting N-terminal residue as has been observed previously (Hirel et al., 1989; Kawakami & Ohmori, 1994). A. brasilense ACP contains a high number of hydrophobic, and more acidic than basic amino acid residues (35, 20 and 8 out of 79, respectively). The calculated molecular mass of the mature ACP of A. brasilense comprising 78 amino acid residues is 8752 kDa (comparing with 8847 kDa for E. coli ACP), taking into account the addition of a 4′ phosphopantetheine prosthetic group at Ser-36.

**DISCUSSION**

We found it almost impossible to clone the full-length ACP gene of A. brasilense in E. coli. A search for a clone containing the ACP gene in the library failed even when a homologous acpP gene fragment was used as a probe. Moreover, our attempts to isolate a full-length acpP clone from a pUC18-based subgenomic library of A. brasilense were unsuccessful. This is in contrast to the finding of Revill & Leaday (1991) who reported high-level expression of part of the A. brasilense ACP gene in E. coli. However, cloning and sequencing of part of the A. brasilense ACP gene was achieved by determining the N-terminal amino acid sequence of purified ACP and using this sequence to construct A. brasilense-specific degenerate primers. The gene was then cloned by inverse PCR using primers based on the sequence determined from the initial PCR products.

The presence of two ACP protein bands having different N-terminal amino acid residues suggested the possibility of two ACPs of similar structure. However, when several independently derived clones were sequenced no evidence was found for more than one ACP gene. The nucleotide sequence of the 5′ and 3′ flanking region of the ACP gene revealed features suggestive of a stable ACP transcript. For example, in addition to the stem–loop structure of a typical rho-independent terminator at the 3′ downstream sequence, the second potential hair-pin structure around the stop codon (position 349–376, Fig. 4) could be a barrier to 3′ endonucleolytic degradation (Higgins, 1991). The nine nucleotide A-T sequence between the start site and the proposed ribosome-binding site, and the A following the ATG, might also impart mRNA stability as suggested from a study of the highly stable lpp mRNA of E. coli by Nakamura et al. (1980). A stable transcript might be a contributory factor to the high level of expression of ACP.

The role of ACP in bacteria is an enigma. It is not only essential as a cofactor for de novo fatty acid biosynthesis in bacteria, but it (or its acyl derivative) also plays important roles in some essential and accessory cellular functions (Beyreuther et al., 1978; Niki et al., 1992; Shearman et al., 1986; Therisod & Kennedy, 1987; Geiger et al., 1991; Platt et al., 1990). However, diverse functional roles of constitutive ACPs have not been investigated to any significant extent in bacteria other than enterobacteria.

**ACKNOWLEDGEMENTS**

The authors thank the United Nations Development Program (Grant IND/87/017), the Department of Biotechnology and the Department of Science and Technology, Government of India, for financial support of this work, and the Council for Scientific and Industrial Research, Government of India, for providing a Senior Research Fellowship to M. K. M. We are indebted to R. Angeletti (Albert Einstein College of Medicine, New York) for N-terminal amino acid sequence data, and the staff of the Distributive Information Centre (Bose Institute) for their expert help with computer analysis. We also thank Professor S. K. Sen for providing laboratory facilities to do some critical experiments reported here and for his keen interest in this work.

**REFERENCES**


protein is governed by the side chain length of the penultimate amino acid. Proc Natl Acad Sci USA 86, 8247–8251.


Received 7 December 1995; revised 8 March 1996; accepted 18 March 1996.