Purification, characterization, DNA sequence and cloning of a pimeloyl-CoA synthetase from *Pseudomonas mendocina* 35

Andrew BINIEDA*, Martin FUHRMANN*, Bruno LEHNER*, Claudine REY-BERTHOD*, Séverine FRUTIGER-HUGHES†, Graham HUGHES† and Nicholas M. SHAW*¹

*Biotechnology Department, Lonza AG, CH-3930 Visp, Switzerland, and †Département de Biochimie Médicale, Centre Médical Universitaire, CH-1211 Genève 4, Switzerland

A pimeloyl-CoA synthetase from *Pseudomonas mendocina* 35 was purified and characterized, the DNA sequence determined, and the gene cloned into *Escherichia coli* to yield an active enzyme. The purified enzyme had a pH optimum of ≈ 8.0 , K_m values of 0.49 mM for pimelic acid, 0.18 mM for CoA and 0.72 mM for ATP, a subunit M_r of ≈ 80000 as determined by SDS/PAGE, and was found to be a tetramer by gel-filtration chromatography. The specific activity of the purified enzyme was 77.3 units/mg of protein. The enzyme was not absolutely specific for pimelic acid. The relative activity for adipic acid (C_6) was 72% and for azelaic acid (C_9) was 18% of that for pimelic acid (C_7). The N-terminal amino acid was blocked to amino acid sequencing, but controlled proteolysis resulted in three peptide fragments for which amino acid sequences were obtained. An oligonucleotide gene probe corresponding to one of the amino

acid sequences was synthesized and used to isolate the gene (*pauA*, pimelic acid-utilizing *A*) coding for pimeloyl-CoA synthetase. The *pauA* gene, which codes for a protein with a theoretical M_r of 74643, was then sequenced. The deduced amino acid sequence of the enzyme showed similarity to hypothetical proteins from *Archaeoglobus fulgidus*, *Methanococcus jannaschii*, *Pyrococcus horikoshii*, *E. coli* and *Streptomyces coelicolor*, and some limited similarity to microbial succinyl-CoA synthetases. The similarity with the protein from *A. fulgidus* was especially strong, thus indicating a function for this unidentified protein. The *pauA* gene was cloned into *E. coli*, where it was expressed and resulted in an active enzyme.

Key words: *Archeoglobus fulgidus*, biotin synthesis, hybrid pathway, *pauA* gene, sequence similarity.

INTRODUCTION

Pimeloyl-CoA is the first common intermediate on the biotin-biosynthetic pathway in a number of bacteria (Scheme 1). It is synthesized either directly from pimelic acid and CoA by pimeloyl-CoA synthetase, or by a modified fatty acid-synthesis pathway [1,2]. The pimeloyl-CoA synthetase from *Bacillus megaterium* has been partially purified and characterized [3], and that from *Bacillus sphaericus* has been expressed in *Escherichia coli* and purified to homogeneity [4]. This enzyme is highly specific for pimelic acid and is the product of the *bioW* gene, suggesting that it has a unique role in biotin biosynthesis in this bacterium. It is now clear that, in *E. coli*, pimeloyl-CoA is synthesized by a modification of fatty acid biosynthesis, but functions have not yet been allocated to the products of the *bioC* and *bioH* genes, which are known to be involved before pimeloyl-CoA in the pathway.

The final aim of our work is to develop a biotechnological process for the production of biotin. To this end we have cloned the entire biotin (*bio*) operon from *E. coli* and over-expressed the enzymes, both in *E. coli* and in an *Agrobacterium/Rhizobium* strain [5]. In *E. coli*, one problem is the supply of precursors to the pathway. Since *bioC* overexpression has been reported to inhibit cell growth, and biotin production is growth-dependent, we have been exploring other routes to supply pimeloyl-CoA to the pathway *in vivo*. One strategy was to clone a highly active pimeloyl-CoA synthetase gene from another organism to form a hybrid pimeloyl-CoA synthetase-*bio* operon, and to produce

biotin starting from pimelic acid instead of from a carbon source such as glucose or sucrose. To do this we selected a catabolic enzyme from *Pseudomonas mendocina* 35. Here we describe its purification, characterization, partial amino acid sequence, complete DNA sequence and cloning of the corresponding gene into *E. coli*, where the product of the cloned gene was active. This is the first report of the characterization, DNA sequence and cloning of a catabolic pimeloyl-CoA synthetase. The work also serves to identify a protein coded by the genome of *A. fulgidus*.

METHODS

Bacterial strains

Ps. mendocina 35, *Acinetobacter calcoaceticus* 19, *Ps. aeruginosa* PAO1 (wt), *E. coli* K38 and *E. coli* XL1-Blue [6] were from the Lonza strain collection.

Growth media

Ps. mendocina 35 and *A. calcoaceticus* 19 were grown at 30 °C on minimal medium containing either 0.2% (w/v) pimelic acid or 0.2% pimelic acid plus 0.2% succinic acid. The medium contained, per litre: Solution (A+N), 100 ml; Solution SL4, 1 ml; Solution B, 25 ml; 20% (w/v) pimelic acid, 10 ml [or 20% (w/v) pimelic acid plus 20% succinic acid, 10 ml]. Solution (A+N) was autoclaved separately and added to the medium afterwards. It contained, per litre: $(\text{NH}_4)_2\text{SO}_4$, 20 g; $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$, 25 g; KH_2PO_4 , 10 g; NaCl, 30 g. Solution SL4

Abbreviation used: DTT, dithiothreitol.

¹ To whom correspondence should be addressed (e-mail nicholas.shaw@lonza.ch).

The DNA sequence of the *pauA* gene has been deposited in the EMBL, GenBank® and DDBJ Nucleotide Sequence Databases under the accession number AJ012480.

Scheme 1 Biotin-biosynthetic pathway in *E. coli*

Abbreviations used: KAPA, 7-keto(= oxo)-8-aminopelargonic acid; PLP, pyridoxal 5'-phosphate; DAPA, 7,8-diaminopelargonic acid; SAM, S-adenosyl-L-methionine; 'S', unidentified sulphur donor.

contained, per litre: ZnSO₄·7H₂O, 100 mg; MnCl₂·4H₂O, 90 mg; H₃BO₃, 300 mg; CoCl₂·6H₂O, 200 mg; CuCl₂·2H₂O, 10 mg; NiCl₂·6H₂O, 20 mg; Na₂MoO₄·2H₂O, 30 mg; Na₂EDTA·2H₂O, 5 g; FeSO₄·7H₂O, 2 g. Solution B contained, per litre: MgCl₂·6H₂O, 16 g; CaCl₂·2H₂O, 0.58 g; FeCl₃·6H₂O, 0.032 g. Cells grown for 48 h were harvested in late-exponential phase [attenuance at 600 nm (*D*₆₀₀) = 2] and washed once in saline [0.9% (w/v) NaCl] before storage at -20 °C.

Ps. aeruginosa PAO1 was grown on the minimal medium described above, with 0.2% pimelic acid as carbon source. Cells were harvested in the exponential phase.

E. coli XL1-Blue was grown at 30 °C on NYB medium [Nutrient Broth No. 2 (Oxoid), 25 g/litre; Yeast-Extract (Oxoid), 5 g/litre]. The cells were harvested in exponential phase and washed once with 0.9% saline before cell breakage.

Preparation of cell-free extracts

Cells from 100 ml of culture were resuspended in 2 ml of 100 mM Hepes buffer, pH 7.5, containing 1 mM dithiothreitol (DTT) and placed in an ice bath. A 20 µl portion of 50 mM PMSF was added and the cell suspension sonicated seven times at 50 W for 10 s

with 50 s cooling in between. Cell debris was removed by centrifugation.

Chemicals and radioisotopes

[1,7-¹⁴C]Pimelic acid (8.0 mCi/mmol) was from ICN. Pimeloyl-CoA and [1,7-¹⁴C]pimeloyl-CoA (1.72 mCi/mmol) were custom-synthesized by Dr. B. Martinoni, Eidgenössische Technische Hochschule, Zurich, Switzerland. All other reagents were from either Sigma or Fluka.

Measurement of pimeloyl-CoA synthetase activity

One unit of activity is defined as 1 μ mol of pimeloyl-CoA synthesized/min. Pimeloyl-CoA synthetase activity *in vitro* was measured by a coupled spectrophotometric assay system with pig muscle myokinase (1800 units/ml), rabbit muscle pyruvate kinase (2000 units/ml) and rabbit muscle lactate dehydrogenase (2750 units/ml). The oxidation of NADH to NAD⁺ was monitored spectrophotometrically at 340 nm. The production of 1 mol of AMP results in the oxidation of 2 mol of NADH. The molar absorption coefficient of NADH at 340 nm was taken to be 6.22×10^3 litre $\cdot$ mol⁻¹ $\cdot$ cm⁻¹. The reaction mixture contained, in a final volume of 1 ml: 10 mM MgCl₂, 1.5 mM ATP, 1.5 mM phosphoenolpyruvate, 0.24 mM NADH, 0.4 mM CoASH, lactate dehydrogenase (2.75 units), pyruvate kinase (2.0 units), myokinase (7.2 units), 5 mM pimelic acid and cell extract, plus 100 mM Hepes buffer, pH 7.5, containing 1 mM DTT. All assay components were prepared in the Hepes/DTT buffer, and the 50 mM pimelic acid stock solution was also adjusted to pH 7.5 with NaOH. The rate of reaction was determined at 30 °C (a) without enzyme (cell extract) and pimelic acid (b) without pimelic acid, and (c) with all the assay components present.

Purification of the pimeloyl-CoA synthetase from *Ps. mendocina* 35

All materials for column chromatography were from Pharmacia.

DEAE-Sepharose CL-6B anion-exchange chromatography

Cell-free extract was loaded on to a 50 mm $\times$ 73 mm column (bed vol. 143 ml) that had been equilibrated with 100 mM Hepes buffer, pH 7.5, containing 1 mM DTT. The column was washed with the same buffer containing 100 mM NaCl at a flow rate of 5 ml/min until no further protein was eluted from the column. Pimeloyl-CoA synthetase was eluted with a linear gradient of NaCl from 100 mM to 300 mM in 100 mM Hepes buffer, pH 7.5, containing 1 mM DTT. The gradient volume was 750 ml, the flow rate 1 ml/min and 10 ml fractions were collected. The enzyme was eluted from the column at about 200 mM NaCl. Fractions active in the spectrophotometric assay were pooled and concentrated by ultrafiltration (Amicon YM10 membrane).

Superose 12 Prep gel-filtration chromatography

Concentrated sample from the DEAE-Sepharose step was loaded on to a 100 ml volume Superose-12 Prep column (16 mm $\times$ 500 mm), equilibrated with 100 mM Hepes buffer, pH 7.5, containing 1 mM DTT. The flow rate was 0.5 ml/min and 2 ml fractions were collected. The enzyme was eluted in a volume of $\approx$ 10 ml, with an elution volume between 40 and 50 ml (fractions 20–25).

Chromatofocusing on Mono-P

The buffer in the sample was exchanged to the start buffer for the Mono-P column, namely 25 mM BisTris, pH 6.3, by means of a Sephadex G25 PD-10 gel-filtration column.

Chromatofocusing was carried out as described in the Pharmacia instruction brochure with a linear gradient from pH 6.0 to pH 4.0. The enzyme was eluted as a single sharp peak at pH 5.3.

Identification of the pimeloyl-CoA synthetase reaction product

A crude desalted preparation of pimeloyl-CoA synthetase from *Ps. mendocina* 35 (250 μ l, 8 mg/ml protein) in 100 mM Hepes buffer, pH 7.5, was incubated in a total volume of 2.5 ml with 13 mM CoASH, 12 mM pimelic acid, 12 mM ATP, 10 mM MgCl₂, 1.5 mM phosphoenolpyruvate, 10 μ l of myokinase and 2.5 μ l of pyruvate kinase. The stock solution of pimelic acid was made up in 0.5 M Hepes buffer, pH 7.13, and 600 μ l of this was added to the assay. [¹⁴C]Pimelic acid (1 μ Ci) was added as tracer to facilitate purification of the reaction product. After 22 h incubation at 30 °C the pH was adjusted to 3.0 with 51 μ l of formic acid, and precipitated protein removed by centrifugation. The sample was then loaded on to a Pharmacia HR 10/10 Mono Q column, and eluted with a linear gradient of ammonium formate, pH 3.0, from 10 mM to 1.0 M. The flow rate was 4 ml/min, the gradient volume was 272 ml and 8 ml fractions were collected. Radioactivity in the fractions was determined by liquid-scintillation counting. Two peaks of radioactivity were seen, and the constituent fractions of each peak were pooled and freeze-dried. The peak corresponding to pimeloyl-CoA contained a calculated 4.2 mg of dry material, which was analysed by NMR, in ²H₂O with a Nicolet NT 300 spectrometer, operating at 300 MHz.

PAGE

SDS/PAGE was carried out essentially as described in [7].

Analytical gel filtration

This was carried out with a Superose-12 analytical column equilibrated with 100 mM Tris/HCl buffer, pH 6.8, containing 15 mM NaCl. The following proteins were used as markers: ribonuclease, M_r 13 700; chymotrypsin, M_r 25 000; BSA, M_r 67 000; aldolase, M_r 158 000; catalase, M_r 232 000; ferritin, M_r 440 000. The standards were prepared at a concentration of 5 mg/ml, except for ferritin, which was 2 mg/ml. The sample volume was 200 μ l.

Determination of kinetic constants

All assays were carried out in 100 mM Bicine buffer, pH 8.0. Before determining the kinetic constants of the purified enzyme, the linearity of the reaction with respect to protein concentration was established. For K_m determinations, 0.42 μ g of the purified enzyme was used and the standard concentrations of substrates in the assays were: ATP, 1.5 mM; CoASH, 0.4 mM; pimelic acid, 5 mM. The concentration of the substrate for which the K_m was being determined was then varied. Kinetic values were calculated with the Enzfitter and Fig. P computer programs from Biosoft. Two determinations were carried out, with similar results. Typical results are presented.

Cloning of the *Ps. mendocina* 35 *pauA* gene

A mixed oligonucleotide deduced from the amino acid sequence of pimeloyl-CoA synthetase tryptic peptide 3 (Ala-Leu-Phe-Glu-

Asp-Pro-Thr) was synthesized (Microsynth, Windisch, Switzerland) and used as a specific DNA probe. It was end-labelled with [γ - ^{32} P]ATP and hybridized to a Southern blot of *Ps. mendocina* 35 total DNA cut with different restriction enzymes. Specific restriction fragments binding the radioactive probes were: *EcoRI*, 9.4 kb; *EcoRV*, 2.1 kb; *PstI*, 2.65 kb; *SalI*, 1.8 kb; and *SmaI*, 6.0 kb. To clone the 9.4 kb *EcoRI* fragment a partial gene bank of the *Ps. mendocina* 35 genome was prepared. Fragments (8.5–10.5 kb) from *Ps. mendocina* 35 DNA cut with *EcoRI* were isolated and ligated with plasmid pBluescript KS⁻ that was also cut with *EcoRI*. *E. coli* XL1-Blue was transformed with the ligation mixture. After plating on selective agar [Nutrient agar with 100 μ g/ml ampicillin, 30 μ g/ml 5-bromo-4-chloroindol-3-yl β -D-galactopyranoside ('X-Gal') and 0.5 mM isopropyl β -D-thiogalactoside ('IPTG')], 76 individual clones with recombinant plasmids were isolated.

With these a 'colony blot' was prepared on a nitrocellulose membrane, which was then probed with the same radiolabelled oligonucleotide as above. One colony bound the oligonucleotide. Restriction mapping showed that it contained a recombinant plasmid with the desired 9.4 kb *EcoRI* fragment from the *Ps. mendocina* 35 genome. This fragment included the other specific restriction fragments mentioned above. The plasmid was named pCRB1.

DNA sequence of the pimeloyl-CoA synthetase gene

DNA sequencing was carried out as described in [8]. Single-strand DNA of pBluescript derivatives were used as DNA templates. Sequence reactions were performed with the Sequenase Kit from United States Biochemicals (Cleveland, OH, U.S.A.). Both DNA strands were completely sequenced, and for all cloning sites overlapping sequences were obtained. Analysis of DNA sequences was done with PC/Gene programs from IntelliGenetics Inc. (Mountain View, CA, U.S.A.). The deduced amino acid sequence was compared with existing sequences using the BLAST network at the Swiss Institute for Experimental Cancer Research and through the Expasy web site.

Radioactive labelling of plasmid-encoded pimeloyl-CoA synthetase

Genes cloned into plasmid pBluescript can be selectively expressed via its bacteriophage T₇ promoter, pT7. Recombinant plasmids were transformed into *E. coli* K38 with the plasmid pGP1-2 that contains the gene for bacteriophage T₇ RNA polymerase under control of the bacteriophage- λ promoter P_L [9].

The method for induction of T₇ RNA polymerase and for selective [^{35} S]methionine labelling of proteins encoded by pT7-recombinant plasmids is described in [10].

Determination of pimeloyl-CoA synthetase activity in strains containing the cloned enzyme

For the measurement of pimeloyl-CoA synthetase activity in these strains, determinations were either made in crude cell-free extracts or after chromatography on a 10 mm $\times$ 100 mm (8 ml volume) column of DEAE-Sephadex CL-6B anion-exchange resin. A 30 ml gradient was used, from 100 mM Hepes buffer, pH 7.5, containing 1 mM DTT and 300 mM NaCl to the same buffer containing 1 M NaCl. This simple purification step was included because it was often difficult to measure pimeloyl-CoA synthetase activity in the crude cell-free extracts owing to background ATPase activity.

Protein sequencing

Protein sequencing was by Edman degradation with either a model 473A or a model 477A Microsequencer (Applied Biosystems) equipped with Problott reaction cartridges. Preparation of the peptides derived from a tryptic digest of purified pimeloyl-CoA synthetase for sequencing was as described in [11].

RESULTS AND DISCUSSION

Selection of a bacterium as a source of pimeloyl-CoA synthetase

The utilization of saturated dicarboxylic acids containing from six to ten carbon atoms can support the growth of aerobic bacteria belonging to the genera *Pseudomonas* and *Acinetobacter*. The dicarboxylic acids are metabolized by β -oxidation, and the initial step is activation to the corresponding CoA derivative [12]. Bacteria were screened for growth on pimelic acid (C₇) and crude cell-free extracts of the pimelic acid-utilizing bacteria *A. calcoaceticus* 19, *Ps. mendocina* 35 and *Ps. aeruginosa* PAO1 were prepared by sonication and tested for pimeloyl-CoA synthetase activity *in vitro* using the spectrophotometric assay (Table 1). Pimeloyl-CoA synthetase from *Ps. mendocina* 35 was the most active, especially when pimelic acid was the sole carbon source for growth. This strain was therefore chosen as a source of pimeloyl-CoA synthetase for further study and for cloning. We suppose that pimelic acid as a carbon source for this bacterium is activated to pimeloyl-CoA as a prerequisite to its catabolism, and consequently the pimeloyl-CoA synthetase that we chose for further study is a catabolic enzyme.

Identification of the pimeloyl-CoA synthetase reaction product

The NMR spectra of the putative pimeloyl-CoA synthetase reaction product synthesized with a crude cell-free extract of *Ps. mendocina* 35 and that of the synthetic pimeloyl-CoA were virtually identical, and consistent with the structure of pimeloyl-CoA. The product of the enzyme reaction was therefore pimeloyl-CoA.

Purification of pimeloyl-CoA synthetase from *Ps. mendocina* 35

The enzyme was purified in three steps, using anion-exchange and gel-filtration chromatography, followed by chromatofocusing. The results of a typical purification are summarized in Table 2 and analysis of the purified enzyme by SDS/PAGE is shown in Figure 1.

Amino acid sequence

After purification, the buffer in the sample was exchanged to 50 mM ammonium bicarbonate, pH 8.0, the sample freeze-dried and a partial amino acid sequence determined. The N-terminal amino acid was blocked, so the protein was reduced, carboxymethylated and cleaved specifically with trypsin, and three peptides isolated by reverse-phase HPLC. The sequences of the peptides were:

- (1) Ala-Xaa-Ile-Glu-Ser-Ala-Pro-Glu
- (2) Ile-Leu-Pro-Xaa-Asn-Pro-Asn
- (3) Ala-Leu-Phe-Glu-Asp-Pro-Thr

Table 1 Comparison of pimeloyl-CoA synthetase activity in various bacteria

Bacterial strain	Growth medium	Pimeloyl-CoA synthetase activity (units/mg of protein)
<i>A. calcoaceticus</i> 19	0.2% Pimelic acid + 0.2% succinic acid	Not detected
<i>Ps. mendocina</i> 35	0.2% Pimelic acid + 0.2% succinic acid	0.16 ± 0.016 ($n = 7$)
	0.2% Pimelic acid	0.56 ± 0.05 ($n = 3$)
<i>Ps. aeruginosa</i> PA01	0.2% Pimelic acid	0.08 ± 0.03 ($n = 4$)
	0.2% Pimelic acid + 0.2% glucose	0.11 ± 0.088 ($n = 2$)

Table 2 Purification of the pimeloyl-CoA synthetase from *Pseudomonas mendocina* 35

Step	Vol. (ml)	Protein (mg)	Activity (units)	Sp. activity (units/mg)	Yield (%)	Purification (fold)
Cell-free extract	25	180	(100.8)	(0.56*)	(100)	—
Anion exchange	1.5	8.9	13.35	1.5	13.24	2.7
Gel filtration	4.0	1.7	9.86	5.8	9.78	10.4
Chromatofocusing	1.0	0.025	1.93	77.3	1.91	138

* Estimated activity.

Figure 1 SDS/PAGE of purified pimeloyl-CoA synthetaseLanes A and C, M_r markers; lane B, pimeloyl-CoA synthetase. **Figure 2** Determination of pH optimumEach assay mixture contained 2.3 μ g of purified enzyme.**Absorption spectrum**

The absorption spectrum for the purified enzyme between 200 nm and 700 nm showed a peak at 280 nm, but otherwise there were no other characteristic peaks.

Substrate specificity

Five dicarboxylic acids (50 mM in 100 mM Hepes buffer, pH 7.5, containing 1 mM DTT) were tested as substrates for the enzyme in a crude cell-free extract of *Ps. mendocina* 35. Compared with pimelic acid (C_7 , 100 %), the relative activities were as follows: oxalic acid (C_2), 0 %; glutaric acid (C_5), 0 %; adipic acid (C_6), 72 %; azaleic acid (C_9), 18 %.

Stability

The crude cell-free extract steadily lost activity when stored at 4 °C such that only 10 % of the original activity remained after 8 days. The enzyme, both in crude cell-free extracts and after purification, was most stable when stored at -20° or -80° C.

pH optimum

The pH optimum of the purified enzyme was determined at 30 °C with the coupled assay system, and was $\approx$ pH 8.0 (Figure 2). Because of the complex assay system, this value should be considered as a functional pH optimum.

Native and subunit M_r

The subunit M_r of the pure enzyme was calculated to be between 74000 and 80000 from SDS/PAGE (Figure 1). The sequence of the *pauA* (pimeloyl-CoA synthetase) gene showed that it codes

Figure 3 Restriction map of pCRB1

The fragments removed to give pCRB1-1 (*Cla*I) and pCRB1-2 (*Xho*I–*Mlu*I) are indicated by the cross-hatched boxes.

for a protein with an M_r of 74643. The native M_r was estimated by analytical gel filtration to be between 200000 and 400000 (elution volumes of 11.6–13.6 ml). A more accurate estimate was not possible with this method. The enzyme is therefore probably a tetramer, with an M_r of 298572.

Determination of kinetic constants

Typical Michaelis–Menten kinetics were observed, with K_m values of 0.72 mM for ATP, 0.18 mM for CoASH and 0.49 mM for pimelic acid. The average V_{max} for the enzyme was 25 μ mol/min per mg of protein, which indicated that the enzyme lost some activity on storage or thawing compared with the freshly prepared enzyme (Table 2). In comparison with the biotin-pathway-specific enzyme from *Bacillus sphaericus* (1.0 μ mol/min per mg of protein) [4], the *Ps. mendocina* enzyme is between one and two orders of magnitude more active (77.0 μ mol/min per mg of protein for the freshly prepared enzyme). This difference is probably due to the metabolic function of the *B. sphaericus* enzyme in vitamin synthesis and that of the *Ps. mendocina* enzyme in catabolism.

Cloning of the *Ps. mendocina* 35 *pauA* gene

We named the pimeloyl-CoA synthetase gene *pauA*. To isolate the *pauA* gene, a mixed radiolabelled oligonucleotide corresponding to the sequence of tryptic peptide 3 (Ala-Leu-Phe-Glu-Asp-Pro-Thr) was synthesized and used as a specific DNA probe. The probe was hybridized to a Southern blot of *Ps. mendocina* 35 total DNA cut with different restriction enzymes, and bound to five fragments, the largest being a 9.4 kb *Eco*RI fragment. Since this fragment had the greatest chance of containing the gene for the complete enzyme, a partial gene bank of 8.5–10.5 kb *Eco*RI fragments was prepared from *Ps. mendocina* 35, and ligated with plasmid pBluescript KS[−]. *E. coli* XL1-Blue was transformed with the ligation mixture and clones with recombinant plasmids isolated. One colony contained a recombinant plasmid with the desired 9.4 kb *Eco*RI fragment from the *Ps. mendocina* 35 genome. A restriction map of this plasmid (pCRB1) was prepared (Figure 3). In pCRB2, the restriction fragment was

cloned in the opposite direction. All of the restriction fragments hybridizing with the synthetic oligonucleotide spanned the same area, thus indicating the approximate location of the gene coding for pimeloyl-CoA synthetase. This region was examined by DNA sequencing and by specific expression using the pT7 promoter of pBluescript. Subclones were generated by a two-step strategy. First, several restriction fragments generated from pCRB1 were isolated and re-cloned into the plasmid vectors pBluescript KS[−], KS⁺ or SK⁺, resulting in a series plasmids (pCRB3 to pCRB12; results not shown). The cloning of a DNA fragment in both possible orientations made DNA sequencing and gene expression starting from either end possible.

The plasmids pCRB3 to pCRB12 were further modified by deletion mutagenesis, either by cutting suitable restriction sites and re-ligating the shortened plasmids or by using the exonuclease *Bal*31 when suitable restriction sites were absent. More than 25 of these plasmids were constructed and used for the DNA sequence determination.

DNA sequence of the pimeloyl-CoA synthetase gene

The DNA sequence of the *pauA* gene (2916 nt) was determined and is shown in Figure 4. Computer analysis showed an open reading frame of 711 codons, from position 204 to 2336, yielding a theoretical protein with a M_r of 74643. This corresponds well with the M_r of pimeloyl-CoA synthetase as calculated from SDS/PAGE. Furthermore, the amino acid sequence deduced from the DNA sequence contained the peptides that were isolated and sequenced after a tryptic digest of pimeloyl-CoA synthetase.

The GC content of *pauA* (70% GC) is typical of *Ps.* genes, as is the codon usage [13]. A comparison of the pimeloyl-CoA synthetase amino acid sequence with other proteins showed similarities with hypothetical proteins from *A. fulgidus* [14], *Methanococcus jannaschii* [15], *Pyrococcus horikoshii*, *E. coli* [16] and *Streptomyces coelicolor*. The similarity was especially strong (normalized alignment score 936 [17]) with the conserved hypothetical protein from *A. fulgidus*, thus indicating a function for this unidentified protein. The hypothetical proteins from the organisms mentioned above also had other areas of strong similarity with one another, indicating the possibility of a

<i>P. mendocina</i>	249	KPVILHKVGRSELGGRAAMSH ^T ASLAGDD
<i>A. fulgidus</i>	229	KPVVVMKAGKSQSGAKAASSHTGSLAGSY
<i>M. jannaschii</i>	233	KPIIALKSGRTEVVGKKA ^A AKSHTGSLAGED
<i>P. horikoshii</i>	248	KPIIALKSGRTEYGA ^A KAASSHTGSLAGQD
<i>E. coli</i>	228	KPILVIKSGRSPAAQR--LLNTT--AGMD
<i>S. coelicolor</i>	483	KPVVVLKAGRTAAGAKAAGSHTGALAGDD
<i>P. mendocina</i>	520	AEQIGFPVVLKIASADILHKSEIGGVLLNVG
<i>A. fulgidus</i>	503	AE ^S IGYPVVLKVSPDVVHKSDVGGVKLNVG
<i>M. jannaschii</i>	521	CKKLG-KCVMKIVSPQIIHKTEAGGVIINPK
<i>E. coli</i>	513	AEQIGYPVALKLRSPDIPHKSEVQGVML
<i>S. coelicolor</i>	57	AARFGGPVVMKIVSPDILHKTDAGGVVVGVE
<i>P. mendocina</i>	571	PHARLDGVLVARQLQGGVECFMGIQRDPQFGPVAVFGLGGIFVEVLKDVVFR
<i>A. fulgidus</i>	553	PKARIEGVLVQKMVDGKELIVGMKRD ^P QFGPMIMFGMGGVYVEVLKDV ^S FR
<i>M. jannaschii</i>	592	EIIIGAKRDDIFG ^S VMVGLGGVFVEVLKDV ^S FG
<i>S. coelicolor</i>	107	ASARIEGVQIQELLPQGEVIVGAVTDPTEGKVVAFGLGGVIVEVLKDV ^T FR

A. fulgidus conserved hypothetical protein (AF1211) 685 amino acids. Score 936, identity 36%.

M. jannaschii hypothetical protein (MJ0590) 704 amino acids. Score 788, identity 31%.

P. horikoshii hypothetical protein (PHCY008) 472 amino acids. Score 604, identity 33%.

E. coli hypothetical 98.0 kD protein in UNGPSSA intergenic region (YFIQ) 886 amino acids. Score 581, identity 28%.

S. coelicolor hypothetical 75.3 kD protein (SCA6.03C) 714 amino acids. Score 549, identity 35%.

Figure 5 Amino acid sequence similarities between the pimeloyl-CoA synthetase from *Ps. mendocina* and hypothetical proteins found in the BLAST database

Amino acids identical with those in the *Ps. mendocina* 35 enzyme are shown in **bold** type. To show the similarities between the proteins found in the database search, all other identical amino acids are underlined.

CODFICK and CODRNY), this is a coding region. The GC content and codon usage are very similar to the *pauA* gene. The open reading frame is closely linked to *pauA*; in fact five codons are overlapping. Furthermore, it is preceded by the prominent Shine-Dalgarno sequence AGGAGG. Therefore we assign to it a hypothetical gene *pauB* that forms an operon with *pauA* and may also be involved in pimelic acid metabolism. This hypothesis is, however, weakened by the presence of a 9 bp inverted repeat, centred around nt 2403. This structure might act as a transcriptional terminator approx. 50 bp downstream of the *pauA* 3' end. Neither possibility could be excluded in further experiments and, up to now, we have no final proof for the existence of a *pauB* gene. A database comparison did reveal some sequence similarities, for example with the multifunctional β -oxidation protein from *Neurospora crassa*. The significance of this observation remains to be established.

Expression and specific labelling of the protein encoded by the *pauA* gene

The expression of the proteins encoded by various pCRB plasmids was analysed by specific expression of plasmid-encoded proteins from the bacteriophage T₇ promoter, also contained in the pCRB plasmids. We were able to identify unambiguously the *pauA* gene product as an approx.-74 000-*M_r* protein (Figure 6).

Determination of pimeloyl-CoA synthetase activity in cell-free extracts from strains containing the cloned enzyme

The results for the various strains tested show the successful cloning of the active *Ps. mendocina* 35 enzyme in *E. coli* and confirm the identity of the enzyme (Table 3). The levels of

Figure 6 Expression and specific labelling of proteins encoded by the *pau* genes

Expression of plasmid-encoded proteins from the bacteriophage T₇ promoter. Lane A, pCRB2; lane B, pCRB2-1 with the *XhoI*-*MluI* fragment shown in Figure 3 deleted; lane C, plasmid without insert (control); lane D, *M_r* markers.

activity achieved were comparable with those in *Ps. mendocina* 35. The extract from cells containing pCRB1 had no pimeloyl-CoA synthetase activity, but removal of either a *ClaI* or a *XhoI*-*MluI* fragment (pCRB1-1 and pCRB1-2; Figure 3) resulted in activity, presumably by removal of a regulatory element that blocks the transcription of *pauA*. We have not determined the nature of this regulatory mechanism.

Table 3 Pimeloyl-CoA synthetase activity in strains containing the cloned *pauA* gene

Strain	Plasmid	Activity in cell-free extract (units/mg of protein)	Activity in the most active column fraction (units/mg of protein)
<i>E. coli</i> XL1 Blue	pCRB 1	Not detected	Not detected
	pCRB 2	0.047	0.32
	pCRB 1-1 (— IPTG)	Not detected	Not detected
	pCRB 1-1	0.122	0.99
	pCRB 1-2	0.055	0.58
<i>Ps. mendocina</i> 35	—	0.33	1.65

After completion of the work described above, the *pauA* gene was available for further cloning to form hybrid *bio* operon-*pauA* plasmids for biotin production from exogenous pimelic acid.

On the basis the activity of the enzyme, we suggest classification as pimelate-CoA ligase (6.2.1.-).

REFERENCES

- Sanyal, I., Lee, S.-L. and Flint, D. (1994) *J. Am. Chem. Soc.* **116**, 2637–2638
- Ifuku, O., Miyaoka, H., Koga, N., Kishimoto, J., Haze, S., Wachi, Y. and Kajiwara, M. (1994) *Eur. J. Biochem.* **220**, 585–591
- Izumi, Y., Morita, H., Tani, Y. and Ogata, K. (1974) *Agric. Biol. Chem.* **38**, 2257–2262
- Ploux, O., Soularue, P., Marquet, A., Gloeckler, R. and Lemoine, Y. (1992) *Biochem. J.* **287**, 685–690
- Birch, O., Brass, J., Fuhrmann, M. and Shaw, N. (1994) *PCT Pat. Appl. WO* 94/08023
- Bullock, W. O., Fernandez, J. M. and Short, J. M. (1987) *Biotechniques* **5**, 376–379
- Ferreira, R. B. and Davies, D. D. (1987) *Plant Physiol.* **83**, 869–877
- Sanger, F., Niklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Tabor, S. and Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1074–1078
- Streber, W. R., Timmis, K. N. and Zenk, M. H. (1987) *J. Bacteriol.* **169**, 2950–2955
- Hughes, G. J., Frutiger, S., Paquet, N. and Jaton, J.-C. (1990) *Biochem. J.* **271**, 641–647
- Hoet, P. P. and Stanier, R. Y. (1970) *Eur. J. Biochem.* **13**, 65–70
- West, S. E. H. and Jglewski, B. H. (1988) *Nucleic Acids Res.* **16**, 9323–9335
- Klenk, H. P., Clayton, R. A., Tomb, J., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D. et al. (1997) *Nature (London)* **390**, 364–370
- Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D. et al. (1996) *Science* **273**, 1058–1073
- Blattner, F. R., Plunkett, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F. et al. (1997) *Science* **277**, 1453–1474
- Doolittle, R. F. (1986) *Of Urfs and Orfs*, University Science Books, Mill Valley, CA

Received 3 February 1999/17 March 1999; accepted 12 April 1999