THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 273, No. 40, Issue of October 2, pp. 25974–25986, 1998

Printed in U.S.A.

Catabolism of Phenylacetic Acid in *Escherichia coli*

CHARACTERIZATION OF A NEW AEROBIC HYBRID PATHWAY*

(Received for publication, May 8, 1998, and in revised form, July 16, 1998)

Abel Ferrández‡§, Baltasar Miñambres¶, Belén García¶, Elías R. Olivera¶, José M. Luengo¶, José L. García‡, and Eduardo Díaz‡‖

From the ‡Department of Molecular Microbiology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, 28006 Madrid and the ¶Department of Biochemistry and Molecular Biology, Facultad de Veterinaria, Universidad de León, 24007 León, Spain

The *paa* cluster of *Escherichia coli* W involved in the aerobic catabolism of phenylacetic acid (PA) has been cloned and sequenced. It was shown to map at min 31.0 of the chromosome at the right end of the *mao* region responsible for the transformation of 2-phenylethylamine into PA. The 14 *paa* genes are organized in three transcription units: *paaZ* and *paaABCDEFGHIJK*, encoding catabolic genes; and *paaXY*, containing the *paaX* regulatory gene. The *paaK* gene codes for a phenylacetyl-CoA ligase that catalyzes the activation of PA to phenylacetyl-CoA (PA-CoA). The *paaABCDE* gene products, which may constitute a multicomponent oxygenase, are involved in PA-CoA hydroxylation. The PaaZ protein appears to catalyze the third enzymatic step, with the *paaFGHIJ* gene products, which show significant similarity to fatty acid β-oxidation enzymes, likely involved in further mineralization to Krebs cycle intermediates. Three promoters, *Pz*, *Pa*, and *Px*, driven the expression of genes *paaZ*, *paaABCDEFGHIJK*, and *paaX*, respectively, have been identified. The *Pa* promoter is negatively controlled by the *paaX* gene product. As PA-CoA is the true inducer, PaaX becomes the first regulator of an aromatic catabolic pathway that responds to a CoA derivative. The aerobic catabolism of PA in *E. coli* represents a novel hybrid pathway that could be a widespread way of PA catabolism in bacteria.

Escherichia coli living in the animal gut encounters aromatic compounds such as phenylacetic acid (PA),[1] phenylpropionic acid, and their hydroxylated derivatives, as a result of the action of intestinal microflora on plant constituents, the amino acids phenylalanine and tyrosine, fatty acids with a terminal phenyl substituent, and some of their metabolites (1–3). The aerobic catabolism of these aromatic compounds by *E. coli* could occur close to the epithelial cells in the guts of warm-blooded animals, as well as in soil, sediment, and water once *E. coli* is excreted from its intestinal residence (4). The ability of *E. coli* to mineralize 3- and 4-hydroxyphenylacetic acids (5), 3-phenylpropionic, 3-(3-hydroxyphenylpropionic), and 3-hydroxycinnamic acids (2, 6), and phenylacetic acid (2, 7) has been reported previously. Recently, the molecular characterization of these catabolic pathways, with the only exception of that for PA degradation, has been carried out (1, 8–11), demonstrating that *E. coli* is endowed with its own set of genes and enzymes for the catabolism of aromatic compounds, and that they are similar to those of other microorganisms more relevant in the environment such as bacteria of the genus *Pseudomonas*.

Although PA is a common source of carbon and energy for a wide variety of microorganisms, the bacterial catabolism of this natural aromatic compound is still poorly understood (12, 13). Earlier reports suggested that aerobic PA catabolism implicated the typical initial attack by hydroxylation of the aromatic ring with the formation of the corresponding 2,5- or 3,4-dihydroxyphenylacetate as intermediates (13). However, much of this evidence was circumstantial, and none of the typical aerobic routes that could explain PA degradation were responsible of this catabolism in different PA-degrading bacteria (13, 14). According to these data, it has been recently shown that *Pseudomonas putida* U mineralizes PA aerobically through a novel catabolic pathway, which does not follow the conventional routes for the aerobic catabolism of aromatic compounds and whose first step is the activation of PA to phenylacetyl-coenzyme A (PA-CoA) by the action of a PA-CoA ligase (12, 15). In this sense, the participation of a PA-CoA ligase in the aerobic catabolism of PA has been also inferred from its specific induction during growth on PA of different bacterial strains (13, 16).

Here we present the cloning, genetic characterization, mechanism of regulation, and a partial biochemical characterization of the PA biodegradation pathway from *E. coli* W. This work reveals that the PA degradation in *E. coli* follows an unusual route for the aerobic catabolism of aromatic compounds, which involves CoA derivatives. With the molecular characterization of the *paa*-encoded pathway, all aromatic catabolic routes so far reported in *E. coli* are now described at the molecular level.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Plasmids, and Growth Conditions—The *E. coli* strains used were *E. coli* W ATCC11105 (17); *E. coli* W14 (18); *E. coli* C (8); and the *E. coli* K-12 strains MV1190 (19), C600 (20), TG1 (20), ET8000 (21), W3110 (20), MG1655 (11), HB101 (20), DH1 (20), CC118 (1), DH5α (20), JM109 (20), and S17–1λ*pir* (22). *E. coli* W14Rif (this work) is a spontaneous rifampicin-resistant mutant from *E. coli* W14; *E. coli* AF141 (this work) is a *lacZ*$^-$ mutant obtained from the *E. coli*

* This work was supported by Grants AMB94-1038-C02-02 and AMB97-063-C02-02 from Comisión Interministerial de Ciencia y Tecnología, Spain. 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. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) X97452.

§ Recipient of a predoctoral fellowship from the Plan Nacional de Formación de Personal Investigador-MEC.

‖ Supported by a Contrato Temporal de Investigadores from the Consejo Superior de Investigaciones Científicas. To whom all correspondence should be addressed: Dept. of Molecular Microbiology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Velázquez 144, 28006 Madrid, Spain. Tel.: 34-91-5611800; Fax: 34-91-5627518; E-mail: cibdf4f@fresno.csic.es.

[1] The abbreviations used are: PA, phenylacetic acid; bp, base pair(s); FNR, ferredoxin-NADP$^+$ reductase; 2-HPA, 2-hydroxyphenylacetate; HPLC, high performance liquid chromatography; kb, kilobase pair(s); ORF, open reading frame; PA-CoA, phenylacetyl-coenzyme A; PCR, polymerase chain reaction.

W14Rif strain by *N*-methyl-*N*′-nitro-*N*-nitrosoguanidine treatment as described by Miller (23). For cloning and expression purposes we have used two chloramphenicol-resistant low copy number cloning vectors, plasmids pCK01 (1) and pSJ19Not (like pCK01 but containing the polylinker region from pUC19), as well as pUC18, pUC19 (20), and pUC18Not (22) vectors. The promoter-probe pSJ3 vector is a pUJ9 derivative (22) bearing a 85-bp *Sma*I/*Bam*HI fragment that contains the 5′-leader region of IS*10* transposase (19). To integrate the *lacZ* fusions into the chromosome, plasmid pUTmini-Tn*5*Km2 was used (22). Unless otherwise stated, bacteria were grown in Luria-Bertani (LB) medium (20) at 37 °C. When used as carbon sources, PA and/or glycerol were supplied at 5 and 20 mM, respectively, to M63 minimal medium (23) containing the corresponding necessary nutritional supplements, and the cultures were incubated at 30 °C. Where appropriate, antibiotics were added at the following concentrations: ampicillin (100 μg/ml), chloramphenicol (35 μg/ml), kanamycin (50 μg/ml), and rifampicin (50 μg/ml).

DNA Manipulations and Sequencing—Plasmid DNA was prepared by the rapid alkaline lysis method (20). Transformation of *E. coli* was carried out using the RbCl method (20). DNA manipulations and other molecular biology techniques were essentially as described (20). DNA fragments were purified by using low melting point agarose. Southern blot analyses were performed as previously reported (20), using as probes DNA fragments labeled with [α-^{32}P]dCTP by the random primer method (Amersham Pharmacia Biotech). Pulsed-field gel electrophoresis of total DNA from *E. coli* W was carried out as described (24). Oligonucleotides were synthesized on an Oligo-1000M nucleotide synthesizer (Beckman Instruments, Inc.). Nucleotide sequences were determined directly from plasmids by using the dideoxy chain termination method (25). Standard protocols of the manufacturer for *Taq* DNA polymerase-initiated cycle sequencing reactions with fluorescently labeled dideoxynucleotide terminators (Applied Biosystems Inc.) were used. The sequencing reactions were analyzed using a model 377 automated DNA sequencer (Applied Biosystems Inc.). Sequences were extended by designing primers based on the previously determined sequence.

Sequence Data Analyses—Nucleotide sequence analyses were done with the DNA-Strider 1.2 program. Amino acid sequences were analyzed with the Protein Analysis Tools at the ExPASy World Wide Web molecular biology server of the Geneva University Hospital and the University of Geneva. Nucleotide and protein sequence similarity searches were made by using the BLASTP, BLASTN, and BLASTX programs (26) via the National Institute for Biotechnology Information server. Pairwise and multiple protein sequence alignments were made with ALIGN (27) and CLUSTAL W (28) programs, respectively, at the Baylor College of Medicine-Human Genome Center server. The *E. coli* data base collection ECDC (29) was accessed via the Internet.[2]

Cloning and Expression of the Pa, Px, and Pz Promoter Regions—The Pa_{424} promoter region was PCR-amplified from plasmid pAAD by using primers PZ5 (5′-GGGGTGAATCAAACGGCTACG-3′; the sequence corresponds to nucleotides 2402–2420 in Fig. 2) and PA5–1 (5′-CAATCTCGGAATGCGCATG-3′; the sequence corresponds to nucleotides 2885–2867 in Fig. 2). The resulting 495-bp DNA fragment was digested with *Kpn*I and *Sph*I and cloned as a 424-bp fragment into the double-digested *Kpn*I+*Sph*I pSJ3 vector to form plasmid pAFPA1 (Fig. 5). The Pa_{255} promoter region was PCR-amplified by using primers PAP (5′-GGTCTAGAGTTATCAAAATAGAGTGCG-3′; the sequence corresponds to nucleotides 2611–2631 in Fig. 2; engineered *Xba*I restriction site is underlined) and PA5–1. The resulting 272-bp DNA fragment was digested with *Xba*I and *Sph*I and cloned as a 255-bp fragment into the double-digested *Xba*I+*Sph*I pSJ3 vector to form plasmid pAFPA2 (Fig. 5). The *Pz* promoter region was PCR-amplified by using primers PZ5–1 (5′-GGGGATCCCGGCCAGACTGCCAGGTAC-3′; the sequence corresponds to nucleotides 2423–2443 in Fig. 2; engineered *Bam*HI restriction site is double-underlined) and PA5–1. The resulting 474-bp DNA fragment was digested with *Sph*I and *Bam*HI and cloned as a 446-bp fragment into the double-digested *Sph*I+*Bam*HI pSJ3 vector to form plasmid pAFPZ (Fig. 5). The *Px* promoter region was PCR-amplified by using primers 5,5f2 (5′-CCGCATGGATCGCATCAGC-3′; the sequence corresponds to nucleotides 11910–11928 in Fig. 2) and PX5–2 (5′-GGGGATCCGAATCACCATACAGAGAGGAG-3′, the sequence corresponds to nucleotides 12456–12436 in Fig. 2; engineered *Bam*HI site is double-underlined). The resulting 554-bp DNA fragment was digested with *Ssp*I and *Bam*HI, and cloned as a 214-bp fragment into the double-digested *Sma*I+*Bam*HI pSJ3 vector to form plasmid pAFPX (Fig. 5).

Tn1000 Transposition Mutagenesis of paa Genes—Transposition with Tn*1000* was carried out according to the previously described method (30). The pAAD plasmid was mutagenized with Tn*1000* by selecting for conjugative transfer of this plasmid from the *recA* F^+ donor strain *E. coli* MG1063 to the F^- recipient strain *E. coli* CC118. Before mating, both parents were grown at 37 °C, without shaking, to an optical density at 600 nm of about 0.5. Two milliliters of donor cells and one milliliter of recipient cells were mixed and incubated at 37 °C without shaking for 2 h. After addition of 12 ml of LB medium and further incubation with vigorous shaking for 3 h, exconjugants were selected on LB medium containing rifampicin and chloramphenicol. The pAAD derivatives bearing Tn*1000* insertions within the *paa* genes were further analyzed. The primer Tn5900 (5′-AAAAGGGGAACTGAGAGCTC-3′) that hybridized with the δ terminus of transposon Tn*1000* was used to sequence the insertion sites.

Construction of E. coli AF1411 and AF1412 Strains—By means of RP-4 mediated mobilization, the plasmids pAFPA1T and pAFPA2T, which contain mini-Tn*5* hybrid transposons expressing *Pa-lacZ* fusions (Fig. 5), were transferred from *E. coli* S17–1λ*pir* into *E. coli* AF141. Filter matings were performed as described previously (22). Exconjugants containing the *lacZ* translational fusions stably inserted into the chromosome, *E. coli* AF1411 and *E. coli* AF1412, were selected for the transposon marker, kanamycin, on rifampicin-containing LB medium.

Determination of the Transcription Start Sites by Primer Extension—E. coli CC118 cells containing plasmids pAFPA1, pAFPA2, pAFPZ, pAFPX, or pSJ3 (Fig. 5), were grown in minimal medium containing glycerol and PA until the cultures reached an optical density at 600 nm of about 1.0. Total RNA was isolated using the Qiagen RNA/DNA Midi Kit (Qiagen) according to the instructions of the supplier. Primer extension reactions were carried out with the avian myeloblastosis virus reverse transcriptase as described previously (31), using primers LAC-57 (5′-CGATTAAGTTGGGTAACGCCAGGG-3′, which hybridized at 57 nucleotides downstream of the *lacZ* start codon) and PA5–4 (5′-CGGGCATCCAGTCCTGTGGCTCG-3′, which hybridized at 55 nucleotides downstream of the *paaA* start codon). Products were analyzed on 6% polyacrylamide-urea gels.

Resting Cell Reactions—E. coli W14 cells harboring different pAAD::Tn*1000* derivatives were grown in minimal medium containing glycerol and PA to an optical density at 600 nm of about 0.8. Cell cultures were then centrifuged (3,000 × *g,* 10 min at 20 °C), and cells were washed and resuspended in a 0.05 volume of minimal medium. The resting cell reactions were performed in a final volume of 5 ml containing 4.5 ml of M63 minimal medium and 0.5 ml of the cell suspension. The reaction was started by the addition of 1 mM PA, and the tubes were incubated on a rotary shaking platform at a temperature of 30 °C. Samples of 0.5 ml were taken at different times and centrifuged (10,000 × *g*, 5 min) to remove the cells.

To detect the accumulation of intracellular intermediates in PA catabolism, *E. coli* W14 (pAFK5) and *E. coli* W14 (pAAD::Tn*1000* derivatives) cells were grown in LB medium containing 0.2 mM isopropyl-1-thio-β-D-galactopyranoside and minimal medium containing glycerol and PA, respectively. When the cultures reached an optical density at 600 nm of about 1.0, cells were centrifuged, washed as described above, and resuspended in 0.005 volume of minimal medium supplemented with 10 mM glucose. The reaction was started by the addition of 50 μM [1-^{14}C]PA (4 μCi/ml) (Sigma) and incubated for 5 min at 30 °C. The following treatments were performed at 4 °C. Cells were centrifuged, washed with minimal medium containing 10 mM glucose, and resuspended in 0.01 volume of 0.5 M $HClO_4$. After vortexing vigorously for 1 min, the sample was frozen and thawed once and then centrifuged (10,000 × *g*, 5 min). While the supernatant was collected, the sediment was treated again with 0.01 volume of 0.5 M $HClO_4$. Finally, the two supernatants were mixed, filtered through cellulose-triacetate membrane filters (10,000 M_r) (Alltech Associates, Inc.), vacuum-dried, and resuspended in 200 μl of H_2O.

Phenylacetyl-Coenzyme A Ligase Assay—E. coli W14 cells containing different pAAD derivatives were grown overnight in minimal medium containing glycerol and vitamin B_{12} (500 ng/ml) in the presence or absence of PA, and then diluted into fresh medium to an optical density of about 0.1. Growth was resumed at 30 °C until the cultures reached an optical density at 600 nm of about 1.0. Cells were harvested by centrifugation, washed, and resuspended in 0.05 volume of 0.5 M potassium phosphate buffer, pH 8.2, prior to disruption by passage through a

[2] The programs listed are available via the Internet: Protein Analysis Tools (http://EXPASY.HCUGE.CH./www/tools.html); BLASTP, BLASTN, and BLASTX programs (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST); ALIGN and CLUSTAL W programs (http://kiwi.imgen.bcm.tmc.edu:8088/search-launcher/launcher.html); and the *E. coli* data base collection ECDC (http://susi.bio.uni-giessen.de/ecdc.html).

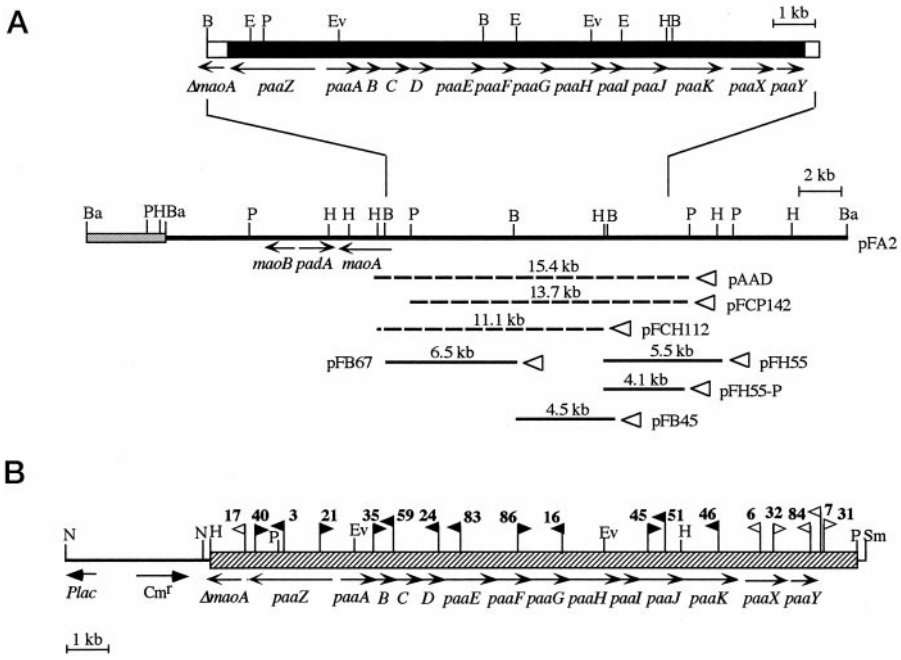


FIG. 1. **Genetic organization of the *paa* genes responsible of the catabolism of PA in *E. coli*, and locations of Tn*1000* insertions.** *A*, physical and genetic map of the chromosomal region containing the *paa* genes, and different cloned DNA fragments. Locations of the genes are shown relative to those of some relevant restriction sites. *Arrows* indicate the direction of gene transcription. Enlarged is shown the *paa* cluster (*black box*). In plasmid pFA2 the nucleotide sequence of the vector (pBR322) is represented by the *dotted box*. Plasmids containing different subcloned DNA fragments (the sizes of the inserts are shown on the *top*) are indicated with *continuous* (pUC derivatives) or *discontinuous* (pCK01 derivatives) *lines*. The *open arrowheads* represent the *Plac* promoter. *B*, locations of transposon Tn*1000* in pAAD-derived mutant plasmids. The 15.4-kb DNA insert in plasmid pAAD is represented with a *striped box*. The γ-δ and δ-γ orientations of Tn*1000* insertions are shown by > and < *symbols*, respectively. Insertions which affect (*filled symbols*) or do not affect (*empty symbols*) catabolic functions are indicated. Δ, a truncated gene. *Cm*^{r}, the gene that confers chloramphenicol resistance. Restriction sites are: *B*, *Bgl*II; *Ba*, *Bam*HI; *E*, *Eco*RI; *Ev*, *Eco*RV; *H*, *Hin*dIII; *N*, *Not*I; *P*, *Pst*I; *Sm*, *Sma*I.

French press (Aminco Corp.) operated at a pressure of 20,000 p.s.i. The cell debris was removed by centrifugation at 26,000 × *g* for 30 min. The clear supernatant fluid was carefully decanted and used as crude extract. Protein concentration was determined by the method of Bradford (32) using bovine serum albumin as standard. Phenylacetyl-coenzyme A ligase was assayed as described previously (33). One unit of enzyme activity is defined as the catalytic activity leading to the formation of 1 nmol of phenylacetylhydroxamate in 1 min at 37 °C.

β-Galactosidase Assay—β-Galactosidase activities were measured with permeabilized cells from cultures grown to mid-log phase, as described by Miller (23).

Analytical Methods—The N-terminal sequence of PaaK was determined by Edman degradation with a model 477A automated protein sequencer (Applied Biosystems Inc.). The protein was directly electroblotted from a SDS-polyacrylamide gel onto a polyvinylidene difluoride membrane as described previously (18).

The metabolites accumulated in the supernatants of resting cells and in culture supernatants, were analyzed with a Gilson HPLC equipment using a Lichrosphere 5 RP-8 column (150 × 4.6 mm) and an isocratic flow of a 40% methanol-H_2O mobile phase pumped at a flow rate of 1 ml/min. Peaks with retention times of 5.6 and 10.6 min, corresponding to those of authentic standard 2-HPA and PA, respectively, were monitored at 220 nm.

The intracellular ^{14}C-labeled metabolites accumulated in resting cell assays were analyzed with the HPLC equipment described above but using an isocratic flow of 0.2 M KH_2PO_4 (pH 4.2), isopropyl alcohol (92:8, v/v) as mobile phase pumped at a flow rate of 1 ml/min. Samples (400 μl) were collected, and ^{14}C-labeled products were detected in aliquots of the fractions by liquid scintillation counting. Peaks with retention times of 11.2 and 20.2 min, corresponding to those of authentic standard PA and PA-CoA, respectively, were monitored at 220 nm.

Thin-layer chromatography was carried out on 0.2-mm Silica gel 60 F_{254} plates (Merck) with chloroform-acetone (2:1, v/v) as solvents. Radiolabeled compounds were visualized by autoradiography.

The 2-HPA was extracted from culture supernatants with an equal volume of ethyl acetate and dried with sodium sulfate. Samples were derivatized with *N, O*-bis(trimethylsilyl)trifluoroacetamide and subjected to gas chromatography-mass spectrometry analysis as described elsewhere (34).

RESULTS

Identification of the paa Genes for the Catabolism of PA—We had recently reported the isolation of an *E. coli* W mutant, strain W14, unable to grow on PA and 2-phenylethylamine as the sole carbon and energy source (18). A recombinant plasmid, pFA2, which contains a 33.3-kb *Bam*HI DNA insert from the chromosome of the wild-type *E. coli* W strain (Fig. 1*A*), had been also constructed and was able to confer to *E. coli* W14 the ability to grow on either PA or phenylethylamine as the sole carbon source (18). When the *Hin*dIII-digested total DNA from *E. coli* W14 and *E. coli* C, a strain also unable to grow on PA as the sole carbon source (2), was analyzed by Southern blot using the 33.3-kb *Bam*HI fragment of pFA2 as probe, no hybridization bands were observed (data not shown). These results indicated that *E. coli* strains W14 and C have a deletion encompassing at least the 33.3-kb DNA fragment cloned in pFA2, and therefore it is likely that they lack the genes involved in PA catabolism (*paa* genes). Moreover, we have observed that, whereas the *E. coli* K-12 strains MV1190, C600, TG1, ET8000, W3110, and MG1655 grew on PA, the K-12 strains HB101, DH1, CC118, DH5α, and JM109 did not grow on this aromatic compound. A Southern blot analysis of the *Hin*dIII-digested total DNA from *E. coli* K-12 PA$^+$ (W3110, MG1655, and C600) and PA$^-$ (DH5α, HB101, and DH1) strains revealed the same pattern of hybridization bands, thus suggesting that the PA$^-$ phenotype in *E. coli* K-12, in contrast to that in *E. coli* W14 and *E. coli* C, could be produced by point mutations or small deletions or insertions. Interestingly, all *E. coli* PA$^-$ strains were able to grow on this aromatic compound when harboring plasmid pFA2, and the loss of this plasmid after several rounds of cultivation of the recombinant strains in the absence of selective pressure was also accompanied by the loss of the PA$^+$ phenotype.

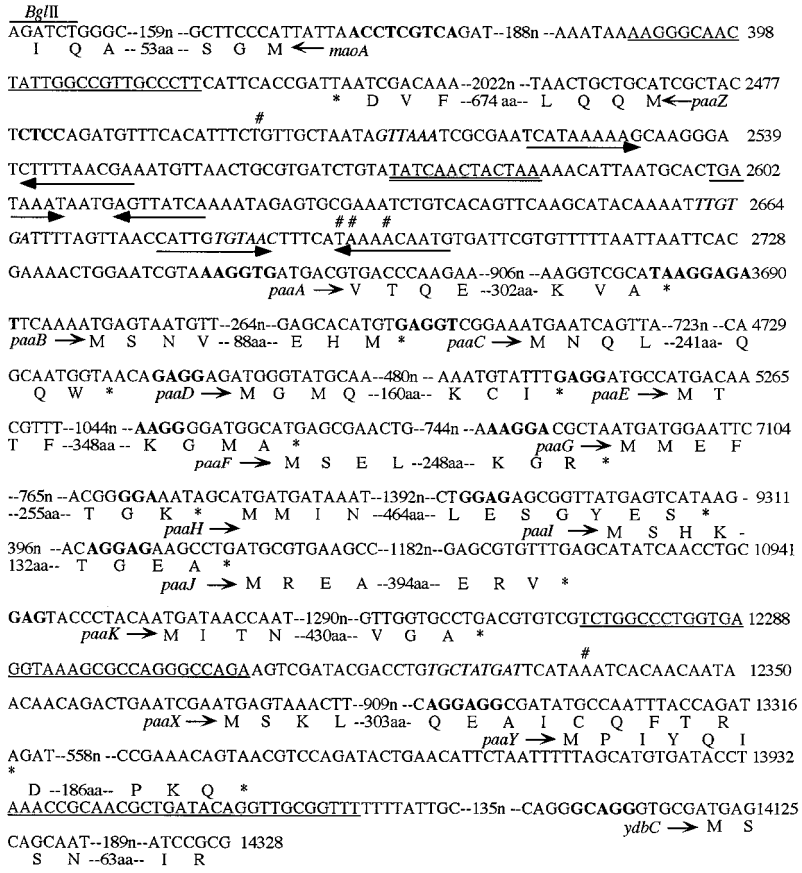


FIG. 2. **Nucleotide and derived amino acid sequences of the PA catabolic pathway.** Only the sequences of the 5′- and 3′-end coding regions of the *paa* genes are shown. The 5′-end coding regions of the *maoA* and *ydbC* genes are also indicated. The *Bgl*II restriction site at the 5′-end of the sequence is shown. Amino acids are represented by their standard *one-letter code*. *Short arrows* show the direction of gene transcription. *Asterisks* indicate stop codons. *Boldface nucleotides* represent potential Shine-Dalgarno sequences. Putative transcriptional terminators are *underlined*. #, a transcription start site. Sequences of potential −10 and −35 boxes are shown in *italics*. Inverted repeats in promoter regions are marked with *convergent arrows* underneath the sequence. A sequence that shows good similarity to the core consensus integration host factor-binding motif (1) is *double-underlined*.

As several aromatic catabolic pathways are encoded by plasmids and *E. coli* W was shown to host cryptic plasmids (35), it was checked whether the *paa* genes were also located in a plasmid. When a pulse field electrophoresis of unrestricted total DNA from *E. coli* W was analyzed by Southern blot using the 33.3-kb DNA fragment as probe, hybridization signals were only found in the sample wells (data not shown), thus indicating a chromosomal location (36) of the *paa* genes.

To localize the *paa* genes within the cloned 33.3-kb DNA fragment, its physical map was determined and different subclones were constructed (Fig. 1*A*) and checked for their ability to restore the growth of *E. coli* W14 on PA-containing minimal medium. Interestingly, although plasmids pFCP142 and pFCH112 (Fig. 1*A*) did not confer to *E. coli* W14 the ability to grow on PA, plasmid pAAD restored the growth of strain W14 on this aromatic compound, thus locating the *paa* genes in a 15.4-kb DNA fragment at the right end of the previously identified *mao* region (Fig. 1*A*) responsible of the transformation of 2-phenylethylamine into PA (18, 37, 38).

Sequencing and Gene Arrangement of the paa Cluster—To genetically characterize the PA catabolic pathway of *E. coli* W, the 15.4-kb insert of plasmid pAAD (Fig. 1*A*) was sequenced. The nucleotide sequence of a 14,328-bp region of this insert is shown in Fig. 2. Computer analysis of this sequence revealed the presence of 14 ORFs, all of which are transcribed in the same direction with the sole exception of *paaZ* (Figs. 1 and 2). Data bases were searched for similar proteins to the *paa* gene products, and those showing the highest similarity values were then retrieved and compared (Table I). The putative Shine-Dalgarno sequences of *paaFGHIJ* and *paaY* overlap the preceding ORFs (Fig. 2), suggesting that the most common mechanism of translational coupling (39) may occur. Downstream of the *paaZ*, *paaK*, and *paaY* genes we found inverted repeat sequences (Fig. 2) predicted to form hairpin loops with ΔG values of −13.4, −25.3, and −15.7 kcal/mol, respectively, which could act as transcriptional terminators. The G+C content of the *paa* cluster averaged 52.5%, a value close to the mean G+C content of *E. coli* genomic DNA (51.5%) (40). At the 3′-end of the sequenced fragment, a partial ORF corresponded to the 5′-end of *ydbC* (Fig. 2), a gene of unknown function from *E. coli* K-12 (29, 41).

To define the essential genes required for PA catabolism in *E. coli* W, transposon mutagenesis of the 15.4-kb *Not*I-DNA cassette in pAAD was carried out. A collection of Tn*1000* insertion derivatives was obtained, and to determine the physical location and orientation of each insertion in pAAD::Tn*1000* mutant plasmids, DNA from each derivative was isolated and analyzed by cleavage with different restriction enzymes. In most of the cases, the sites of the Tn*1000* insertions were sequenced with a primer that hybridized with the δ terminus of transposon Tn*1000*. Twenty different Tn*1000* insertions were identified and checked for their ability to avoid the growth on PA of *E. coli* W14 harboring the corresponding pAAD::Tn*1000* derivatives (Fig. 1*B*). On the basis of these studies, it was shown that only Tn*1000* insertions within genes *paaX* and *paaY* did not affect PA catabolic functions, suggesting that the 12 remaining *paa* genes were essential for the catabolism of this aromatic compound (Fig. 1*B*).

Analysis of PA Pathway Intermediates—To identify possible intermediates of the PA catabolic pathway in *E. coli* W, we first checked the capability of this organism to grow on different aromatic compounds that were thought to be produced during PA degradation, *e.g.* benzoate, phenylpyruvate, *p*-hydroxyphenylpyruvate, mandelate, and phenylglyoxylate. The strain was patched onto PA-containing mineral agar plates to induce the PA pathway, and then replicated onto mineral agar plates containing the aromatic compounds at a concentration of 5 mM. As none of the compounds tested allowed the growth of *E. coli*

TABLE I
PA pathway genes, their products, and related gene products

Gene	Gene product (Da/aa)[a]	Prosite signature	Related gene products				
			Name	Function	Organism	% Identity/aa	Accession no.
paaZ	73,045/681		MaoC	Putative regulator	*K. aerogenes*	78.7 /335 (336–681)[b]	P49251
			NodN	Unknown	*R. leguminosarum*	28.3 /161 (545–681)	P08634
			BadH	Betaine-aldehyde DH[c]	*B. vulgaricus*	24.5 /500 (1–527)	P28237
			XylC	Benzaldehyde DH	*P. putida*	24.2 /487 (1–527)	U15151
paaA	35,437/309						
paaB	10,942/95						
paaC	27,878/248						
paaD	18,606/167						
paaE	39,333/356	2Fe-2S binding region	TdnB	Aniline dioxygenase reductase	*P. putida*	30.1 /337	D85415
			TsaB	Toluenesulfonate monooxygenase reductase	*C. testosteroni*	28.3 /317	U32622
			Pdr	Phathalate dioxygenase reductase	*B. cepacia*	25.6 /321	P33164
paaF	27,164/255	Enoyl-CoA hydratase-isomerase	Echm	Enoyl-CoA hydratase	*R. norvegicus*	41.7 /290	P14604
			BadK	Cyclohexene carboxyl-CoA hydratase	*R. palustris*	40.3 /226	U75363
			Crt	Crotonase	*C. acetobutylicum*	37.8 /261	P52046
			ORF257	Enoyl-CoA hydratase homologue	*R. capsulatus*	36.2 /257	P24162
			CaiD	Carnitine racemase	*E. coli*	31.0 /297	P31551
paaG	28,450/262		ORF257	Enoyl-CoA hydratase homologue	*R. capsulatus*	44.9 /257	P24162
			Crt	Crotonase	*C. acetobutylicum*	31.3 /261	P52046
			BadK	Cyclohexene carboxyl-CoA hydratase	*R. palustris*	34.0 /226	U75363
			Echm	Enoyl-CoA hydratase	*R. norvegicus*	29.7 /290	P14604
paaH	51,704/475	3-hydroxyacyl-CoA DH[c]	Hbd	3-hydroxybutyryl-CoA DH[c]	*C. acetobutylicum*	43.3 /282 (1–289)	M31799
			Schad	3-hydroxyacyl-CoA DH[c]	*H. sapiens*	38.1 /314 (1–289)	X96752
paaI	14,850/140		ComA	Competence protein	*B. subtilis*	25.5 /126	P14205
			YbdB	Unknown	*E. coli*	21.9 /137	P15050
			FcbC	Thioesterase	*Arthrobacter sp.*	20.4 /151	M93187
paaJ	42,276/401	Thiolase	PcaF	β-ketoadipyl-CoA thiolase	*P. putida*	66.1 /400	U10895
			CatF	β-ketoadipyl-CoA thiolase	*A. calcoaceticus*	61.3 /401	M76991
			Thil	Acetyl-CoA acetyltransferase	*C. acetobutylicum*	45.3 /392	U08465
paaK	48,966/437	AMP-binding	FtsA	Coenzyme F390 synthetase	*M. thermoautotrophicum*	24.9 /377	X92347
			FcbA	4-Cl-benzoate-CoA ligase	*Arthrobacter sp.*	20.8 /522	M93187
paaX	35,266/316						
paaY	21,439/196	Transferase	CaiE	Unknown	*E. coli*	56.7 /203	P39206
			Fbp	Ferripyochelin binding protein	*P. aeruginosa*	33.7 /174	P40882

[a] aa, amino acid.
[b] In parentheses is indicated the partial amino acid sequence of the *paa* gene product that has been used for comparison.
[c] DH means dehydrogenase.

W, they appeared not to be intermediates in PA catabolism.

To check whether the different pAAD::Tn*1000* derivatives caused accumulation of PA pathway intermediates, *E. coli* W14 cells harboring these plasmids were grown in minimal medium containing PA and glycerol. Supernatants of the cultures were then analyzed by HPLC, revealing that only the clones containing plasmids with Tn*1000* insertions in genes *paaX*, *paaY*, and *paaZ*, showed PA consumption. However, although disruptions of genes *paaX* and *paaY* did not cause the accumulation of any compound, disruption of the *paaZ* gene caused the conversion of PA into a metabolite whose retention time in HPLC was identical to that of standard 2-HPA. Gas chromatography-mass spectrometry analysis confirmed this metabolite as 2-HPA (data not shown). Similar results were obtained when resting cell assays of the *paaZ* insertion mutants were performed in the presence of 1 mM PA and then analyzed by HPLC (Fig. 3*A*). These data were also in agreement with the observation that plasmid pFCP142, which contains the *paa* cluster with a truncated *paaZ* gene (Fig. 1*A*), conferred to *E. coli* W14 cells the ability to transform PA into 2-HPA.

To analyze whether 2-HPA was an intermediate in PA degradation, *E. coli* W cells were grown in PA-containing minimal medium and then inoculated into minimal medium containing 2-HPA as the sole carbon source; however, this compound did not support bacterial growth. Moreover, *E. coli* W cells grown in minimal medium containing both PA and 2-HPA did not attack the latter compound, as observed by HPLC analysis of the culture supernatants. Therefore, these data suggested that 2-HPA could be a product derived from an unstable intermediate rather than a true intermediate of the PA catabolic pathway. As it has been shown that the *cis*-dihydrodiols formed during the catabolism of different aromatic compounds readily dehydrate nonenzymatically under acidic conditions leading to the corresponding monohydroxy derivatives (2, 42), we checked whether 2-HPA could be also the product of dehydration of a dihydrodiol by monitoring the reported spectral changes associated with such decomposition (42). However, the UV spectrum of the supernatants from *E. coli* W14 cells containing the pAAD::Tn*1000* derivative 3 (Fig. 1*B*) did not show any change under acidic conditions, suggesting that a putative dihydrodiol of PA was not present in the culture medium.

As analyses of culture supernatants did not reveal any true intermediate in PA catabolism, intracellular accumulation of possible metabolites was monitored by thin-layer chromatography of resting cell assays of *E. coli* W14 (pAAD::Tn*1000*) mutants in the presence of radioactive [^{14}C]PA. It was observed that only mutations in the *paaZ* gene caused incorporation of radioactivity into the cells. To identify the radioactive products accumulated, extracts of *E. coli* W14 (pAAD::Tn*1000* derivative 3) were analyzed by HPLC and the detected radioactive peaks were shown to cochromatograph with authentic PA and PA-CoA (data not shown). Furthermore, radioactive 2-HPA was detected by HPLC analysis of the supernatants derived from *E. coli* W14 (pAAD::Tn*1000* derivative 3) in resting cell assays.

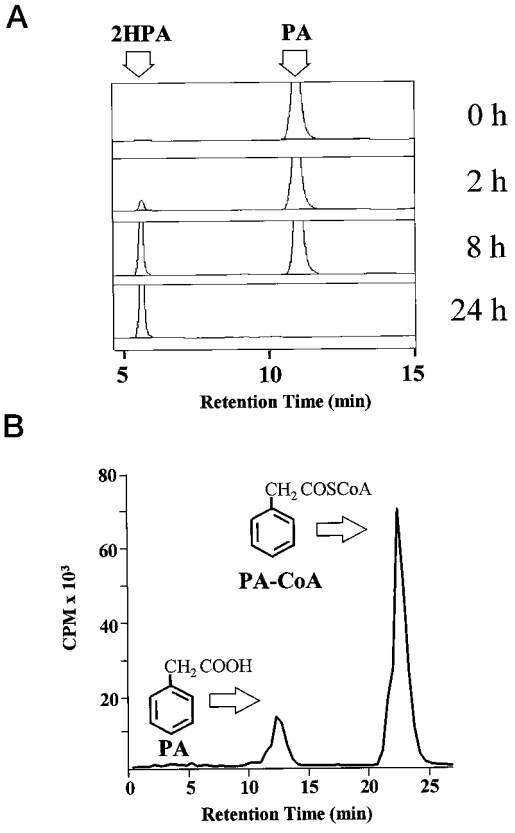


FIG. 3. **HPLC chromatograms of the formation of 2-HPA and PA-CoA from PA in resting cell assays.** *A*, time course of PA consumption and 2-HPA formation in the supernatants of the resting cell assays from *E. coli* W14 (pAAD::Tn*1000* derivative 3). Assays were performed as described under "Experimental Procedures." Samples (20 μl) were retrieved after 0, 2, 8, and 24 h, and analyzed by HPLC as indicated under "Experimental Procedures." *B*, conversion of [^{14}C]PA to [^{14}C]PA-CoA in *E. coli* W14 (pAFK5) resting cell assays. The elution profile of ^{14}C-cellular labeled compounds separated by HPLC is shown. For details, see "Experimental Procedures."

Therefore, all these data taken together suggest that 2-HPA is secreted to the culture medium when the catabolism of PA is blocked, CoA derivatives such as PA-CoA being the true intermediates of the pathway.

The paaK Gene Encodes a PA-CoA Ligase—The formation of PA-CoA during the catabolism of PA in *E. coli* W suggests the existence of a PA-CoA ligase activity involved in the activation of PA to its CoA derivative. As the derived amino acid sequence of the *paaK* gene product revealed a putative AMP-binding consensus motif (Table I), and showed a high identity to that of the PhaE, former Pcl (65.6%), and PaaK_Y2 (67.3%) proteins responsible for the conversion of PA to PA-CoA in *P. putida* U (12) and *Pseudomonas* sp. Y2 (16), respectively, we assumed that PaaK could be the PA-CoA ligase of *E. coli* W. To experimentally demonstrate that *paaK* encoded a PA-CoA ligase, this gene was subcloned in plasmid pUC19 under the control of the *lac* promoter resulting in plasmid pAFK5 (Fig. 4*A*). Crude extracts of *E. coli* W14 (pAFK5) cells grown in glycerol-containing minimal medium showed a high level of PA-CoA ligase activity (450 units/mg protein), and this activity was dependent on the presence of ATP, CoA, and PA. SDS-polyacrylamide gel electrophoresis analysis of crude lysates from these cells revealed the presence of an intense band corresponding to a protein with an apparent molecular mass of 49 kDa (data not shown), which is in good agreement with the predicted molecular mass for the PaaK protein (Table I). The N-terminal sequence, MITNTK, of the overproduced protein corresponded with that deduced from the nucleotide sequence of the *paaK* gene, thus confirming it as the *paaK* gene product and showing that no processing of its N-terminal end occurs.

To demonstrate that the product of the reaction catalyzed by PaaK was PA-CoA, we performed resting cell assays of *E. coli* W14 (pAFK5) in the presence of [^{14}C]PA. Two radioactive peaks that cochromatographed with authentic PA and PA-CoA were observed (Fig. 3*B*). To confirm that the second radioactive peak corresponded to PA-CoA, we performed the hydrolytic removal of the CoA moiety by treatment of the sample with NaOH at 65 °C for 30 min, and, as expected, the resulting ^{14}C-labeled product was shown to elute in HPLC as standard PA.

Analyses of crude extracts of *E. coli* W14 (pAAD::Tn*1000*) mutants grown in minimal medium containing PA and glycerol revealed that insertions of transposon Tn*1000* in genes *paaB*, *paaC, paaD, paaE, paaF, paaG,* and *paaJ* caused a significant decrease in the PA-CoA ligase activity, such activity being below detection limits when the Tn*1000* insertions were located within the *paaK* gene (Table II). These results may indicate that genes *paaABCDEFGHIJK* constitute an operon and that insertions of transposon Tn*1000* within this operon exert strong polar effects on the genes located downstream of the insertion site. In agreement with this hypothesis, Tn*1000* insertions in genes *paaZ, paaX*, and *paaY* did not reduce the PA-CoA ligase activity of the corresponding mutants (Table II).

Functional Organization of the paa Genes—To study the functional organization of the *paa* genes within the PA catabolic pathway, plasmids pAFK3 (*paaK*), pAFZ1 (*paaZ*), pFB67 (*paaZpaaABCDE*), pAFAF1 (*paaABCDE*), and pAFFK (*paaFGHIJKpaaXY*) (Figs. 1 and 4), were constructed. *E. coli* W14 cells harboring these plasmids were grown in minimal medium containing both glycerol and PA, and the supernatants of the cultures were analyzed by HPLC. We could only observe PA consumption when genes *paaABCDE* were expressed simultaneously to the *paaK* gene. Thus, *E. coli* W14 (pAFK5, pAFAF1) cells removed PA with the concomitant accumulation of 2-HPA in the culture medium (data not shown). However, PA remained unaltered in the culture medium when *E. coli* W14 cells containing plasmid pAFK5, pAFAF1, pAFFK, pAFZ1, or pFB67 were analyzed. Therefore, these results indicate that all or some of the *paaABCDE* genes are involved in 2-HPA formation and that this hydroxylation reaction requires the expression of the *paaK* gene responsible of PA-CoA formation. As 2-HPA is not a true intermediate in the PA catabolic pathway, it can be assumed that after the first catabolic step in the PA degradation, *i.e.* activation to PA-CoA, a hydroxylation reaction occurs leading to the formation of a hydroxylated derivative of PA-CoA. A blockade in the PA degradation pathway preventing further catabolism of the hydroxylated PA-CoA derivative could cause the secretion of the latter to the culture medium as 2-HPA.

Although the simultaneous expression of the *paaABCDE* and *paaK* genes gave rise to 2-HPA, the additional presence of the *paaZ* gene did not reveal accumulation of 2-HPA in the supernatants of *E. coli* W14 (pFB67, pAFK3) cultures. Therefore, these data suggest that the *paaZ* gene product could be responsible of the third enzymatic step in the aerobic catabolism of PA in *E. coli* W, genes *paaFGHIJ* being likely involved in further catabolism of PA to Krebs cycle intermediates.

Regulation of the paa Cluster—Analysis of the *paaK* expression showed that the *paa*-encoded pathway was inducible. Thus, although crude extracts of *E. coli* W14 (pAAD) cells grown in glycerol-containing minimal medium in the absence of PA did not reveal PA-CoA ligase activity, a significant level of PaaK activity was observed when the cells were grown in minimal medium containing glycerol and PA (Table III). Interestingly, extracts from *E. coli* W14 (pAAD::Tn*1000* derivatives

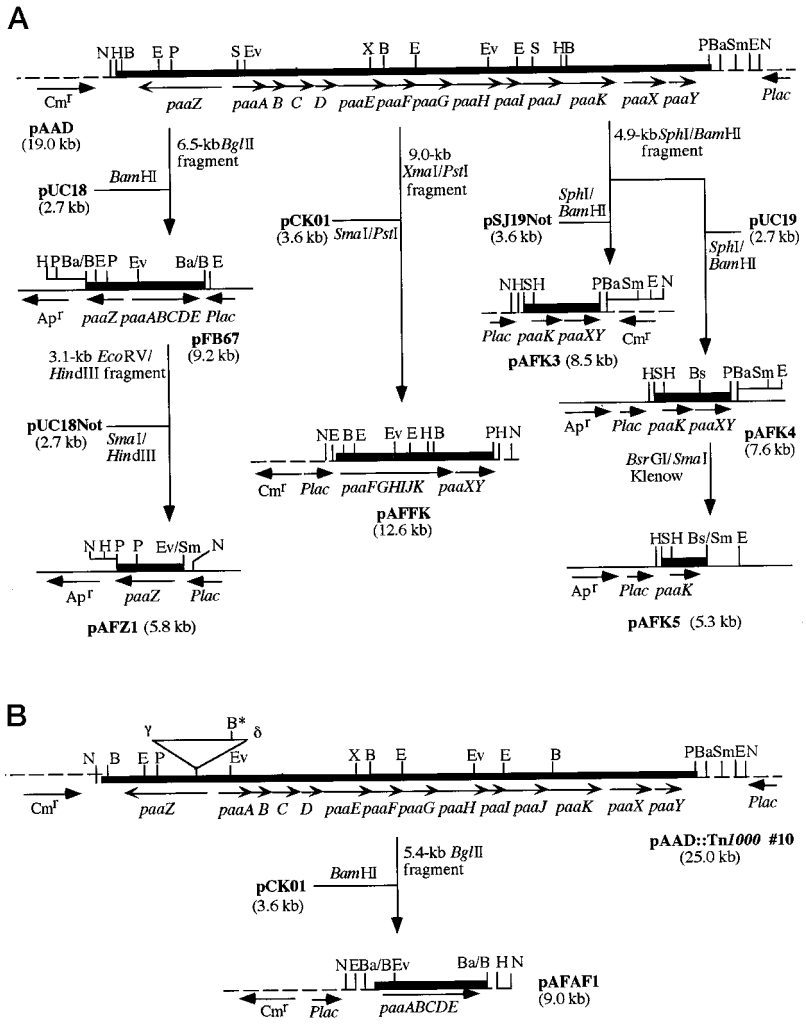


FIG. 4. **Schematic representation of the subcloning and expression of different *paa* genes.** Plasmids are drawn with the relevant elements and restriction sites indicated. DNA containing the *paa* genes from *E. coli* W is represented with a *black box*. Vector-derived sequences are indicated with *continuous* (pUC derivatives) or *discontinuous* (pCK01 derivatives) *lines*. *Arrows* show the *Plac* promoter and the direction of transcription of the genes. Plasmid constructs derived from pAAD (*A*) and pAAD::Tn*1000* derivative 10 (*B*), are shown. γ and δ, the gamma and delta ends of transposon Tn*1000*, respectively. Ap^r and Cm^r, the genes that confer ampicillin and chloramphenicol resistance, respectively. Restriction sites are: *B*, *Bgl*II; *Ba*, *Bam*HI; *Bs*, *Bsr*GI; *E*, *Eco*RI; *Ev*, *Eco*RV; *H*, *Hin*dIII; *N*, *Not*I; *P*, *Pst*I; *S*, *Sph*I; *Sm*, *Sma*I; *X*, *Xba*I. *B**, additional *Bgl*II restriction sites are present in the sequence of transposon Tn*1000*.

6 and 32) cells grown in the presence of PA showed levels of PA-CoA ligase activity similar to those of extracts from the same cells grown in the absence of PA (Table III), thus indicating that Tn*1000* insertions in gene *paaX* cause a constitutive expression of the *paa*-encoded pathway. However, no constitutive expression of the *paaK* gene was observed when *paaY* was disrupted by Tn*1000* insertion (Table III). Therefore, these data suggest that the *paaX* gene product behaves as a negative regulator of the *paa* catabolic genes.

The arrangement of the *paa* genes and the polar effects derived from the Tn*1000* insertions in the *paa* cluster suggest the existence of three transcription units, two of them, *paaZ* and *paaABCDEFGHIJK*, responsible of catabolic functions, and a third one, *paaXY*, involved in regulation. To study the promoter regions of the *paa* cluster, DNA fragments containing the potential *Pz*, *Pa*, and *Px* promoters of genes *paaZ*, *paaABCDEFGHIJK*, and *paaXY*, respectively, were PCR-isolated and ligated to the *lacZ* gene of the promoter-probe vector pSJ3. The resulting translational fusion plasmids pAFPA1 (Pa_{424}-*lacZ*, Pa_{424} being a 424-bp DNA fragment containing *Pa*), pAFPA2 (Pa_{255}-*lacZ*, Pa_{255} being the Pa_{424} fragment lacking 169 bp at its 5′-end), pAFPZ (*Pz-lacZ*), and pAFPX (*Px-lacZ*) (Fig. 5), conferred to the host strain *E. coli* CC118 the ability to produce blue colonies on media containing the β-galactosidase indicator 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside, thus indicating the presence of a functional promoter in the four cloned fragments.

To determine the transcription initiation sites in the *Pa, Pz*, and *Px* promoters, primer extension analyses were performed with total RNA isolated from *E. coli* CC118 cells containing plasmids pAFPA1, pAFPX, and pAFPZ (Fig. 6). The transcription initiation sites of the *paaA* gene were mapped utilizing two different primers (Fig. 6, *A* and *B*), and potential −10 (TGTAAC) and −35 (TTGTGA) boxes typical of σ70-dependent promoters were identified in the *Pa* promoter region (Fig. 2). The same results were obtained when plasmid pAFPA2 (Pa_{255}-*lacZ*) was used as source of RNA for the primer extension analyses (data not shown). The transcription initiation site in the *Pz* promoter (Fig. 6*C*) was located 27 nucleotides upstream of the ATG translation initiation codon of the *paaZ* gene, showing a putative −10 box (TTTAAC) but lacking a consensus −35 sequence (Fig. 2). Analysis of the 194-bp region between the transcription start sites of the *paaZ* and *paaA* genes showed a high A+T content (70%), and revealed several inverted repeat sequences and a putative integration host factor-binding consensus motif (Fig. 2), which might be involved in the control of gene expression. Although these potential regulatory elements are present in Pa_{424}, they are absent in Pa_{255}. Transcription of *paaX* (Fig. 6*D*) starts 29 nucleotides upstream of the ATG translation initiation codon (Fig. 2). The presence in the *Px* promoter of a putative extended −10 box (TGCTATGAT) could explain the absence of a consensus −35 hexamer (Fig. 2). It is worth noting that the putative stem-loop structure that could act as a transcriptional terminator of the *paaABCDEFGHIJK* operon is located only 15 bp upstream from the extended −10 box of the *Px* promoter (Fig. 2).

We have shown above that expression of the *paa* catabolic operon is controlled by the *paaX* gene product; therefore, to

TABLE II
PA-CoA ligase activity of E. coli W14 recombinant strains

Enzymatic activities were assayed using crude extracts of *E. coli* W14 containing different plasmids. Cells were grown in glycerol-containing minimal medium in the presence of 5 mM PA. Enzymatic assays were performed as described under "Experimental Procedures." Results of one experiment are given; values were reproducible in three separate experiments.

E. coli W14 containing plasmid	*paa* genes	PA-CoA ligase activity
		units/mg protein
None	*paa* cluster absent	BD[a]
pAAD	Wild-type *paa* cluster	22
pAAD::Tn*1000* 40	*paaZ*	30
pAAD::Tn*1000* 3	*paaZ*	55
pAAD::Tn*1000* 21	*paaZ*	59
pAAD::Tn*1000* 35	*paaB*	6
pAAD::Tn*1000* 59	*paaC*	5
pAAD::Tn*1000* 24	*paaD*	5
pAAD::Tn*1000* 83	*paaE*	5
pAAD::Tn*1000* 86	*paaF*	5
pAAD::Tn*1000* 16	*paaG*	5
pAAD::Tn*1000* 45	*paaJ*	6
pAAD::Tn*1000* 46	*paaK*	BD
pAAD::Tn*1000* 6	*paaX*	27
pAAD::Tn*1000* 32	*paaX*	20
pAAD::Tn*1000* 84	*paaY*	16

[a] BD, below detection limits.

TABLE III
Analysis of the paaK expression

E. coli W14 cells containing different plasmids were grown in glycerol-containing minimal medium in the absence (uninduced) or presence of 5 mM PA (induced). Preparation of cellular extracts and determination of PA-CoA ligase activity were done as described under "Experimental Procedures." Results of one experiment are given; values were reproducible in three separate experiments.

E. coli W14 plasmid	Tn*1000* insertion	PA-CoA ligase activity	
		Uninduced	Induced with 5 mM PA
		units/mg protein	
pAAD	None	BD[a]	18
pAAD::Tn*1000* 6	*paaX*	20	22
pAAD::Tn*1000* 32	*paaX*	15	20
pAAD::Tn*1000* 84	*paaY*	BD	16

[a] BD, below detection limits.

further analyze faithfully this regulatory system, we have engineered the reporter *Pa-lacZ* fusion within a mini-Tn*5* vector (Fig. 5). The resulting constructions, pAFPA1T (Pa_{424}-*lacZ*) and pAFPA2T (Pa_{255}-*lacZ*) were used to deliver by transposition the corresponding translational fusions into the chromosome of *E. coli* AF141, a rifampicin-resistant *E. coli* W14 (*lacZ*$^-$) mutant strain devoid of β-galactosidase activity, giving rise to the reporter strains *E. coli* AF1411 and AF1412, respectively. The presence of a strong T7 phage transcriptional terminator downstream of the *lacZ* fusions and their orientation within the mini-Tn*5* elements (Fig. 5), prevented read-through transcription from nearby chromosomal promoters after insertion, thus facilitating the regulatory studies. To check the influence of the PaaX protein on the expression of the reporter fusions, *paaX* was cloned in plasmid pAFX2 (Fig. 7) and expressed into the reporter strains. The β-galactosidase assays of permeabilized *E. coli* AF1411 and AF1412 cells harboring control plasmid pCK01 showed a similar and constitutive expression of the reporter fusions (Table IV). When the gene *paaX* was expressed in *trans*, we observed a drastic decrease (more than 2 orders of magnitude) in the β-galactosidase levels of *E. coli* AF1411 (pAFX2) and *E. coli* AF1412 (pAFX2) cells (Table IV), thus indicating that PaaX behaves as a transcriptional repressor of *Pa* both in the Pa_{424} and Pa_{255} promoter fragments.

The repressor effect of PaaX on *Pa* could not be significantly alleviated by growing the cells in the presence of 5 mM PA (Table IV), which indicates that this aromatic compound is not the true inducer of the pathway. Interestingly, when the *paaK* and *paaX* genes were simultaneously expressed in the reporter strains, β-galactosidase activities were shown to be inducible by 5 mM PA (Table IV), suggesting that the reaction product of the PaaK enzyme, *i.e.* PA-CoA, is the inducer of the *Pa* promoter. Furthermore, it is worth noting that Pa_{424} and Pa_{255} responded similarly to the PaaX-mediated regulation, thus suggesting that the 169-bp region that is absent in Pa_{255} is not directly involved in the promoter-operator sites of the *paa* catabolic operon.

DISCUSSION

In this report, we describe the molecular characterization of the PA catabolic pathway of *E. coli*. Previous work had shown that, whereas *E. coli* K-12 and *E. coli* W were able to grow on PA as the sole carbon source, this catabolic ability was lacking in *E. coli* C (2). The molecular analysis presented here confirm the previous observations, indicating that a 33.3-kb DNA fragment that appears to contain the *paa* genes responsible of the PA catabolism in *E. coli* W is lacking in *E. coli* C as well as in the mutant strain *E. coli* W14. However, we have shown here that the ability of *E. coli* K-12 to grow on PA was strain-dependent, with point mutations or small gene rearrangements being the most probable reason for the PA$^-$ phenotype of some K-12 laboratory strains such as DH5α, HB101, and DH1.

The *paa* genes from *E. coli* W were located in a chromosomal 15.4-kb DNA fragment cloned in plasmid pAAD, and they mapped at the right end of the *mao* region (Fig. 1*A*), which is involved in the transformation of 2-phenylethylamine into PA (18, 37, 38). As the equivalent *mao* genes in *E. coli* K-12 have been mapped at min 31.0 on the chromosome (43), and two PA$^-$ mutants of *E. coli* K-12 had been located in this chromosomal region (7), a similar location of the *paa* genes in the chromosome of *E. coli* W can be suggested.

The nucleotide sequence of the *paa* cluster revealed the presence of 14 ORFs, *paaZpaaABCDEFGHIJKpaaXY* (Figs. 1 and 2), that corresponded with those of unknown function whose Protein Identification Database accession numbers are g1787653–g1787664, g1787666, and g1787667, respectively, and that have been recently sequenced in *E. coli* K-12 (accession numbers AE000236, AE000237, D90777, and D90778) (41). Although the left end of the *paa* cluster was near to the *maoA* gene both in *E. coli* W and K-12, the right end of the *paa* cluster was different in the two strains. Thus, although the *paaY* stop codon was found 231 bp upstream of the ATG start codon of the *ydbC* gene in *E. coli* W (Fig. 2), a 9.2-kb sequence encoding a long ORF (*ydbA*) disrupted by two insertion sequences (IS*2* and IS*30*) was found between *paaY* (Protein Identification Database accession number g1787667) and *ydbC* in *E. coli* K-12 (29). The presence of insertion sequences near the *paa* cluster and the location of this cluster in a nonessential region of the chromosome (44) provide some clues on the possible mechanisms of gene mobilization of a catabolic cassette that would enhance bacterial adaptability, and could explain the heterogeneity observed among different *E. coli* strains respect to their ability to mineralize PA. It is also noteworthy that the *mao* genes for the metabolism of 2-phenylethylamine, an aromatic amine whose degradation gives rise to PA, lie adjacent to the *paa* cluster responsible for the further catabolism of PA. This association between genes belonging to the same catabolon (15), *i.e.* genes involved in convergent degradative routes, could be considered as an important evolutionary and adaptive advantage. Another example of such association within a PA catabolon has been recently described in the path-

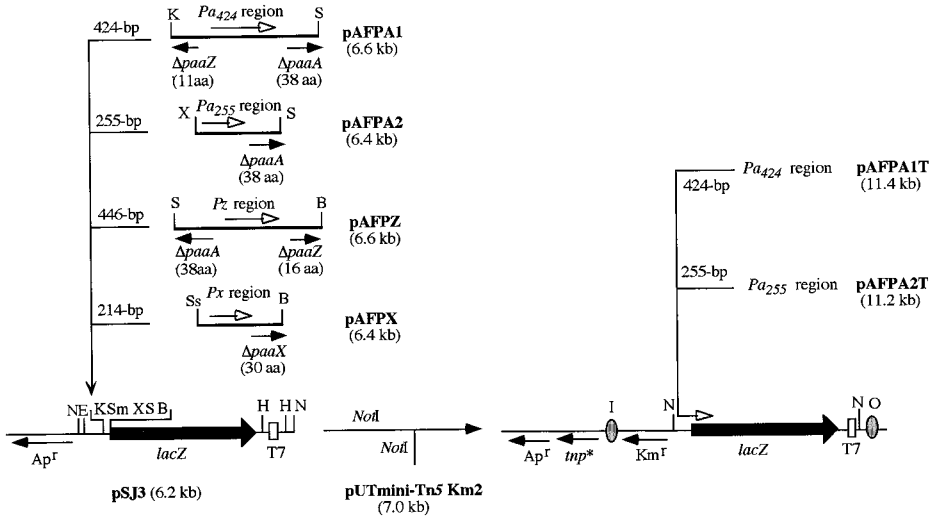


FIG. 5. **Schematic representation of the construction of *lacZ* translational fusion cassettes.** The PCR-amplified promoter regions (detailed under "Experimental Procedures") cloned in the promoter-probe plasmid pSJ3 are shown. The promoterless *lacZ* reporter gene is indicated with a *thick arrow*. The promoters and the direction of transcription of the genes are represented with *white* and *black thin arrows*, respectively. Δ, a truncated gene. The number of amino acid residues fused to the LacZ protein is indicated in *parentheses*. The *white box* represents the early T7 transcriptional terminator (19). The I and O termini of the hybrid mini-Tn*5* transposons are also indicated. Ap^r and Km^r, the genes that confer ampicillin and kanamycin resistance, respectively. *tnp**, gene devoid of *Not*I sites encoding the Tn*5* transposase. Restriction sites are: *B*, *Bam*HI; *E*, *Eco*RI; *H*, *Hin*dIII; *K*, *Kpn*I; *N*, *Not*I; *S*, *Sph*I; *Sm*, *Sma*I; *Ss*, *Ssp*I; *X*, *Xba*I.

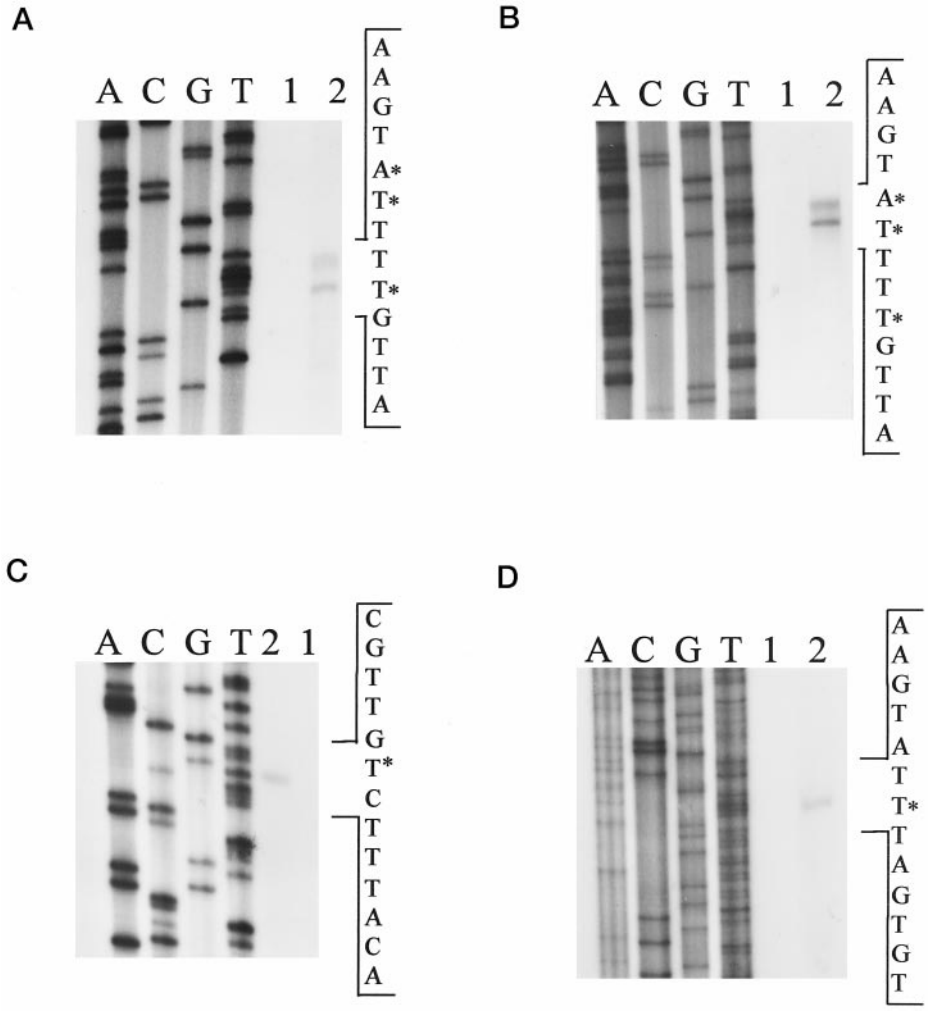


FIG. 6. **Identification of the 5′ transcription start sites in the *Pa*, *Px*, and *Pz* promoters.** Total RNA was isolated from *E. coli* CC118 cells bearing the *lacZ* translational fusion plasmids pAFPA1 (*A* and *B*, *lane 2*), pAFPZ (*C*, *lane 2*), and pAFPX (*D*, *lane 2*), and the control plasmid pSJ3 (*lane 1* in the four panels), as described under "Experimental Procedures." The sizes of the extended products were determined by comparison with the DNA-sequencing ladder of the *Pa* (*A* and *B*), *Pz* (*C*), and *Px* (*D*) promoter regions (A, C, G, and T), using plasmids pAFPA1, pAFPZ, and pAFPX as the templates, respectively. Primer extension and sequencing reactions were performed with primers LAC-57 (*A*, *C*, and *D*), and PA5–4 (*B*), as described under "Experimental Procedures." To the *right* of each panel, an expanded view of the nucleotide sequence surrounding the transcription initiation site(s) (*) is shown. Note that the sequence corresponds to the coding strand (Fig. 2).

way for styrene degradation in *Pseudomonas* sp. Y2, where the *sty* genes responsible of the oxidation of styrene to PA are in tight association with the genes involved in PA degradation (16).[3]

The genetic arrangement of the *paa* cluster and the mutagenesis of pAAD with transposon Tn*1000* revealed that the 14 *paa* genes are organized in three transcriptional units, two of them, *paaZ* and *paaABCDEFGHIJK*, essential for the catabolism of PA, and a third one, *paaXY*, that contains the *paaX* regulatory gene. An overall sequence comparison analysis of the *paa* gene products showed that they were homologous to the recently described *pha* genes responsible of the catabolism of PA in *P. putida* U (15) (Fig. 8*B*). Here, we have presented experimental evidence that the *paaK* gene product is the PA-CoA ligase of *E. coli* W (Fig. 3*B*), an activity that had been

[3] A. Velasco, S. Alonso, J. Perera, J. L. García, and E. Díaz, unpublished data.

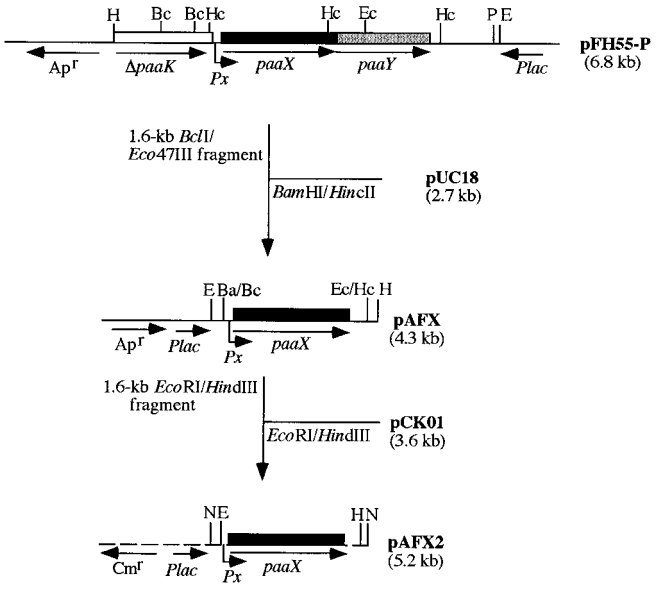


FIG. 7. **Subcloning and expression of the *paaX* regulatory gene.** Plasmids are drawn with the relevant elements and restriction sites indicated. Plasmid pFH55-P has been described in Fig. 1*A*. Vector-derived sequences are indicated with *continuous* (pUC derivatives) or *discontinuous* (pCK01 derivatives) *lines*. *Arrows* show the *Plac* promoter and the direction of transcription of the genes. The *bent arrows* indicate the *Px* promoter. Ap^r and Cm^r, the genes that confer ampicillin and chloramphenicol resistance, respectively. Δ, a truncated gene. Restriction sites are: *Ba*, *Bam*HI; *Bc*, *Bcl*I; *E*, *Eco*RI; *Ec*, *Eco*47III; *H*, *Hin*dIII; *Hc*, *Hin*cII; *N*, *Not*I; *P*, *Pst*I.

detected in this strain when it was grown in PA-containing medium (13). Analysis of the primary structure of PaaK (Fig. 2) revealed that residues ^{103}SSGTTGKPTV^{112} match the AMP-binding site consensus sequence T(SG)-S(G)-G-(ST)-T(SE)-G(S)-*X*-P(M)-K-G(LAF) in acyl-adenylate-forming enzymes (residues that predominate at that position are underlined, with alternates given in parentheses; *X* represents a hypervariable position) (45). It is worth noting that the Lys residue of this signature motif is substituted by Thr in all phenylacetyl-CoA ligases so far sequenced, *i.e.* PaaK, PhaE (12) and PaaK_Y2 (16), an observation that supports recent studies showing that this residue does not assume a major role in ATP binding (45). The sequences ^{236}DIYGLSE^{242} and ^{302}YRTRD^{306} (underlined are the stringently conserved residues) in PaaK also match the conserved motifs II and III that may contribute to the substrates binding sites in acyl-adenylate-forming enzymes (45).

The detection of radiolabeled PA-CoA inside *E. coli* W14 (pAAD::Tn*1000* derivative 3) cells, indicates that disruption of the *paaZ* gene causes a blockade of the PA catabolic pathway leading to the accumulation of this CoA derivative, and confirms the physiological role of PaaK in the catabolism of PA in this microorganism. Assuming that the *paaK* gene product catalyzes the first enzymatic step of the PA catabolic pathway, the polar effects caused by the Tn*1000* insertions within the potential *paa* catabolic operon containing the *paaK* gene at its 3′-end, can explain why pathway intermediates did not accumulate in *E. coli* W14 cells expressing the corresponding pAAD::Tn*1000* derivatives. The degradation of PA in *P. putida* U also appears to require PA-CoA as the first intermediate of the pathway (12), and a similar situation could be inferred in other bacteria that are able to use aerobically PA as the sole carbon source (13, 16). The aerobic catabolism of aromatic compounds via their initial activation to CoA derivatives constitutes an unusual strategy that resembles anaerobic degradation mechanisms (46), and could be a widespread way of PA catabolism in bacteria. The participation CoA ligases in the initial step of the aerobic catabolism of 2-aminobenzoate (47) and benzoate (48) in *Azoarcus evansii* KB740 (formerly *Pseudomonas* sp. KB740), ferulate in *P. putida* (49) and *Pseudomonas fluorescens* (50), and 2-furoic acid in *P. putida* Fu1 (51) has been also reported, and the existence of a CoA ligase has been suggested for the aerobic catabolism of salicylate in *Rhodococcus* sp. strain B4 (52) and thiophen-2-carboxylate (53). Moreover, some dehalogenation mechanisms of aromatic compounds also involve CoA thioester formation in aerobiosis (54). Although the rationale for utilizing such hybrid pathways, *i.e.* aerobic catabolic pathways endowed with typical features of an anaerobic catabolism, is not known, it has been suggested that they could represent a strategy of facultative microorganisms to cope with the fluctuations of oxygen supply (55). In this sense, the existence of a hybrid pathway for the catabolism of PA in *E. coli* could reflect the facultative anaerobe character of this bacterium.

All or some of the *paaABCDE* genes appear to be responsible of the second enzymatic step in the catabolism of PA in *E. coli*. Thus, the expression of *paaK* and *paaABCDE* genes in *E. coli* W14 caused the consumption of PA and the accumulation of 2-HPA in the culture medium. However, 2-HPA appears not to be a true intermediate in the PA catabolic pathway as it does not support growth of *E. coli* W and is not consumed even when *E. coli* W cells are growing also in the presence of PA. Interestingly, a similar lack of growth on 2-HPA and accumulation of this compound after adding PA to some cultures of PA^- mutant strains from *E. coli* K-12 (7) and *P. putida* U (15), has been also observed. Although the possibility that exogenous 2-HPA does not enter the cells cannot be ruled out, the fact that 2-HPA formation requires the simultaneous expression of the *paaK* and *paaABCDE* genes strongly suggests that 2-HPA is not a true intermediate in PA degradation but derives from the accumulation of a hydroxylated PA-CoA intermediate that cannot be further degraded. The excretion to the culture medium of a hydroxylated aromatic compound as a dead-end product derived from the intracellular accumulation of a hydroxylated CoA derivative has been also reported in the hybrid pathway for the catabolism of 2-aminobenzoate (47), and could be a general strategy of the cells to prevent the possible metabolic risk of depletion of the intracellular pool of CoA (15, 56).

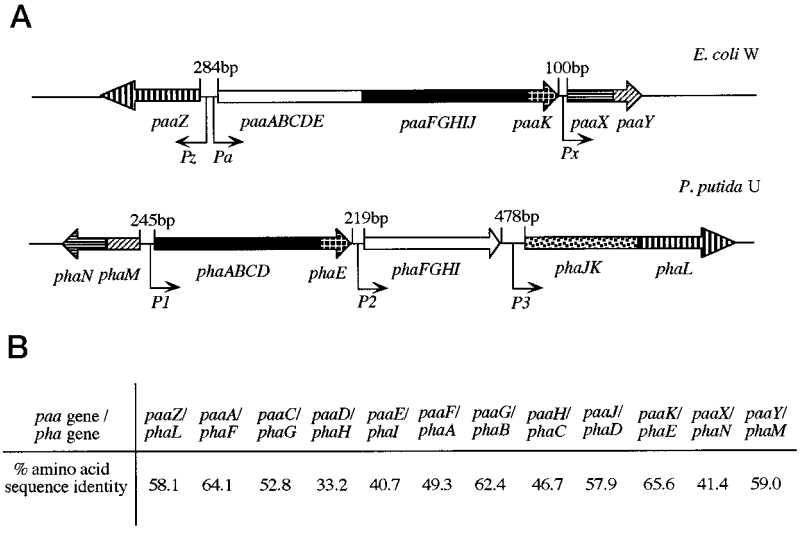
The second catabolic step in PA degradation in *E. coli* seems to be, therefore, the hydroxylation of PA-CoA. Although we could not detect a hydroxylated CoA derivative in *E. coli* W14 (pAAD::Tn*1000* derivative 3) cells, intracellular accumulation of 2-HPA-CoA has been observed during the catabolism of PA by a PA^- *P. putida* U mutant strain (15). Sequence comparison analyses of the *paaABCDE* gene products revealed that the PaaE protein (356-amino acid length) showed significant similarity with the class IA-like reductases (Table I). These enzymes are members of the ferredoxin-$NADP^+$ reductase (FNR) family and they contain a FNR-like domain consisting of a FMN(FAD)- and a NAD(P)-binding region (57). The residues ^{55}RCYS^{58} in PaaE fit the R*X*YS consensus motif for binding of the isoalloxazine ring of the flavin cofactor, and residues ^{121}GS-GITP^{126} and ^{216}CGPAAM^{221} match the G*X*G(X)$_{2-3}$P and CG(X)$_{3-4}$M sequences for the binding of the NAD(P) ribose and NAD(P)-pyrophosphate-nicotinamide moieties of the nicotinamide cofactor, respectively (58). At the C terminus of the FNR-like domain, residues 299–337 in PaaE correspond to the CX_4C*XX*CX_{24-34}C conserved motif of the plant-type ferredoxin [2Fe-2S] binding domain (58). Other members of the extended FNR family are the reductase components of the methane, alkene, phenol, and toluene diiron monooxygenases (59–63), a group of bacterial hydrocarbon oxidation enzymes that com-

TABLE IV
Regulation of expression from the Pa promoter by PaaX

E. coli strains were grown in glycerol-containing minimal medium in the absence (uninduced) or in the presence of 5 mM PA (induced). β-Galactosidase activities were measured with permeabilized cells as described under "Experimental Procedures." Results of one experiment are shown; values were reproducible in three separate experiments.

E. coli strain	Plasmid(s)/gene(s)	Relevant background	β-Galactosidase activity (Miller units)		
			Uninduced	Induced (5 mM PA)	-Fold induction
AF141	None	*lacZ*⁻	BD[a]	BD	
AF1411	pCK01	Pa_{424}::*lacZ*	3111	4156	1.3
AF1411	pAFX2/*paaX*	Pa_{424}::*lacZ*	10	46	4.6
AF1411	pAFX2/*paaX* pAFK5/*paaK*	Pa_{424}::*lacZ*	12	1114	93
AF1412	pCK01	Pa_{255}::*lacZ*	3441	3877	1.1
AF1412	pAFX2/*paaX*	Pa_{255}::*lacZ*	16	59	3.6
AF1412	pAFX2/*paaX* pAFK5/*paaK*	Pa_{255}::*lacZ*	25	1737	69

[a] BD, below detection limits.



B

paa gene / *pha* gene	*paaZ*/ *phaL*	*paaA*/ *phaF*	*paaC*/ *phaG*	*paaD*/ *phaH*	*paaE*/ *phaI*	*paaF*/ *phaA*	*paaG*/ *phaB*	*paaH*/ *phaC*	*paaJ*/ *phaD*	*paaK*/ *phaE*	*paaX*/ *phaN*	*paaY*/ *phaM*
% amino acid sequence identity	58.1	64.1	52.8	33.2	40.7	49.3	62.4	46.7	57.9	65.6	41.4	59.0

FIG. 8. **Comparison of the *paa* cluster of *E. coli* W with the *pha* cluster of *P. putida* U.** *A*, comparison of the genetic organization of the *paa* and *pha* clusters. *Blocks* with similar *shading* or *hatching* indicate homologous regions encoding potential functional units in both gene clusters. The location and size of the intergenic regions, are also indicated. *Bent arrows* represent the promoters. *B*, percentages of amino acid sequence identity between the analogous *paa* and *pha* gene products. Note that genes *phaJK* do not have counterparts in the *paa* cluster, and that genes *paaB* and *paaI* have not been described in the *pha* cluster.

prises an evolutionarily related protein family (60). These soluble multicomponent monooxygenases contain, in addition to the reductase component, a heteromultimeric ($\alpha\beta\gamma$) oxygenase component, a low molecular weight activator protein (61–63), and, in some cases, a Rieske-type ferredoxin (59, 60). Interestingly, the primary structure of the PaaA protein (309-amino acid length) shows the two repeats of residues EX_2H separated by approximately 100 amino acids (positions 155–158 and 249–252) that characterize the dinuclear iron binding-site of the large (α) oxygenase subunit of the methane, phenol, and toluene diiron monooxygenases (60). Moreover, the amino acid sequence of PaaB (95-amino acid length) shows the strictly conserved residues found in the low molecular weight dissociable activator protein that is required for optimal turnover of the oxygenase component in multicomponent diiron monooxygenases (64). Therefore, these sequence comparison analyses suggest that genes *paaABCDE* may encode the five subunits of a diiron multicomponent oxygenase, with PaaB being the effector protein and PaaE the reductase that mediates electron transfer between NAD(P)H and the PaaACD oxygenase component. It is worth noting that the *paaE* gene product can constitute the first example of a reductase subunit from a multicomponent oxygenase that shows a reversed domain order, *i.e.* a FNR-like N-terminal domain and a plant-type ferredoxin C-terminal domain, which supports the previous hypothesis that class IA-like reductases may have been recruited for a variety of aromatic ring oxidation reactions (65). Moreover, the putative PaaABCDE oxygenase, and its counterpart encoded by the *phaFGHI* operon of *P. putida* U (Fig. 8*B*), may represent the first reported multicomponent oxygenase acting on a CoA-activated aromatic acid.

The *paaZ* gene product appears to be responsible of the third enzymatic step of the PA catabolic pathway. The putative PaaZ protein (681-amino acid length) presents an N-terminal region (residues 1–527) whose primary structure shows similarity with that of aldehyde dehydrogenases (Table I). In this sense, the PaaZ residues ^{229}FTGSAATG^{236} and ^{291}GQKCTAIR^{298}, respectively, match the consensus NAD(P)$^+$-binding site and the active site motif spanning the catalytic cysteine (underlined) of aldehyde dehydrogenases (18, 66). Moreover, the sequence ^{254}*M*EADSLN^{260} in PaaZ encompasses the potential catalytic glutamic acid residue (italicized) of aldehyde dehydrogenases (18, 66). The amino acid sequence of the C-terminal region of PaaZ shows similarity to that of the *maoC* and *nodN* gene products of unknown function (Table I). As has been suggested for the analogous PhaL protein of *P. putida* U (15) (Fig. 8*B*), the *paaZ* gene product in *E. coli* might catalyze the aromatic ring cleavage of the hydroxylated CoA derivative formed during PA degradation. Nevertheless, the formation by PaaZ of a non-aromatic CoA cyclic intermediate, similar to that described as the product of the reaction catalyzed by the aminobenzoyl-CoA monooxygenase-reductase during the aerobic catabolism of 2-aminobenzoate (67), cannot be ruled out.

The *paaF*, *paaG*, *paaH*, and *paaJ* gene products show significant sequence similarities to fatty acid β-oxidation enzymes (Table I), and therefore can tentatively constitute a β-oxidation-like pathway involved in the successive oxidation reactions of the non-aromatic CoA intermediate. Interestingly, a

β-oxidation-like mechanism is another typical feature of the anaerobic catabolism of aromatic compounds (46). The primary structure of the putative PaaF (255-amino acid length) and PaaG (262-amino acid length) proteins shows similarity with that of members of the enoyl-CoA hydratase/isomerase superfamily (54, 68) (Table I). The *paaH* gene encodes a protein (475-amino acid length) that shares the signature sequence motives of 3-hydroxyacyl-CoA dehydrogenases (69) (Table I), thus suggesting that it could attack the product of the reaction catalyzed by the PaaF and PaaG enzymes. Although the *paaI* gene product (140-amino acid length) did not show a high level of sequence similarity with other proteins in the data bases, the *paaJ* gene product (401-amino acid length) presented a significant sequence similarity with the PcaF and CatF β-ketoadipyl-CoA thiolases (Table I), residues 90 and 386 in PaaJ being the putative catalytic cysteines. As PcaF and CatF catalyze the last step in the *ortho*-cleavage pathway for the aerobic degradation of protocatechuate and catechol, respectively (70), it is tempting to speculate that PaaJ and its analogous PhaD protein in *P. putida* U (Fig. 8) are also responsible for the last enzymatic step in PA degradation.

In the *paa* cluster, we have identified three promoters, *Pz, Pa*, and *Px*, which drive the expression of genes *paaZ*, *paaABCDEFGHIJK*, and *paaXY*, respectively (Figs. 6 and 8*A*). The expresion of the *paa*-encoded catabolic pathway is inducible, and it has been shown that the *Pa* promoter is negatively controlled by the *paaX* gene product (Table III and IV). The PaaX protein (316-amino acid length) contains a stretch of 25 residues at amino acids 39–64 that shares similarity with the helix-turn-helix motif predicted to be important for DNA recognition and binding in other transcriptional repressors such as GntR (71) and FadR (72). The GntR and FadR binding sites within the respective promoters contain a region of dyad symmetry, which is located very close to the transcription initiation sites (71–73). Interestingly, a region of dyad symmetry can also be found centered near the transcription initiation sites in the *Pa* promoter (Fig. 2). As the repression caused by PaaX was only alleviated by PA in the presence of the PaaK protein (Table IV), PA-CoA appears to be the true inducer of the *paa*-encoded pathway. In this sense, gel retardation assays have confirmed PA-CoA as the effector molecule.[4] Therefore, PaaX constitutes the first reported transcription factor regulated by CoA derivatives that controls the catabolism of aromatic compounds. It is worth noting that the FadR transcriptional repressor, which is regulated by acyl-CoA compounds and shows local similarity to PaaX, is also controlling the expression of genes involved in β-oxidation mechanisms (73).

Overlapping the 3′-end of *paaX*, we have found the putative ATG translation initiation codon of the *paaY* gene. A palindromic sequence (ΔG value of −15.7 kcal/mol) followed by a $(T)_7$ tract is located 42 bp downstream of the TAA stop codon of *paaY* (Fig. 2), and may act as a ρ-independent transcription terminator of the putative *paaXY* operon. Although the primary structure of the *paaY* gene product (196-amino acid length) and its analogous PhaM protein from *P. putida* U (Fig. 8*B*) show several repeats of the hexapeptide (LIV)GX_4 motif that characterizes the members of the bacterial transferases family, *e.g.* the CaiE protein from the carnitine operon of *E. coli* and the Fbp ferripyochelin-binding protein of *P. aeruginosa* (Table I), the physiological role of these proteins in PA catabolism is still unknown.

Comparative studies of the whole structure and organization of the *paa* and *pha* clusters from *E. coli* and *P. putida* U, respectively (Fig. 8*A*), revealed interesting functional and evolutionary data. Thus, although the *pha* genes appear to be cotranscribed in four discrete DNA segments or modules encoding the six different functional units for the catabolism of PA, *i.e.* the β-oxidation and activation (*phaABCDE*), hydroxylation (*phaFGHI*), transport and dearomatization (*phaJKL*), and regulation (*phaMN*) units, the *paa* cluster showed the transcriptional coupling of the hydroxylation-β-oxidation-activation functional units into the single operon *paaABCDEFGHIJK* (Fig. 8*A*). As there is good evidence that operons coding for the catabolism of aromatic compounds are assembled in a stepwise manner from existing catabolic genes (74), it is tempting to speculate that the *paa* cluster from *E. coli* arose by the fusion of some gene blocks that are contiguous but separately regulated in the *pha* cluster of *P. putida* U, and therefore it could be considered as a further step in the evolution toward a single regulon of a common ancestral gene cluster involved in PA catabolism. Moreover, the differences in gene order within some of the DNA modules, and the relative locations of these modules in the *paa* and *pha* clusters, suggest that various DNA rearrangements have occurred during their evolution. As the G+C content of the *paa* (52.5%) and *pha* (63.5%) genes averaged a value close to the mean G+C content of *E. coli* (51.5%) and *P. putida* (60%) genomic DNA (40), it could be thought that these two set of genes have been imprisoned within each host over a long period of evolution. Especially remarkable is the observation that the *phaJ* and *phaK* genes of *P. putida* U, encoding a permease and a specific-channel-forming protein for the uptake of PA, respectively (15), are absent in the *paa* cluster from *E. coli* W (Fig. 8*A*). Interestingly, the *phaJ* gene product shows significant amino acid sequence identity (62.1%) with the product of the *yjcG* gene that is located at min 92.2 of the *E. coli* K-12 chromosome (29). Whether a permease, such as the putative YjcG protein, and a channel-forming protein are required for the catabolism of PA in *E. coli* is still an open question.

The identification and genetic characterization of the hybrid *paa*-encoded pathway complete our knowledge on the pathways so far described for the aerobic catabolism of aromatic compounds in *E. coli*. Although, in some *Pseudomonas* and *Acinetobacter* species, a supraoperonic clustering of the aromatic catabolic genes has been observed in a limited region of the chromosome, the aromatic catabolic clusters are dispersed throughout the genome in *E. coli*, with cluster *mhp* (3-(3-hydroxyphenyl)propionate and 3-hydroxycinnamate) at min 8 (1, 10), *paa* at min 31, *hca* (3-phenylpropionate) at min 57.5 (11), and *hpa* (3- or 4-hydroxyphenylacetate) at min 98 (8). These data also indicate that *E. coli* is not an "empty box" for the catabolism of aromatic compounds; on the contrary, it is endowed with typical aerobic degradation routes as well as with a novel hybrid pathway, which are considered among the most ubiquitous aromatic compound catabolic systems and therefore are thought to be closer to the central catabolism than those involved in the degradation of xenobiotic compounds (75).

The results presented in this work provide a framework for additional studies to determine the role and properties of the enzymes involved in PA catabolism through a hybrid aerobic pathway that is likely to be a widespread route for the metabolism of this aromatic compound. In this sense, the cloned *paa* genes should be useful as probes to identify homologous genes from distinct groups of bacteria. Moreover, we anticipate that the unique features of the aerobic *paa*-encoded pathway will reveal novel catabolic activities that can be of great biotechnological interest to improve some microorganisms for the degradation of PA-related aromatic environmental pollutants (*e.g.*, styrene), and for the synthesis of pathway intermediates that

[4] A. Ferrández, J. L. García, and E. Díaz, manuscript in preparation.

can be useful for the production of new or modified antibiotics and plastics (15).

Acknowledgments—We thank M. K. B. Berlyn for strains MG1063 and MG1655, S. Jaenecke for plasmids pSJ3 and pSJ19Not, A. Prieto and J. Varela for assistance with the gas chromatography-mass spectrometry and N-terminal amino acid sequence analyses, respectively, and A. Díaz and G. Porras for their help with the sequencing. The excellent technical assistance of E. Cano, M. Carrasco, and F. Morante is gratefully acknowledged. We are indebted to K. N. Timmis and D. Pieper for facilitating the short visit of A. Ferrández to their laboratory.

REFERENCES

1. Ferrández, A., García, J. L., and Díaz, E. (1997) *J. Bacteriol.* **179,** 2573–2581
2. Burlingame, R., and Chapman, P. J. (1983) *J. Bacteriol.* **155,** 113–121
3. Mohamed, M., and Fuchs, G. (1993) *Arch. Microbiol.* **159,** 554–562
4. Savage, D. C. (1977) *Annu. Rev. Microbiol.* **31,** 107–133
5. Cooper, R. A., and Skinner, M. A. (1980) *J. Bacteriol.* **143,** 302–306
6. Burlingame, R. P., Wyman, L., and Chapman, P. J. (1986) *J. Bacteriol.* **168,** 55–64
7. Cooper, R. A., Jones, D. C. N., and Parrott, S. (1985) *J. Gen. Microbiol.* **131,** 2753–2757
8. Prieto, M. A., Díaz, E., and García, J. L. (1996) *J. Bacteriol.* **178,** 111–120
9. Stringfellow, J. M., Turpin, B., and Cooper, R. A. (1995) *Gene* (*Amst.*) **166,** 73–76
10. Spence, E. L., Kawamukai, M., Sanvoisin, J., Braven, H., and Bugg, T. D. H. (1996) *J. Bacteriol.* **178,** 5249–5256
11. Díaz, E., Ferrández, A., and García, J. L. (1998) *J. Bacteriol.* **180,** 2915–2923
12. Miñambres, B., Martínez-Blanco, H., Olivera, E. R., García, B., Díez, B., Barredo, J. L., Moreno, M. A., Schleissner, C., Salto, F., and Luengo, J. M. (1996) *J. Biol. Chem.* **271,** 33531–33538
13. Vitovski, S. (1993) *FEMS Microbiol. Lett.* **108,** 1–6
14. Olivera, E. R., Reglero, A., Martínez-Blanco, H., Fernández-Medarde, A., Moreno, M. A., and Luengo, J. M. (1994) *Eur. J. Biochem.* **221,** 375–381
15. Olivera, E. R., Miñambres, B., García, B., Muñiz, C., Moreno, M. A., Ferrández, A., Díaz, E., García, J. L., and Luengo, J. M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 6419–6424
16. Velasco, A., Alonso, S., García, J. L., Perera, J., and Díaz, E. (1998) *J. Bacteriol.* **180,** 1063–1071
17. Davis, B. D., and Mingioli, E. S. (1950) *J. Bacteriol.* **60,** 17–28
18. Ferrández, A., Prieto, M. A., García, J. L., and Díaz, E. (1997) *FEBS Lett.* **406,** 23–27
19. Jaenecke, S., de Lorenzo, V., Timmis, K. N., and Díaz, E. (1996) *Mol. Microbiol.* **21,** 293–300
20. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
21. McNeil, D. (1981) *J. Bacteriol.* **146,** 260–268
22. de Lorenzo, V., and Timmis, K. N. (1994) *Methods Enzymol.* **235,** 386–405
23. Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
24. Arrecubieta, C., López, R., and García, E. (1994) *J. Bacteriol.* **176,** 6375–6383
25. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467
26. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
27. Wilbur, W. J., and Lipman, D. J. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80,** 726–730
28. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680
29. Kröger, M., and Wahl, R. (1997) *Nucleic Acids Res.* **25,** 39–42
30. Guyer, M. S. (1983) *Methods Enzymol.* **101,** 362–369
31. Prieto, M. A., and García, J. L. (1997) *Biochem. Biophys. Res. Commun.* **232,** 759–765
32. Bradford, M. M. (1976) *Anal. Biochem.* **72,** 248–254
33. Martínez-Blanco, H., Reglero, A., Rodríguez-Aparicio, L. B., and Luengo, J. M. (1990) *J. Biol. Chem.* **265,** 7084–7090
34. Ahrazem, O., Prieto, A., Leal, J. A., Gómez-Miranda, B., Domenech, J., Jiménez-Barbero, J., and Bernabé, M. (1997) *Carbohydr. Res.* **303,** 67–72
35. Prieto, M. A., Pérez-Aranda, A., and García, J. L. (1993) *J. Bacteriol.* **175,** 2162–2167
36. Cheng, H.-P., and Lessie, T. G. (1994) *J. Bacteriol.* **176,** 4034–4042
37. Yamashita, M., Azakami, H., Yokoro, N., Roh, J. H., Suzuki, H., Kumagai, H., and Murooka, Y. (1996) *J. Bacteriol.* **178,** 2941–2947
38. Hanlon, S. P., Hill, T. K., Flavell, M. A., Stringfellow, J. M., and Cooper, R. A. (1997) *Microbiology* **143,** 513–518
39. Gold, L.(1988) *Annu. Rev. Biochem.* **57,** 199–233
40. Nakamura, Y., Gojobori, T., and Ikemura, T. (1997) *Nucleic Acids Res.* **25,** 244–245
41. Blattner, F. R., Plunkett III, G., Bloch, C. A., Perna, N. T., Burland V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science* **277,** 1453–1462
42. DeFrank, J. J., and Ribbons, D. W. (1977) *J. Bacteriol.* **129,** 1356–1364
43. Sugino, H., Sasaki, M., Azakami, H., Yamashita, M., and Murooka, Y. (1992) *J. Bacteriol.* **174,** 2485–2492
44. Henson, J. M., Kopp, B., and Kuempel, P. L. (1984) *Mol. Gen. Genet.* **193,** 263–268
45. Chang, K.-H., Xiang, H., and Dunaway-Mariano, D. (1997) *Biochemistry* **36,** 15650–15659
46. Egland, P. G., Pelletier, D. A., Dispensa, M., Gibson, J., and Harwood, C. S. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6484–6489
47. Altenschmidt, U., and Fuchs, G. (1992) *Eur. J. Biochem.* **205,** 721–727
48. Altenschmidt, U., Oswal, B., Steiner, E., Herrmann, H., and Fuchs, G. (1993) *J. Bacteriol.* **175,** 4851–4858
49. Zenk, M. H., Ulbrich, B., Busse, J., and Stöckigt, J. (1980) *Anal. Biochem.* **101,** 182–187
50. Gasson, M. J., Kitamura, Y., McLauchlan, W. R., Narbad, A., Parr, A. J., Parsons, E. L. H., Payne, J., Rhodes, M. J. C., and Walton, N. J. (1998) *J. Biol. Chem.* **273,** 4163–4170
51. Koenig, K., and Andreesen, J. R. (1989) *Appl. Environ. Microbiol.* **55,** 1829–1834
52. Grund, E., Denecke, B., and Eichenlaub, R. (1992) *Appl. Environ. Microbiol.* **58,** 1874–1877
53. Cripps, R. E. (1973) *Biochem. J.* **134,** 353–366
54. Dunaway-Mariano, D., and Babbitt, P. C. (1994) *Biodegradation* **5,** 259–276
55. Niemetz, R., Altenschmidt, U., Brucker, S., and Fuchs, G. (1995) *Eur. J. Biochem.* **227,** 161–168
56. Chohnan, S., Furukawa, H., Fujio, T., Nishihara, H., and Takamura, Y. (1997) *Appl. Environ. Microbiol.* **63,** 553–560
57. Correl, C. C., Batie, C. J., Ballou, D. P., and Ludwig, M. L. (1992) *Science* **258,** 1604–1610
58. Rosche, B., Tshisuaka, B., Hauer, B., Lingens, F., and Fetzner, S. (1997) *J. Bacteriol.* **179,** 3549–3554
59. Small, F. J., and Ensign, S. A. (1997) *J. Biol. Chem.* **272,** 24913–24920
60. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Münck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) *Biochemistry* **35,** 9106–9119
61. Johnson, G. R., and Olsen, R. H. (1995) *Appl. Environ. Microbiol.* **61,** 3336–3346
62. Powlowski, J., Sealy, J., Shingler, V., and Cadieux, E. (1997) *J. Biol. Chem.* **272,** 945–951
63. Lipscomb, J. D. (1994) *Annu. Rev. Microbiol.* **48,** 371–399
64. Qian, H., Edlund, U., Powlowski, J., Shingler, V., and Sethson, I. (1997) *Biochemistry* **36,** 495–504
65. Nakatsu, C. H., Straus, N. A., and Wyndham, R. C. (1995) *Microbiology* **141,** 485–495
66. Liu, Z.-J., Sun, Y.-J., Rose, J., Chung, Y.-J., Hsiao, C.-D., Chang, W.-R., Kuo, I., Perozich, J., Lindahl, R., Hempel, J., and Wang, B.-C. (1997) *Nat. Struct. Biol.* **4,** 317–326
67. Langkau, B., and Ghisla, S. (1995) *Eur. J. Biochem.* **230,** 686–697
68. Müller-Newen, G., Janssen, U., and Stoffel, W. (1995) *Eur. J. Biochem.* **228,** 68–73
69. He, X.-Y., Deng, H., and Yang, S.-Y. (1997) *Biochemistry* **36,** 261–268
70. Harwood, C. S., and Parales, R. E. (1996) *Annu. Rev. Microbiol.* **50,** 553–590
71. Fujita, Y., and Miwa, Y. (1989) *J. Biol. Chem.* **264,** 4201–4206
72. DiRusso, C. C., Heimert, T. L., and Metzger, A. K. (1992) *J. Biol. Chem.* **267,** 8685–8691
73. Black, P. N., and DiRusso, C. C. (1994) *Biochim. Biophys. Acta* **1210,** 123–145
74. van der Meer, J. R., de Vos, W. M., Harayama, S., and Zehnder, A. J. B. (1992) *Microbiol. Rev.* **56,** 677–694
75. Barnes, M. R., Duetz, W. A., and Williams, P. A. (1997) *J. Bacteriol.* **179,** 6145–6153

Catabolism of Phenylacetic Acid in *Escherichia coli* : CHARACTERIZATION OF A NEW AEROBIC HYBRID PATHWAY

Abel Ferrández, Baltasar Miñambres, Belén Garci?a, Eli?as R. Olivera, José M. Luengo, José L. Garci?a and Eduardo Di?az

J. Biol. Chem. 1998, 273:25974-25986.
doi: 10.1074/jbc.273.40.25974