JOURNAL OF BACTERIOLOGY, May 1997, p. 3171–3180 Vol. 179, No. 10
0021-9193/97/$04.00+0

p-Cymene Catabolic Pathway in *Pseudomonas putida* F1: Cloning and Characterization of DNA Encoding Conversion of *p*-Cymene to *p*-Cumate†

RICHARD W. EATON*

Gulf Ecology Division, National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Gulf Breeze, Florida 32561

Received 10 January 1997/Accepted 13 March 1997

***Pseudomonas putida* F1 utilizes *p*-cymene (*p*-isopropyltoluene) by an 11-step pathway through *p*-cumate (*p*-isopropylbenzoate) to isobutyrate, pyruvate, and acetyl coenzyme A. The *cym* operon, encoding the conversion of *p*-cymene to *p*-cumate, is located just upstream of the *cmt* operon, which encodes the further catabolism of *p*-cumate and is located, in turn, upstream of the *tod* (toluene catabolism) operon in *P. putida* F1. The sequences of an 11,236-bp DNA segment carrying the *cym* operon and a 915-bp DNA segment completing the sequence of the 2,673-bp DNA segment separating the *cmt* and *tod* operons have been determined and are discussed here. The *cym* operon contains six genes in the order *cymBCAaAbDE*. The gene products have been identified both by functional assays and by comparing deduced amino acid sequences to published sequences. Thus, *cymAa* and *cymAb* encode the two components of *p*-cymene monooxygenase, a hydroxylase and a reductase, respectively; *cymB* encodes *p*-cumic alcohol dehydrogenase; *cymC* encodes *p*-cumic aldehyde dehydrogenase; *cymD* encodes a putative outer membrane protein related to gene products of other aromatic hydrocarbon catabolic operons, but having an unknown function in *p*-cymene catabolism; and *cymE* encodes an acetyl coenzyme A synthetase whose role in this pathway is also unknown. Upstream of the *cym* operon is a regulatory gene, *cymR*. By using recombinant bacteria carrying either the operator-promoter region of the *cym* operon or the *cmt* operon upstream of genes encoding readily assayed enzymes, in the presence or absence of *cymR*, it was demonstrated that *cymR* encodes a repressor which controls expression of both the *cym* and *cmt* operons and is inducible by *p*-cumate but not *p*-cymene. Short (less than 350 bp) homologous DNA segments that are located upstream of *cymR* and between the *cmt* and *tod* operons may have been involved in recombination events that led to the current arrangement of *cym*, *cmt*, and *tod* genes in *P. putida* F1.**

While many alkyl-substituted aromatic hydrocarbons occur as components of fossil fuels such as petroleum and coal (70), some are natural products. One of these, the terpene *p*-cymene (*p*-isopropyltoluene), is found in volatile oils from over 100 plant species (34). Bacteria that degrade *p*-cymene are relatively common (86). They initiate catabolism of *p*-cymene by oxidizing the benzylic methyl group to yield *p*-cumate (*p*-isopropylbenzoate) (16, 19, 51–53, 89). Similar methyl group oxidations to the corresponding acids are involved in the transformation of other methyl-substituted aromatic compounds including toluene, xylenes (4), methylnaphthalenes (71), dimethylnaphthalenes (5, 59, 71), and *p*-cresol (15).

The most extensively studied aromatic catabolic pathway initiated by methyl group oxidation is that encoded by the TOL plasmid, pWW0, of *Pseudomonas putida* mt-2 (3). The *xyl* genes of pWW0 that encode the oxidation of toluene, *m*-xylene, and *p*-xylene to benzoate, *m*-toluate, and *p*-toluate, respectively, have been cloned and sequenced, and some of the enzymes have been purified and studied in detail (39, 40, 45, 76–78, 81). The pathway is initiated by a monooxygenase (Fig. 1, enzyme A) which catalyzes the oxidation of toluene (or *m*- or *p*-xylene) (Fig. 1, compound I) to benzyl alcohol (or *m*- or *p*-toluic alcohol) (Fig. 1, compound II). The monooxygenase is a two-component enzyme consisting of a reductase subunit which transfers electrons from NADH through FAD and a [2Fe2S] center to the membrane-associated hydroxylase subunit. There, one atom of activated molecular oxygen is inserted into the methyl group of the aromatic hydrocarbon substrate while the other oxygen atom is reduced to water (77, 81). This enzyme also has activity toward the reaction product, benzyl alcohol (or *m*- or *p*-toluic alcohol) (Fig. 1, compound II). The alcohol is presumably oxidized by the monooxygenase to an unstable *gem*-diol intermediate (Fig. 1, compound V) which is recognizable as the hydrate of benzaldehyde (or *m*- or *p*-tolualdehyde) (Fig. 1, compound III) (39). The conversion of benzyl alcohols to benzaldehydes is also catalyzed by an NAD^+-linked alcohol dehydrogenase (76, 78) (Fig. 1, enzyme B); dehydrogenation is likely to be the major route for this transformation. The benzaldehyde (or *m*-tolualdehyde or *p*-tolualdehyde) formed is then acted on by an NAD^+-linked aldehyde dehydrogenase (Fig. 1, enzyme C) to produce benzoate (or *m*-toluate or *p*-toluate) (Fig. 1, compound IV).

A similar series of reactions has been proposed for the catabolism of *p*-cymene through *p*-cumate (19, 51, 53). The further catabolism of *p*-cumate has been studied in *P. putida* PL and F1 (18–21, 24), which employ an eight-step pathway to convert *p*-cumate to isobutyrate, pyruvate, and acetyl coenzyme A. The genes encoding the enzymes that catalyze this reaction sequence, which make up the *cmt* operon, have been cloned from *P. putida* F1 and sequenced (24). The *cmt* operon is located in the chromosome of strain F1 just upstream of the well-characterized *tod* operon for which strain F1 is generally recognized (33, 92, 93). The genes (the *cym* operon and *cymR*) encoding the enzymes that transform *p*-cymene to *p*-cumate

* Phone: (904) 934-9345. Fax: (904) 934-9201. E-mail: eaton.richard@epamail.epa.gov.

† Contribution 993 from the Gulf Ecology Division, National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Gulf Breeze, Fla.

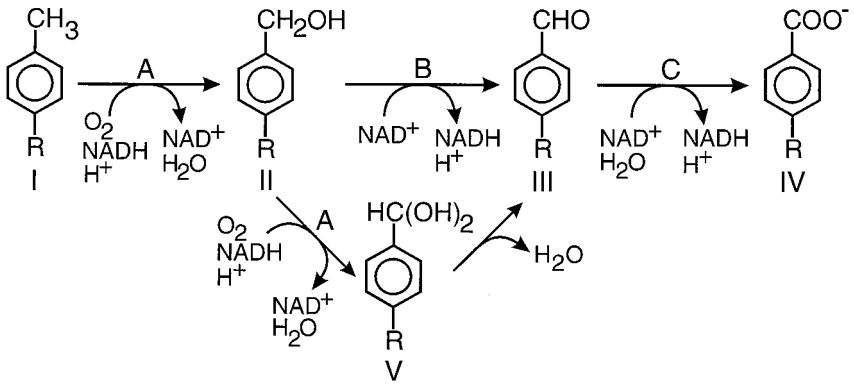

FIG. 1. Pathway for the transformation of toluene, *p*-xylene, and *p*-cymene to benzoate, *p*-toluate, and *p*-cumate, respectively. R represents the following: H, for toluene and metabolites; CH_3, for *p*-xylene and metabolites; and $CH(CH_3)_2$, for *p*-cymene and metabolites. Compounds are represented as follows: I, toluene, *p*-xylene, or *p*-cymene; II, benzyl alcohol, *p*-toluic alcohol, or *p*-cumic alcohol; III, benzaldehyde, *p*-tolualdehyde, or *p*-cumic aldehyde; IV, benzoate, *p*-toluate, or *p*-cumate; and V, α,α′-dihydroxytoluene, α,α′-dihydroxy-*p*-xylene, or α,α′-dihydroxy-*p*-cymene. Enzymes are represented as follows: A, monooxygenase; B, alcohol dehydrogenase; and C, aldehyde dehydrogenase.

and the regulatory protein that controls expression of both the *cym* and *cmt* operons are located in the chromosome of *P. putida* F1 just upstream of the *cmt* operon. This paper describes cloning of this chromosomal DNA and analysis of the *cym* genes and their products.

MATERIALS AND METHODS

Bacterial strains and plasmids. Plasmids and bacterial strains used in this study are listed in Table 1. *P. putida* F1 was originally isolated by using ethylbenzene as the sole carbon and energy source (33). This strain has been used in extensive studies of the genes (*tod*) and enzymes involved in toluene degradation (33, 91–93) and also grows with *p*-cymene and *p*-cumate (24, 26). Molecular genetic studies of *p*-cumate metabolism were initiated previously by ligating a 7.4-kb *Hin*dIII fragment from the chromosome of *P. putida* F1 into the plasmid vector pLV59 to yield pRE611 (see Fig. 2, map coordinates 11.23 to 18.63). Recombinant bacteria carrying this DNA fragment, encoding *p*-cumate dioxygenase and other enzymes of the *p*-cumate pathway, were detected as blue colonies on selective medium containing the chromogenic substrate, indole-2-carboxylate, which is converted to indigo by *p*-cumate dioxygenase (26). The 7.4-kb *Hin*dIII fragment was subsequently used as a probe in colony hybridization experiments to identify recombinant bacteria carrying overlapping cloned fragments, including the 8.15-kb *Bgl*II fragment of pRE764 (see Fig. 2, map coordinates 5.0 to 13.15) and the 14.45-kb *Bgl*II fragment of pRE765 (24). These plasmids and their derivatives form the basis for the work described here. *Escherichia coli* JM109 (90) was the host in cloning experiments.

Media. Luria-Bertani medium (LB) (17) was used for the cultivation of bacteria except where noted. Minimal medium was R medium (29) containing 20 mM sodium lactate. Media were solidified with 1.5% Bacto agar (Difco Laboratories, Detroit, Mich.).

Chemicals and enzymes. *p*-Cymene, *p*-cumic alcohol, *p*-cumic aldehyde, and *p*-cumic acid were obtained from Aldrich Chemical Co., Milwaukee, Wis. NAD^+, coenzyme A, ATP, sodium acetate, and isobutyric acid were from Sigma Chemical Co., St. Louis, Mo. Isopropyl-β-D-thiogalactoside (IPTG) and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) were from Gold Biotechnology, St. Louis, Mo. Preparation of 2,3-dihydroxy-*p*-cumic acid was as described previously (24). Restriction endonucleases and T4 DNA ligase were obtained from New England BioLabs, Beverly, Mass., and used according to the manufacturer's instructions; the 1-kb DNA ladder size standard used was from Bethesda Research Laboratories, Gaithersburg, Md.

Preparation, analysis, and cloning of DNA. DNA isolation and cloning and analyses of clones were carried out as previously described (28). Recombinant bacteria were identified as follows. *E. coli* JM109 carrying different DNA fragments inserted into the multiple cloning site within the *lacZ*α gene of the plasmid vector (pUCBM20, pUCBM21, pBluescriptII KS, or pBluescriptII SK) formed white colonies on selective medium containing ampicillin (100 μg/ml), IPTG (0.25 mM), and X-Gal (0.02%). Bacteria carrying pACYC184 having a *Sal*I fragment inserted into the tetracycline resistance gene were identified as ampicillin-resistant colonies unable to grow on tetracycline-LB agar. Bacteria carrying the positive selection cloning vector pLV59 failed to grow at 37°C on chloramphenicol-LB agar, while bacteria carrying pLV59 having a DNA fragment inserted into a *Hin*dIII or *Bgl*II site within the *Eco*RI restriction endonuclease gene grew and formed colonies. Bacteria carrying DNA fragments that overlap with previously cloned fragments were identified in colony hybridization experiments (35) by using gel-purified (22) ^{32}P-labeled DNA fragments as probes.

DNA sequence determination. Both strands of the DNA segments whose sequences are discussed here were sequenced by the dideoxy chain-termination method, using double-stranded DNA as the template (73). The sequence of plasmid pRE851 was obtained by ACGT, Inc., Northbrook, Ill., by using both manual sequencing with autoradiographic detection and automated sequencing with an Applied Biosystems, Inc. sequencer. The remaining sequences were obtained by the ICBR DNA Sequencing Core Laboratory, University of Florida, Gainesville, by using an Applied Biosystems, Inc., Model 373a sequencer, as previously described (23), with various subclones of pRE847, pRE856, and pRE773. Sequencing reactions were initiated with vector-complementary primers and by using newly synthesized primers which were chosen based on previous sequence data. DNA primers for sequencing were synthesized by the ICBR DNA Synthesis Core Laboratory, University of Florida, Gainesville. Sequence data were aligned and edited by using DNASTAR (DNASTAR, Inc., Madison, Wis.). Searches for specific nucleotide or amino acid sequences in the GenBank database were carried out by using the BLAST program (2). Sequences were retrieved from GenBank and compared with sequences obtained here by using DNASTAR COMPARE, ALIGN, and AALIGN programs (62, 87).

Cultivation of organisms. To cultivate bacteria for chemical transformations or preparation of extracts, a 10- to 20-ml overnight culture of a recombinant strain in LB medium, supplemented with ampicillin (100 μg ml^{-1}) or chloramphenicol (30 μg ml^{-1}), was used to inoculate 500 ml of the same medium. After a 2-h incubation at 30°C, inducer was added (1 mM IPTG, 2 mM *p*-cumate, or 0.5 mM *p*-cymene), and the incubation was continued for 3 h. Cells were harvested by centrifugation, washed in 50 mM sodium-potassium phosphate buffer (pH 7.0), and resuspended in 5 ml of the same buffer.

Preparation of cell extracts. Cell extracts were prepared by passing cell suspensions twice through a chilled French pressure cell (SLM Instruments, Inc., Urbana, Ill.) at 14,000 to 20,000 lb/in^2. Particulate material was removed by centrifugation at 47,800 × *g* for 40 min at 4°C.

Analysis of enzymes produced by recombinant bacteria. UV-visible spectra were recorded with a Perkin-Elmer Lambda 6 double-beam spectrophotometer and were prepared for publication by using CorelDRAW! (Corel Corp., Ottawa, Canada). Acyl coenzyme A synthetase activity in extracts was measured by using the hydroxamate assay of Brown et al. (12) with acetate, isobutyrate, and *p*-cumate as substrates. *p*-Cumic alcohol dehydrogenase was assayed spectrophotometrically at 30°C by recording the rate of increase in absorbance at 261 nm due to the formation of *p*-cumic aldehyde ($\Delta\varepsilon_{261}$ = 16,000 M^{-1} cm^{-1}) over time. Reaction mixtures contained, in 1 ml of 50 mM sodium-potassium phosphate buffer (pH 7), 0.2 mM *p*-cumic alcohol, 0.1 mM NAD^+, and from 5 to 20 μl of cell extract. 2,3-Dihydroxy-*p*-cumate-3,4-dioxygenase was assayed spectrophotometrically at 30°C by recording the increase in absorbance at 346 nm due to the formation of 2-hydroxy-3-carboxy-6-oxo-7-methylocta-2,4-dienoate ($\Delta\varepsilon_{346}$ = 25,000 M^{-1} cm^{-1}) over time (24). Reaction mixtures contained, in 1 ml of 50 mM sodium-potassium phosphate buffer (pH 7), 0.075 mM 2,3-dihydroxy-*p*-cumate and 1 to 20 μl of cell extract. Protein concentrations in extracts were determined by using the bicinchoninic acid protein assay reagent (Pierce, Rockford, Ill.) with bovine serum albumin as standard.

A qualitative screen for *p*-cymene monooxygenase was carried out with strains JM109(pRE897) and JM109 lacking an introduced plasmid. Following growth as described above, washed cells were incubated in R medium supplemented with 0.2% lactate and 1.0 mM *p*-cymene or *p*-cumic alcohol for 20 h at 30°C with shaking. After cells were removed by centrifugation, culture supernatants were acidified to pH 2 with HCl and extracted with 1.0 volume of ethyl acetate. As controls, these substrates, together with *p*-cumic aldehyde and *p*-cumate, were separately incubated under the same conditions but in the absence of cells, and

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Description	Reference(s) or source
Strains		
P. putida F1	Grows with toluene, *p*-cymene, and *p*-cumate	26, 33, 92
E. coli JM109	*recA endA1 gyrA96 thi hsdR17 supE44 relA1* Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q} *ZΔM15*)	90
Plasmids		
pACYC184	Ap^r Tc^r	13
pLV59	Encodes *Eco*RI restriction endonuclease and temperature-sensitive *Eco*RI methylase, positive-selection cloning vector, Cm^r	65
pBluescriptII KS	Ap^r, multiple cloning site in *lacZ*α	Stratagene Cloning Systems; 1
pBluescriptII SK	Identical to pBluescriptII KS but with multiple cloning site reversed	1
pUCBM20	Derived from pUC18 with the following additional cloning sites inserted into *lacZ*α: *Nco*I, *Eco*RV, *Apa*I, *Sac*II, *Xma*III, *Not*I, *Bfr*I, *Mlu*I; Ap^r	Boehringer Mannheim; 90
pUCBM21	Identical to pUCBM20 but with the multiple cloning site reversed	90
pRE611	7.4-kb *Hin*dIII fragment (map coordinates 11.23 to 18.63) from *P. putida* F1 in pLV59, Ind$^+$[a], Cm^r, carries *cmt* operator, *cmtAaAbAcCBAdDI*	24, 26
pRE634	7.4-kb *Hin*dIII fragment (map coordinates 11.23 to 18.63) from pRE611 in pMMB277, Ind$^+$, Cm^r	24
pRE665	pRE611 digested with *Sma*I and *Hpa*I to delete a 2.5-kb fragment and religated, carries *cmt* operator-promoter, *cmtAaAbAc*, and *cmtC*, which encodes 2,3-dihydroxy-*p*-cumate dioxygenase but not *cmtAd*, *cmtB*, or *cmtD*. The *cmt* operon is downstream from and in the same orientation as the *Eco*RI restriction endonuclease promoter of the vector	24
pRE764	8.15-kb *Bgl*II fragment (map coordinates 5.0 to 13.15) from *P. putida* F1 inserted into pLV59, hybridized to 7.4-kb *Hin*dIII fragment from pRE634, Cm^r	24
pRE765	14.45-kb *Bgl*II fragment (map coordinates 16.95 to 31.4) from *P. putida* F1 inserted into pLV59, hybridized to 7.4-kb *Hin*dIII fragment from pRE634, Cm^r	24
pRE773	2.95-kb *Sal*I fragment (map coordinates 21.75 to 26.7) from pRE765 inserted into pUCBM20, Ap^r	24
pRE847	10.22-kb *Hin*dIII fragment (map coordinates 1.01 to 11.23) from strain F1 inserted into pLV59, hybridized to a 1.8-kb *Bgl*II-*Sal*I fragment (map coordinates 5.0 to 6.84) from pRE764	This study
pRE851	6.4-kb *Bam*HI fragment (map coordinates 5.1 to 11.5) from pRE764 inserted into pBluescriptII SK	This study
pRE852	4-kb *Hin*dIII-*Bgl*II fragment (map coordinates 1.01 to 5.0) from pRE847 inserted into *Hin*dIII-*Bam*HI-digested pBluescriptII KS, carries *cymB*	This study
pRE855	*Bgl*II fragments (three) from strain F1, inserted into pLV59, hybridized to a 4-kb *Hin*dIII-*Bgl*II fragment (map coordinates 1.01 to 5.0) from pRE847	This study
pRE856	22.9-kb *Bgl*II fragment (map coordinates −17.9 to 5.0) inserted into pLV59, obtained by recutting and ligating pRE855	This study
pRE890	2.75 kb *Mlu*I fragment (map coordinates −2.75 to 0) from pRE856 inserted into pUCBM20	This study
pRE891	5.0-kb *Mlu*I fragment (map coordinates −14.55 to −9.55) from pRE856 inserted into pUCBM20	This study
pRE892	2.21-kb *Mlu*I fragment (map coordinates 0 to 2.21) from pRE856 inserted into pUCBM20, carries *cymR*	This study
pRE893	3.95-kb *Mlu*I fragment (map coordinates −17.9 to −14.55) from pRE856 inserted into pUCBM20 (the insert contains 0.6 kb of pLV59 DNA)	This study
pRE893	3.95-kb *Mlu*I fragment (map coordinates −17.9 to −14.55) from pRE856 inserted into pUCBM20 (the insert contains 0.6 kb of pLV59 DNA)	This study
pRE894	6.8-kb *Mlu*I fragment (map coordinates −9.55 to −2.75) from pRE856 inserted into pUCBM20	This study
pRE895	6.2-kb *Xma*I-*Hin*dIII fragment (map coordinates −5.2 to 1.01) from pRE856 inserted into pUCBM20	This study
pRE896	2.04-kb *Cla*I-*Nhe*I fragment (map coordinates 9.56 to 11.6) from pRE764 inserted into *Cla*I-*Xba*I-digested pBluescriptII KS, carries *cymE*	This study
pRE897	6.13-kb *Bam*HI-*Hin*dIII fragment (map coordinates 5.1 to 11.23) from pRE847 inserted into pBluescriptII SK, carries *cymAaAbD*	This study
pRE902	2.02-kb *Apa*I-*Sca*I fragment (map coordinates 3.64 to 5.66) from pRE847 inserted into *Apa*I-*Sma*I-digested pUCBM21, carries *cymC*	This study
pRE909	pRE665 recut with *Hin*dIII and ligated, orientation opposite to that of the *Eco*RI restriction endonuclease promoter of the vector, pLV59, Cm^r, carries *cmt* operator-promoter and *cmtC*	This study
pRE912	2.09-kb *Sal*I fragment (map coordinates 2.72 to 4.81) from pRE852 inserted into pACYC184 in the orientation opposite to that of the tetracycline resistance gene promoter, carries *cym* operator-promoter and *cymB*, Cm^r	This study

[a] Ind$^+$, converts indole-2-carboxylate to indigo.

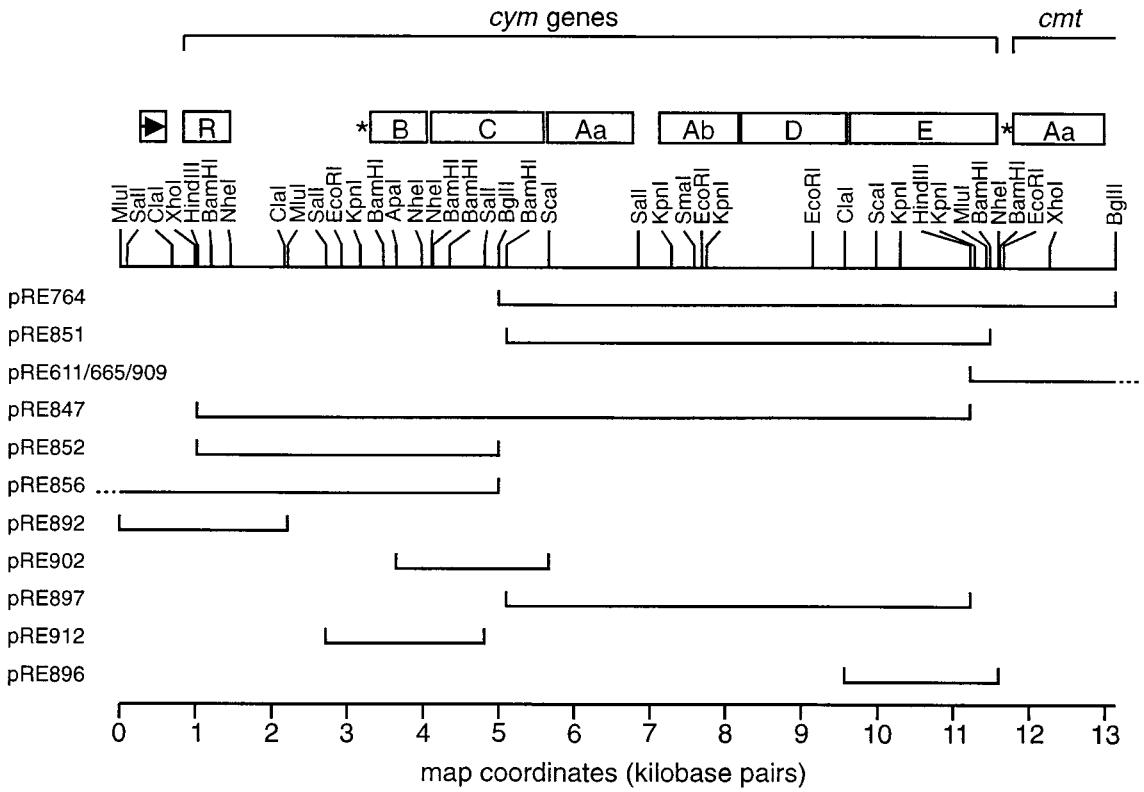

FIG. 2. DNA from *P. putida* F1 encoding the transformation of *p*-cymene to *p*-cumate; restriction map of a 13.14-kb DNA segment. The locations of the genes encoding pathway enzymes (see Fig. 1) are indicated at the top. Below the map are shown DNA fragments that have been cloned or subcloned or deletions that have been obtained. Names of plasmids containing these fragments are indicated at the left. The boxed arrow indicates the upstream repeated element; asterisks indicate operator-promoter sites for *cym* and *cmt* operons.

similarly extracted. After being dried over sodium sulfate and concentrated on the rotary evaporator under reduced pressure, the extracted products were analyzed by thin-layer chromatography on silica gel plates (Kieselgel 60 F_{254}; Merck) developed with ethyl acetate. Chemical spots were detected by their quenching of fluorescence under UV light. Mobilities (R_f) were as follows: *p*-cymene, 0.66; *p*-cumic alcohol, 0.58; *p*-cumic aldehyde, 0.63; and *p*-cumic acid, 0.53.

Nucleotide sequence accession number. The DNA sequences obtained in this study are available from GenBank (accession no. U24215).

RESULTS AND DISCUSSION

Cloning *p*-cymene catabolic genes. Bacteria carrying the plasmids pRE847, pRE855, and pRE856 containing overlapping fragments from the chromosome of *P. putida* F1 extending upstream to map coordinate −17.9 were identified by using DNA fragments derived from pRE764 as probes (Table 1; Fig. 2). Beginning with these plasmids and using information obtained from restriction mapping, DNA sequence analysis, and enzyme assays of selected recombinants, various subclones were constructed. These are described in Table 1 and Fig. 2 and below.

Nucleotide sequence of the DNA encoding the transformation of *p*-cymene to *p*-cumate. The sequence of a DNA segment encoding the transformation of *p*-cymene to *p*-cumate was obtained (data not shown). This new sequence extends from the *Mlu*I site at map coordinate 0 (Fig. 2) 11,236 bp to a *Hin*dIII site. The sequence to the right of the *Hin*dIII site was published previously (24).

Sequence comparisons. The BLAST program (2) was used to search the GenBank database (8) for DNA sequences having homologies to *cym* genes (blastn) and for proteins having identities with the deduced amino acid sequences of *p*-cymene pathway gene products (blastx). Sequences were then retrieved and compared with sequences obtained in this study. These sequence similarities helped in assigning functions to *cym* gene products (described below). The *p*-cymene pathway genes and their products are *cymAa* and *cymAb*, which encode the two components of *p*-cymene monooxygenase, a hydroxylase and a reductase, respectively; *cymB*, encoding *p*-cumic alcohol dehydrogenase; *cymC*, encoding *p*-cumic aldehyde dehydrogenase; *cymD*, encoding an outer membrane protein; *cymE*, encoding an acetyl coenzyme A synthetase; and *cymR*, which encodes a regulatory protein (repressor) that controls expression of both *cym* and *cmt* operons.

***p*-Cymene monooxygenase.** When incubated with *p*-cymene, recombinant strain *E. coli* JM109(pRE897), containing the genes *cymAa* and *cymAb*, which encode the two subunits of *p*-cymene monooxygenase, converted *p*-cymene to a mixture of *p*-cumic alcohol, *p*-cumic aldehyde, and *p*-cumic acid. Incubation of this strain with *p*-cumic alcohol also yielded aldehyde and acidic products. Strain JM109, lacking pRE897, did not transform *p*-cymene or *p*-cumic alcohol (data not shown). A previous study of the comparable enzyme of the TOL plasmid (pWW0)-encoded toluene and *m*- and *p*-xylene catabolic pathway, xylene monooxygenase (39), showed that this enzyme can attack benzylic and toluic alcohols as well as toluene and xylenes, probably with the formation of *gem*-diol intermediates (Fig. 1, compound V) which spontaneously dehydrate to yield the corresponding aldehydes (Fig. 1, compound III). Similarly, *p*-cymene monooxygenase-carrying bacteria act on *p*-cymene and *p*-cumic alcohol to produce *p*-cumic aldehyde. The alde-

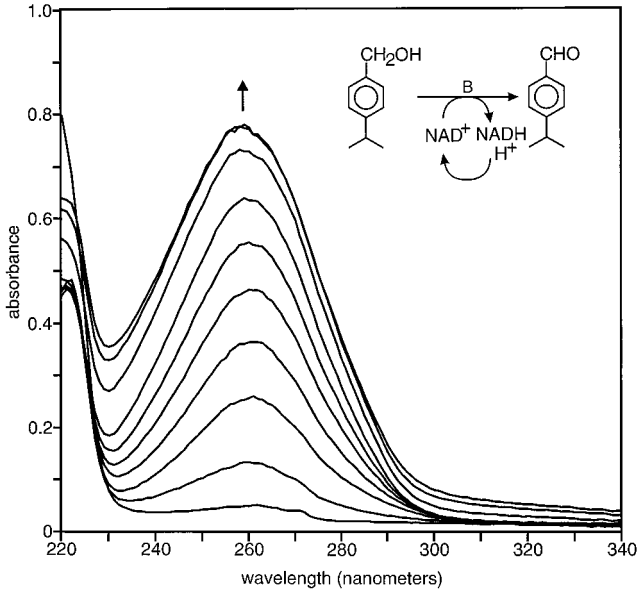

FIG. 3. Conversion of *p*-cumic alcohol to *p*-cumic aldehyde by extracts of *E. coli* JM109(pRE853) at 30°C. The sample and reference cuvettes contained 50 mM potassium-sodium phosphate buffer (pH 7.0) and 30 nmol of NAD^+ in 1-ml volumes. The sample cuvette also contained 50 nmol of *p*-cumic alcohol. Spectra were recorded before the addition of 8 μl of extract containing 140 μg of protein to both cuvettes and after 0.17, 4, 8, 12, 16, 20, 28, and 32 min. This conversion is not catalyzed by extracts of strain JM109 lacking *cymB*. NADH formed in the reaction is converted back to NAD^+ by an NADH oxidase present in extracts and does not accumulate.

hyde product is then apparently oxidized spontaneously to *p*-cumate, a possibility supported by the observation that *p*-cumic aldehyde can be converted to *p*-cumate by simply shaking a solution of the aldehyde in minimal medium in the absence of bacterial cells (data not shown).

Based on studies with xylene monooxygenase and other related enzymes described below, it is likely that the reductase subunit of *p*-cymene monooxygenase, a combined ferredoxin-flavoprotein, acts to transfer electrons from NADH through the flavin and ferredoxin moieties to the hydroxylase subunit. There, one atom of activated molecular oxygen is incorporated into the benzylic methyl group of *p*-cymene to form *p*-cumic alcohol, while the other atom is reduced to water.

Hydroxylase subunit of *p*-cymene monooxygenase. The *cymAa* gene product, the hydroxylase subunit of *p*-cymene monooxygenase, has a deduced molecular weight of 43,117 (376 amino acids). This protein is most closely related to the hydroxylase subunits of xylene monooxygenase (40, 81) encoded by pWW0 of *P. putida* mt-2, which has 39% identity in 363 amino acid residues, and alkane hydroxylase (50) encoded by the OCT (octane catabolic) plasmid of *Pseudomonas oleovorans*, which has 22% identity in 336 amino acids residues. Fifty amino acid residues including 11 histidines are conserved in these three proteins. Eight of the conserved histidine residues form a motif found not only in xylene monooxygenase and the cytoplasmic membrane-associated (82) alkane hydroxylase but also in the rat liver stearoyl coenzyme A Δ^9 desaturase and 16 other membrane-associated eukaryotic and cyanobacterial desaturases (75). This motif, $HX_{(3\text{ or }4)}HX_{(20-50)}HX_{(2\text{ or }3)}HHX_{(100-200)}HX_{(2\text{ or }3)}HH$ (75), is formed in cymene monooxygenase hydroxylase by amino acid residues 112-HELYH-116, 142-HVVGHH-146, and 281-HADHH-285. All eight histidine residues have been shown to be essential for catalytic activity in the rat liver enzyme. That enzyme and other enzymes having this motif are all proposed to have iron-containing active sites involved in the activation of O_2 for catalysis (75). It has been suggested that these histidine residues may act as ligands for two iron atoms (75) and that this group of enzymes may form a new category (class III) of membrane-associated enzymes having oxo-bridged di-iron clusters at their active sites (31, 75).

Reductase subunit of cymene monooxygenase. The *cymAb* gene product, the reductase subunit of cymene monooxygenase, has a deduced molecular weight of 38,295 (349 amino acids). This protein is most closely related to a family of reductases in which flavoprotein and ferredoxin are combined. These include xylene monooxygenase reductase (81), which has 42% identity in 347 amino acid residues; phenol hydroxylase reductases, one of which from *P. putida* (44) has 35% identity in 336 amino acid residues; and naphthalene dioxygenase reductases, of which the NAH7 plasmid-encoded enzyme (80) has 32% identity in 334 amino acid residues. The ferredoxin of the photosynthetic organelle (cyanelle) of cyanobacterium *Cyanophora paradoxa* (64) has 34% identity to the amino terminus in 77 amino acid residues.

The amino terminus of the reductase is that of a typical chloroplast-type ferredoxin (32, 56) with a [2Fe-2S] binding site composed of cysteine residues located at positions 49, 54, 57, and 89. This site is conserved in all members of this family examined to date (63). Neidle et al. (63) have compared the sequence of a closely related protein, xylene monooxygenase reductase, to similar reductases and identified sites (which are also present in cymene monooxygenase reductase) likely to be involved in binding a ribose moiety of the cofactor, NADH, (amino acid residues 217-DAPMVCIAGGSGLAPLISILQHA R-239) and the isoalloxazine ring of FAD (residues 159-RSYS FAN-165) (47, 69).

***p*-Cumic alcohol dehydrogenase.** Extracts prepared from *E. coli* JM109 carrying any of the plasmids in which *cymB* is present, such as pRE853 and pRE912, catalyzed the conversion of *p*-cumic alcohol to *p*-cumic aldehyde in the presence of NAD^+ (Fig. 3).

The *cymB* gene product, *p*-cumic alcohol dehydrogenase, is one of the class II (short chain) alcohol dehydrogenases (68, 72) whose members are dimers or tetramers comprised of subunits of about 250 amino acids in length. *p*-Cumic alcohol dehydrogenase has a deduced molecular weight of 26,149 (252 amino acids) and contains the six amino acid residues present in all members of this group (Gly-12, Gly-18, Asp-63, Gly-133, Tyr-153, and Lys-157) as well as five of seven residues present in most members (Thr-11, Gly-16, Gly-91, Asn-112, and Ser-140). The glycine residues at positions 12, 16, and 18 serve to form the conserved center (72) of a βαβ structure (amino acid residues 6-KVAIVTGAATGIGNAIVRSYLAEGAKVVIAD-36) that is probably involved in binding the ADP moiety of NAD^+ (68, 85). The most closely related enzymes are bacterial oxidoreductases; glucose dehydrogenase of *Bacillus megaterium* (58) has 35% identity in 252 amino acid residues, while 3-ketoacyl-acyl carrier protein reductase of *Vibrio harveyi* (79) has 33% identity in 244 amino acid residues. The TOL plasmid-encoded benzyl or toluic alcohol dehydrogenase (76, 78) is a member of the class I zinc-containing long-chain alcohol dehydrogenases and is not related to *p*-cumic alcohol dehydrogenase (72).

Cumic aldehyde dehydrogenase. Extracts prepared from *E. coli* JM109(pRE902), which contains *cymC*, catalyzed the conversion of *p*-cumic aldehyde to *p*-cumate in the presence of NAD^+ (Fig. 4).

The *cymC* gene product, *p*-cumic aldehyde dehydrogenase, has a deduced molecular weight of 53,175 (494 amino acids). The enzyme belongs to a large family of NAD^+-linked alde-

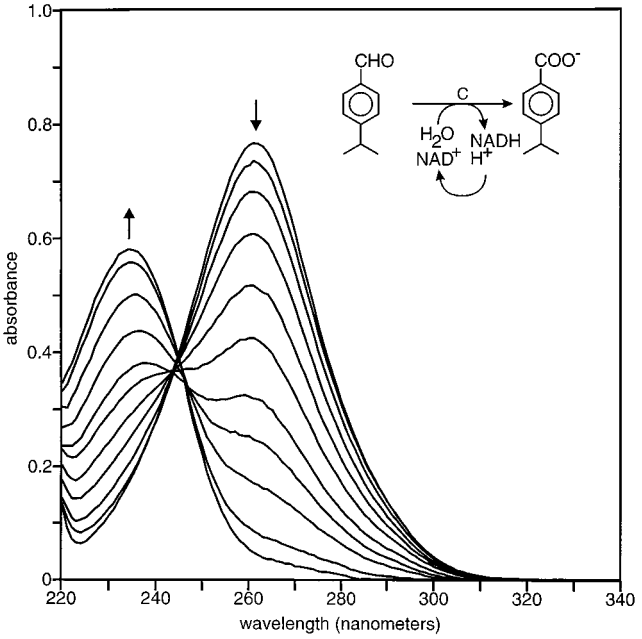

FIG. 4. Conversion of *p*-cumic aldehyde to *p*-cumate by extracts of *E. coli* JM109(pRE902) at 30°C. The sample and reference cuvettes contained 50 mM potassium-sodium phosphate buffer (pH 7.0) and 50 nmol of NAD^+ in 1-ml volumes. The sample cuvette also contained 50 nmol of *p*-cumic aldehyde. Spectra were recorded before the addition of 10 μl of extract containing 160 μg of protein to both cuvettes and after 0.17, 8, 16, 24, 32, 40, 48, 56, 64, and 72 min. This conversion is not catalyzed by extracts of strain JM109 lacking *cymC*. NADH formed in the reaction is converted back to NAD^+ by NADH oxidase present in extracts and does not accumulate.

hyde oxidoreductases (43). Members of this family contain 23 invariant amino acid residues, 27 nearly invariant residues, and 41 invariant similarities (43). All of these are present in cumic aldehyde dehydrogenase except for one of the invariant similarities (Ala at residue 386 is substituted for one of the larger hydrophobic amino acids, Val, Ile, Leu, or Met). The protein lacks the conserved GXGXXG sequence used by Wierenga et al. to predict ADP-binding βαβ folds (85), but it contains a similar sequence conserved in other aldehyde dehydrogenases (43) in which the second glycine is replaced by a threonine (amino acid residues 243-GSTVVG-248). This sequence and the surrounding residues (239-ITFTGSTVVGKKIVEYALG NMKRVTLEL-266) have been proposed to form the coenzyme-binding domain in this enzyme family (43). The most closely related enzymes are aldehyde dehydrogenases of eukaryotes; the mouse mitochondrial enzyme (14) has 48% identity in 481 amino acid residues. The TOL plasmid-encoded tolualdehyde dehydrogenase (45) is more distantly related, having 34% identity in 472 amino acid residues.

Outer membrane protein. The *cymD* gene product, a putative outer membrane protein, has a deduced molecular weight of 48,932 (460 amino acids). It is similar to other proposed outer membrane proteins encoded by genes that are part of operons encoding the catabolism of aromatic compounds, including IpbH, of the isopropylbenzene catabolic operon of *P. putida* RE204 (45% identity in 460 amino acid residues) (25, 30), CumH of the cumene catabolic operon of *Pseudomonas fluorescens* IP01 (44% identity) (36), TodX of the toluene catabolic operon of strain F1 (41% identity in 444 amino acids) (83), XylN of the xylene catabolic operon of the plasmid pWW0 (39% identity in 445 amino acids) (38), as well as a *Haemophilus influenzae* P1 outer membrane protein subtype 3L (23% identity in 431 amino acids) (60) and the *E. coli* FadL fatty acid transport protein (22% identity in 275 amino acids) (9).

The amino terminus of CymD (residues 1-MKKTIYSLSAC GILTCLYCGIASAT-25) has the features of a possible signal peptide required for secretion and integration of the protein into the membrane (66): 1 to 3 positively charged amino acid residues (lysines at positions 2 and 3 in CymD), followed by a core of 14 to 20 neutral, mostly hydrophobic, amino acids and a processing site adjacent to three conserved amino acid residues, (L, V, I, A, G, or S)-X-(A, G, or S). In CymD this processing site could be adjacent to either amino acid residues 20-GIA-22 or 22-ASA-24. Wang et al. (83) demonstrated that the similar protein, TodX, is membrane associated and provided evidence that a possible physiological role for TodX is to facilitate toluene movement across the membrane at low concentrations. The role of CymD in the catabolism of *p*-cymene may be similar, but it has not been studied.

Acetyl coenzyme A synthetase. The *cymE* gene product, presumed to be an acyl coenzyme A synthetase, has a deduced molecular weight of 71,425 (649 amino acids). It is similar to a number of acetyl coenzyme A synthetases, notably those from the cyanobacterium *Synechocystis* sp., (42% identity in 636 amino acid residues [46]), *E. coli* (40% identity in 629 amino acid residues [10]), and the parasitic protozoan *Cryptosporidium parvum* (36% identity in 634 amino acid residues [48]). These enzymes catalyze a two-step reaction. First, acetate is activated by reaction with ATP, producing an acetyl adenylate intermediate and pyrophosphate; second, coenzyme A displaces the adenylate to yield acetyl coenzyme A. Khramtsov et al. (48), in comparing amino acid sequences of fifteen acetyl coenzyme A synthetases, identified two highly conserved motifs (I and II) separated by 289 to 341 amino acids. Both of these are present in all of the acetyl coenzyme A synthetases as well as in other AMP-binding proteins and CymE. Motif I is defined as (Y/F)TS(G/A)(S/T)(T/S)GXPK(G/M); in CymE it is amino acid residues 256-YTSGTTGKPKG-266. Motif II is defined as PKT(R/V/L)SGK(I/V/T)(T/M/V/K)R(R/N); in CymE it is amino acid residues 601-PKTRSGKLLRR-611. All of the enzymes that contain motif I are likely to catalyze the formation of an acyl adenylate intermediate, and it has been proposed (4) that this may be the function of this structure, while motif II has similarity to the phosphate-binding P loop of ATPases and GTPases (48, 74) and may be involved in the binding of ATP.

The role of the *cymE*-encoded acyl coenzyme A synthetase in *p*-cymene metabolism remains unknown. Extracts of IPTG-induced *E. coli* JM109(pRE896), which contains *cymE*, were assayed for acyl coenzyme A synthetase activity by using as substrates three compounds that are intermediates or products of *p*-cymene catabolism: acetate, isobutyrate, and *p*-cumate. Isobutyrate, formed by 2-hydroxy-6-oxo-7-methylocta-2,4-dienoate hydrolase in the *p*-cumate pathway (24), can be further metabolized by a pathway normally employed for the metabolism of branched-chain amino acids once it is converted to its coenzyme A ester (55). Acetate might be formed by the spontaneous oxidation of acetaldehyde, a product of 4-hydroxy-2-oxovalerate aldolase in the *p*-cumate pathway; the acyl coenzyme A synthetase could provide a way to salvage this acetate. Extracts failed to show detectable activity toward *p*-cumate and isobutyrate. Although a high level of acetyl coenzyme A synthetase activity was demonstrated, it was not greater than the level of such activity in extracts of strain JM109 (data not shown).

Regulatory protein. The *cymR* gene product, a putative regulatory protein, has a deduced molecular weight of 23,324 (203

TABLE 2. Regulation of *cmt* and *cym* expression by *cymR*

Bacterial strain	Inducer	2,3-Dihydroxy-*p*-cumate dioxygenase (CmtC) activity (nmol min^{-1} mg of $protein^{-1}$)[a]	Cumic alcohol dehydrogenase (CymB) activity (nmol min^{-1} mg of $protein^{-1}$)[a]
JM109(pRE665)	Uninduced	510 (± 40)	
JM109(pRE665)	Cumate	540 (± 9)	
JM109(pRE665)	Cymene	510 (± 26)	
JM109(pRE665)(pRE892)	Uninduced	3.5 (± 0.2)	
JM109(pRE665)(pRE892)	Cumate	550 (± 120)	
JM109(pRE665)(pRE892)	Cymene	4 (± 0.2)	
JM109(pRE909)	Uninduced	130 (± 3)	
JM109(pRE909)(pRE892)	Uninduced	0 (± 0)	
JM109(pRE909)(pRE892)	Cumate	127 (± 1)	
JM109(pRE909)(pRE892)	Cymene	−0.4 (± 0.3)	
JM109(pRE912)	Uninduced		33 (± 9)
JM109(pRE912)(pRE892)	Uninduced		−0.3 (± 0.6)
JM109(pRE912)(pRE892)	Cumate		34 (± 4)
JM109(pRE912)(pRE892)	Cymene		−0.1 (± 0.2)

[a] Activities are the means of three assays (± standard errors).

amino acids). It has only a very low level of similarity to other regulatory proteins including MtrR, which regulates the permeability of the *Neisseria gonorrhoeae* cell membrane to hydrophobic compounds (67). This similarity is limited to a region of CymR that is likely to form a helix-turn-helix DNA-binding structure (amino acid residues 17-GKLIAAALGVLREKGYA GFRI-37) (11). The characteristics of the helix-turn-helix motif have been determined by comparison of the amino acid sequences of a number of DNA-binding proteins along with structural studies of some of these such as the bacteriophage λ Cro protein (11). In a helix-turn-helix structure, residue 9 is glycine; residues 3 to 8 and 15 to 20 cannot be proline since they are within α helices; residues 4 and 15 should be uncharged since they are buried in Cro; and residue 5 should not be a β-branched residue since it is wedged between the two helices in Cro (11).

Synthesis of the enzymes that catalyze the conversion of *p*-cymene to *p*-cumate and the further metabolism of *p*-cumate is induced in *P. putida* F1 during growth with either *p*-cymene or *p*-cumate (27). To study regulation of the *p*-cymene (*cym*) and *p*-cumate (*cmt*) operons in strain F1, plasmids carrying the regulatory gene and the proposed operator-promoter regions of those operons upstream of genes encoding assayable enzymes were constructed (Table 1; Fig. 2). A DNA fragment carrying the *cmt* operator-promoter upstream of *cmtC*, encoding 2,3-dihydroxy-*p*-cumate 3,4-dioxygenase (Fig. 2, extending from map coordinate 11.23 to the right), was inserted into the vector pLV59 in either orientation to make pRE665 and pRE909. pRE912 was constructed by inserting a 2.09-kb *Sal*I fragment (Fig. 2, map coordinates 2.72 to 4.81) carrying the *cym* operator-promoter upstream of *cymB*, encoding *p*-cumic alcohol dehydrogenase, into pACYC184. To make pRE892, a 2.21-kb *Mlu*I fragment carrying the regulatory protein-encoding gene *cymR* was inserted into vector pUCBM20. *E. coli* JM109 constructs carrying pRE665, pRE909, or pRE912, with or without the compatible plasmid pRE892, were grown in the presence or absence of inducer (*p*-cymene or *p*-cumate); subsequently, the appropriate enzyme was assayed in cell extracts (Table 2). It should be noted that differences in the levels of 2,3-dihydroxy-*p*-cumate-3,4-dioxygenase activity between strains containing pRE665 and pRE909 are probably the result of transcription initiated at the *Eco*RI endonuclease promoter of the vector pLV59. In pRE665, this transcription is in the same direction as normal transcription of *cmtC*; in pRE909, it is in the opposite direction. The results indicate that *p*-cumate is the inducer of both the *cym* and *cmt* operons while *p*-cymene is not an inducer. Moreover, extracts of bacteria lacking the regulatory gene, *cymR*, have a high level of constitutive enzyme activity, while extracts of uninduced bacteria containing *cymR* lack enzyme activity; the regulatory gene product is, therefore, a repressor. With *p*-cymene as a growth substrate, conversion of *p*-cymene to *p*-cumate by enzymes present at low, uninduced levels in *P. putida* F1 would be required prior to induction of both operons by *p*-cumate.

A comparison of the operator-promoter regions of the *cym* and *cmt* operons (Fig. 5) revealed a sequence, located between the putative −35, −10 (σ^{70}) promoters (41) and the beginning of the first genes, that is present in both operons; this may be a recognition site (operator) for the CymR repressor protein.

Nucleotide sequence of the DNA that separates the *cmt* and *tod* operons. In order to complete the sequence of the DNA segment carrying the *cym*, *cmt*, and *tod* operons, the sequence of 915 bp of DNA between *cmt* and *tod* was obtained (data not shown). Inspection of the sequence extending from the end of the *cmt* operon (*cmtG*) to the end of *todR* (*todR* is transcribed from right to left) revealed two interesting features (Fig. 6). The first is a gene encoding a possible enoyl coenzyme A hydratase, and the second is a sequence having significant homology to DNA located upstream of *cymR* and to DNA from other *Pseudomonas* species.

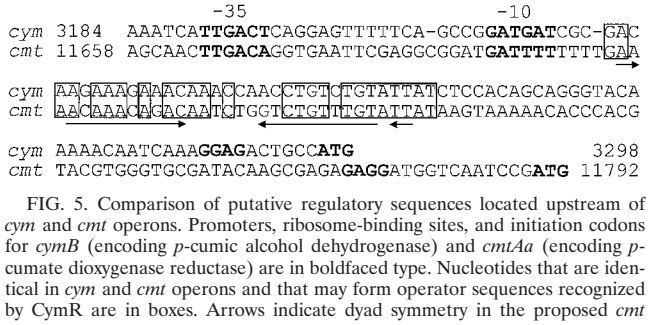

FIG. 5. Comparison of putative regulatory sequences located upstream of *cym* and *cmt* operons. Promoters, ribosome-binding sites, and initiation codons for *cymB* (encoding *p*-cumic alcohol dehydrogenase) and *cmtAa* (encoding *p*-cumate dioxygenase reductase) are in boldfaced type. Nucleotides that are identical in *cym* and *cmt* operons and that may form operator sequences recognized by CymR are in boxes. Arrows indicate dyad symmetry in the proposed *cmt* operator.

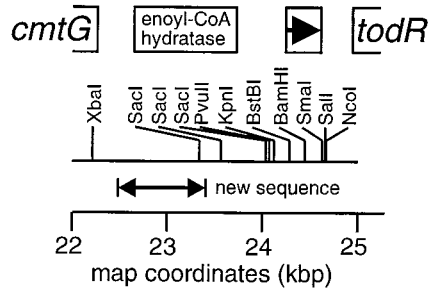

FIG. 6. DNA segment from *P. putida* F1 connecting the *p*-cumate (*cmt*) and toluene (*tod*) catabolic operons. This map is based on newly acquired sequence data extending from base pair 22,488 to 23,404 (data not shown) and previously published data (GenBank accession no. U24215 [24] and U18304 [83]). The boxed arrow indicates the downstream repeated element.

Enoyl coenzyme A hydratase. The presumed product of the gene extending from base pair 22,652 to base pair 23,716 (Fig. 6), enoyl coenzyme A hydratase, has a deduced molecular weight of 39,785 (355 amino acids). It is similar to several enoyl coenzyme A hydratases, most closely to enzymes from *Rhizobium meliloti* (29% identity in 258 amino acid residues) (54), rat mitochondria (34% identity in 182 amino acid residues) (57), and human mitochondria (25% identity in 240 amino acid residues) (61). The function of the gene product has not been examined. Other enoyl coenzyme A hydratases are involved in pathways for the degradation of fatty acids by β oxidation. A role in isobutyrate catabolism is not excluded.

Repeated DNA elements. Comparison of DNA sequences located upstream of *cym* genes (Fig. 2) and downstream from *cmt* genes and the enoyl coenzyme A hydratase gene (between the *cmt* and *tod* operons) (Fig. 6) revealed the presence of regions of sequence homology. Moreover, these sequences have homology to DNA sequences (*ankB* and *ankF*) proposed to encode ankyrins (discussed below) in *Pseudomonas aeruginosa* (42) and *Pseudomonas syringae* (49), respectively. DNA segments carrying these upstream and downstream homologous sequences are aligned with *ankB* in Fig. 7. The right end of the upstream sequence and the left end of the downstream sequence compared here are separated in strain F1 by 23,631 bp. The beginning and end of the regions of homology are not well defined. However, within the sequences aligned in Fig. 7 there are large overlapping stretches of significant homology. Thus, between upstream base pairs 398 to 493 (Fig. 7) and downstream base pairs 24,374 to 24,469, there is 82% homology (79 of 96 bp); this includes a segment of 23 identical base pairs between upstream base pairs 463 to 485 and downstream base pairs 24,439 to 24,461. Between upstream base pairs 400 to 578 and *ankB* base pairs 209 to 386 there is 75% homology. Between downstream base pairs 24,252 to 24,554 and *ankB* base pairs 84 to 382 there is 66% homology.

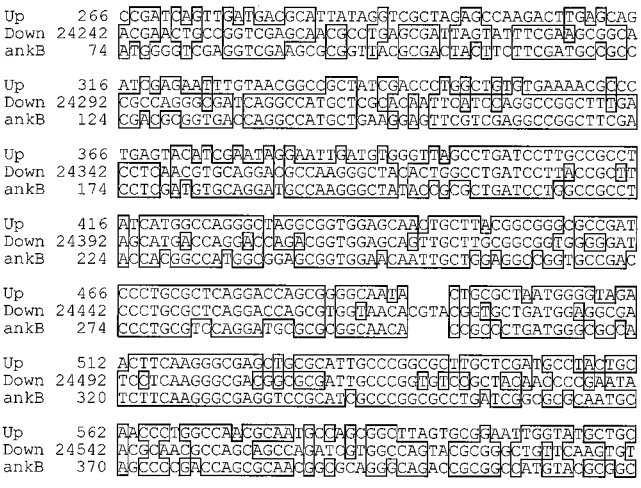

FIG. 7. Comparison of the DNA sequences of directly repeated elements bordering *cym-cmt* DNA. The right end of the upstream compared is 23,631 bp upstream of the left end of the downstream sequence. *ankB* (GenBank accession number U59457) (42) encodes a putative ankyrin protein in *P. aeruginosa*.

Ankyrins are proteins whose role in eukaryotic cells is to link the spectrin-based membrane skeleton to the inner surface of the plasma membrane (6, 7). Eukaryotic ankyrins have as many as 22 to 24 contiguous copies of a 33-amino-acid residue motif (-G-TPLH-AA--GH---V/A--LL--GA--N/D----). The presumed ankyrins from the *Pseudomonas* strains are much smaller than the eukaryotic proteins; that from *P. syringae* has only one domain with 11 of the 15 conserved amino acid residues of the ankyrin repeated motif (49), while the domain from the putative *P. aeruginosa* ankyrin has 12 of 15 residues (25, 42). The function of the prokaryotic ankyrins is unknown.

An alternative role for these DNA elements in *Pseudomonas* species could be as regions of homology involved in, or resulting from, mobilization of the associated DNA. It should be noted, however, that there is no sequence similarity to genetic elements such as integrons (37) or catabolic transposons (88) known to specify that activity. In *P. putida* F1 these elements flank genes encoding the complete catabolism of *p*-cymene through *p*-cumate to isobutyrate, pyruvate, and acetyl coenzyme A and may have been involved in the mobilization of those genes to their current location. It has been proposed (84) that catabolic genes may be clustered in defined regions of the chromosome in *P. putida*. Recombination between homologous genetic elements located downstream of the *cym* and *cmt* operons and upstream of the *tod* operon would have resulted in the clustering of the operons shown here and described in reference 24; this may suggest a general mechanism for the clustering of catabolic operons on the *Pseudomonas* chromosome.

A neighboring gene upstream of *cym*. By using the 4-kb *Hin*dIII-*Bgl*II fragment (Fig. 2, map coordinates 1.01 to 5.0) from pRE847 as a probe, *E. coli* JM109 carrying a recombinant plasmid (pRE856) containing an overlapping 22.9-kb *Bgl*II fragment (map coordinates −17.9 to 5.0) inserted in pLV59 was identified. From pRE856, several subclones were obtained (pRE890 to pRE895) (Table 1). The ends of the DNA fragments cloned in these plasmids were sequenced in order to identify neighboring genes which might help in the future to localize the *cym*, *cmt*, and *tod* operons in *P. putida* F1. One gene encoding a product having significant similarity to a published protein was identified (data not shown). The gene extends from map coordinate −8.9 to the left beyond −10.2 (within pRE891 and pRE894; the sequence of the gene was incomplete) and encodes a protein that has a high degree of identity (42% in the 441 amino acids compared) to DNA topoisomerase III of *E. coli* (21).

Evolution of catabolic pathways. *P. putida* F1 transforms *p*-cymene to *p*-cumate (Fig. 1). This involves three enzymes: a two-component monooxygenase, an alcohol dehydrogenase, and an aldehyde dehydrogenase. Although the monooxygenase and aldehyde dehydrogenase of the *p*-cymene catabolic pathway are related to the analogous enzymes of the TOL plasmid-encoded toluene-xylene catabolic pathway, the respective alcohol dehydrogenases are not related. Since the monooxygenases can catalyze the oxidation of both the methyl and the

benzylic alcohol groups (Fig. 1, enzyme A), it would not have been necessary initially for an organism to have an alcohol dehydrogenase in order to convert *p*-cymene, toluene, *m*-xylene, or *p*-xylene to *p*-cumate, benzoate, *m*-toluate, or *p*-toluate, respectively. After the *p*-cymene and toluene-xylene pathways branched away from each other, subsequent recruitment of alcohol dehydrogenases (Fig. 1, enzyme B) into each pathway would not only have allowed an increased rate of metabolism of *p*-cumic, benzyl, or toluic alcohol but also conversion of alcohol to aldehyde would furnish NADH rather than consume it. This translates to an increased yield of up to 6 mol of ATP per mol of substrate.

The functions of the proposed acyl coenzyme A synthetase and enoyl coenzyme A hydratase in relation to the metabolism of *p*-cymene remain unknown. The proteins involved in the oxidation of the methyl group of *p*-cymene to *p*-cumate may have their origins in a pathway for the catabolism of alkanes through fatty acids. The acyl coenzyme A synthetase and enoyl coenzyme A hydratase genes, while not directly involved in the catabolism of *p*-cymene, may simply be remnants of an alkane or fatty acid catabolic pathway. Such an origin has been suggested for two distantly related enzymes having a very different function, the dehalogenation of 4-chlorobenzoate (4). Thus, 4-chlorobenzoate-coenzyme A synthetase and 4-chlorobenzoyl-coenzyme A dehalogenase have a low level of similarity to acetyl coenzyme A synthetase and enoyl coenzyme A hydratase, respectively, (4) and may have evolved from a fatty acid β-oxidation pathway.

ACKNOWLEDGMENTS

DNA sequence data were obtained from ACGT, Inc., and the ICBR DNA Sequencing Core Laboratory, University of Florida, Gainesville. Synthesis of DNA primers for sequencing at the University of Florida was carried out by the ICBR DNA Synthesis Core Laboratory at the University of Florida. Peter Chapman provided a critical reading of the manuscript. Their help is appreciated.

REFERENCES

1. **Alting-Mees, M. A., and J. M. Short.** 1989. pBluescript II: gene mapping vectors. Nucleic Acids Res. **17:**9494.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
3. **Assinder, S. J., and P. A. Williams.** 1990. The TOL plasmids: determinants of the catabolism of toluene and the xylenes. Adv. Microb. Physiol. **31:**1–69.
4. **Babbitt, P. C., G. L. Kenyon, B. M. Martin, H. Charest, M. Sylvestre, J. D. Scholten, K.-H. Chang, P.-H. Liang, and D. Dunaway-Mariano.** 1992. Ancestry of the 4-chlorobenzoate dehalogenase: analysis of amino acid sequence identities among families of acyl:adenyl ligases, enoyl-coenzyme A hydratases/isomerases, and acyl-coenzyme A thioesterases. Biochemistry **31:** 5594–5604.
5. **Barnsley, E. A.** 1988. Metabolism of 2,6-dimethylnaphthalene by flavobacteria. Appl. Environ. Microbiol. **54:**428–433.
6. **Bennett, V.** 1990. Spectrin-based membrane skeleton: a multipotential adaptor between plasma membrane and cytoplasm. Physiol. Rev. **70:**1029–1065.
7. **Bennett, V.** 1992. Ankyrins. Adaptors between diverse plasma membrane proteins and the cytoplasm. J. Biol. Chem. **267:**8703–8706.
8. **Benson, D., D. J. Lipman, and J. Ostell.** 1993. GenBank. Nucleic Acids Res. **21:**2963–2965.
9. **Black, P. N.** 1991. Primary sequence of the *Escherichia coli fadL* gene encoding an outer membrane protein required for long-chain fatty acid transport. J. Bacteriol. **173:**435–442.
10. **Blattner, F. R., V. Burland, G. Plunkett III, H. J. Sofia, and D. L. Daniels.** 1993. Analysis of the *Escherichia coli* genome. IV. DNA sequence of the region from 89.2 to 92.8 minutes. Nucleic Acids Res. **21:**5408–5417.
11. **Brennan, R. G., and Brian W. Matthews.** 1989. The helix-turn-helix DNA binding motif. J. Biol. Chem. **264:**1903–1906.
12. **Brown, T. D. K., M. C. Jones-Mortimer, and H. L. Kornberg.** 1977. The enzymic interconversion of acetate and acetyl-coenzyme A in *Escherichia coli*. J. Gen. Microbiol. **102:**327–336.
13. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134:**1141–1156.
14. **Chang, C., and A. Yoshida.** 1994. Cloning and characterization of the gene encoding mouse mitochondrial aldehyde dehydrogenase. Gene **148:**331–336.
15. **Dagley, S., and M. D. Patel.** 1957. Oxidation of *p*-cresol and related compounds by a *Pseudomonas*. Biochem. J. **66:**227–233.
16. **Davis, J. B., and R. L. Raymond.** 1961. Oxidation of alkyl-substituted cyclic hydrocarbons by a *Nocardia* during growth on n-alkanes. Appl. Microbiol. **9:**383–388.
17. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **DeFrank, J. J., and D. W. Ribbons.** 1976. The *p*-cymene pathway in *Pseudomonas putida* PL: isolation of a dihydrodiol accumulated by a mutant. Biochem. Biophys. Res. Commun. **70:**1129–1135.
19. **DeFrank, J. J., and D. W. Ribbons.** 1977. *p*-Cymene pathway in *Pseudomonas putida*: initial reactions. J. Bacteriol. **129:**1356–1364.
20. **DeFrank, J. J., and D. W. Ribbons.** 1977. *p*-Cymene pathway in *Pseudomonas putida*: ring cleavage of 2,3-dihydroxy-*p*-cumate and subsequent reactions. J. Bacteriol. **129:**1365–1374.
21. **Digate, R. J., and K. J. Marians.** 1989. Molecular cloning and DNA sequence analysis of *Escherichia coli topB*, the gene encoding topoisomerase III. J. Biol. Chem. **264:**17924–17930.
22. **Dretzen, G., M. Bellard, P. Sassone-Corsi, and P. Chambon.** 1981. A reliable method for the recovery of DNA fragments from agarose and acrylamide gels. Anal. Biochem. **112:**295–298.
23. **Eaton, R. W.** 1994. Organization and evolution of naphthalene catabolic pathways: sequence of the DNA encoding 2-hydroxychromene-2-carboxylate isomerase and *trans-o*-hydroxybenzylidenepyruvate hydratase-aldolase from the NAH7 plasmid. J. Bacteriol. **176:**7757–7762.
24. **Eaton, R. W.** 1996. *p*-Cumate catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA carrying the *cmt* operon. J. Bacteriol. **178:**1351–1362.
25. **Eaton, R. W.** Unpublished observation.
26. **Eaton, R. W., and P. J. Chapman.** 1995. Formation of indigo and related compounds from indolecarboxylic acids by aromatic acid-degrading bacteria: chromogenic reactions for cloning genes encoding dioxygenases that act on aromatic acids. J. Bacteriol. **177:**6983–6988.
27. **Eaton, R. W., and P. J. Chapman.** Unpublished observation.
28. **Eaton, R. W., and J. S. Karns.** 1991. Cloning and analysis of *s*-triazine catabolic genes from *Pseudomonas* sp. strain NRRLB-12227. J. Bacteriol. **173:**1215–1222.
29. **Eaton, R. W., and D. W. Ribbons.** 1982. Metabolism of dibutylphthalate and phthalate by *Micrococcus* sp. strain 12B. J. Bacteriol. **151:**48–57.
30. **Eaton, R. W., and K. N. Timmis.** 1986. Characterization of a plasmid-specified pathway for catabolism of isopropylbenzene in *Pseudomonas putida* RE204. J. Bacteriol. **168:**123–131.
31. **Fox, B. G., J. Shanklin, J. Ai, T. M. Loehr, and J. Sanders-Loehr.** 1994. Resonance raman evidence for an Fe-O-Fe center in stearoyl-ACP desaturase. Primary sequence identity with other diiron-oxo proteins. Biochemistry **33:**12776–12786.
32. **Fukuyama, K., T. Hase, S. Matsumoto, T. Tsukihara, Y. Katsube, N. Tanaka, M. Kakudo, K. Wada, and H. Matsubara.** 1980. Structure of *S. platensis* [2Fe-2S] ferredoxin and evolution of chloroplast-type ferredoxins. Nature **286:**522–524.
33. **Gibson, D. T., J. R. Koch, and R. E. Kallio.** 1968. Oxidative degradation of aromatic hydrocarbons by microorganisms. I. Enzymatic formation of catechol from benzene. Biochemistry **7:**2653–2662.
34. **Gildemeister, E., and F. Hoffmann.** 1960. Aromatische Kohlenwasserstoffe: (a) Benzoide aromaten, *p*-Cymol, p. 402–411. *In* Die ätherischen Öle, vol. IIIa. Akademie-Verlag, Berlin, Germany.
35. **Grunstein, M., and D. Hogness.** 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72:**3961–3965.
36. **Habe, H., K. Kasuga, H. Nojiri, H. Yamane, and T. Omori.** 1996. Analysis of cumene (isopropylbenzene) degradation genes from *Pseudomonas fluorescens* IP01. Appl. Environ. Microbiol. **62:**4471–4477.
37. **Hall, R. M., and H. W. Stokes.** 1993. Integrons: novel DNA elements which capture genes by site-specific recombination. Genetica **90:**115–132.
38. **Harayama, S.** GenBank accession number D63341.
39. **Harayama, S., R. A. Leppik, M. Rekik, N. Mermod, P. R. Lehrbach, W. Reineke, and K. N. Timmis.** 1986. Gene order of the TOL catabolic plasmid upper pathway operon and oxidation of both toluene and benzyl alcohol by the *xylA* product. J. Bacteriol. **167:**455–461.
40. **Harayama, S., M. Rekik, M. Wubbolts, K. Rose, R. A. Leppik, and K. N. Timmis.** 1989. Characterization of five genes in the upper-pathway operon of TOL plasmid pWW0 from *Pseudomonas putida* and identification of the gene products. J. Bacteriol. **171:**5048–5055.
41. **Harley, C. B., and R. P. Reynolds.** 1987. Analysis of *E. coli* promoter sequences. Nucleic Acids Res. **15:**2347–2351.
42. **Hassett, D. J., M. G. Klotz, and M. L. Howell.** 1996. GenBank accession number U59457.
43. **Hempel, J., H. Nicholas, and R. Lindahl.** 1993. Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework. Protein Sci. **2:**1890–1900.
44. **Herrmann, H., C. Müller, I. Schmidt, J. Mahnke, L. Petruschka, and K.**

Hahnke. 1995. Localization and organization of phenol degradation genes of *Pseudomonas putida* strain H. Mol. Gen. Genet. **247:**240–246.
45. **Inoue, J., J. P. Shaw, M. Rekik, and S. Harayama.** 1995. Overlapping substrate specificities of benzaldehyde dehydrogenase (the *xylC* gene product) and 2-hydroxymuconic semialdehyde dehydrogenase (the *xylG* gene product) encoded by TOL plasmid pWW0 of *Pseudomonas putida*. J. Bacteriol. **177:** 1196–1201.
46. **Kaneko, T., A. Tanaka, S. Sato, H. Kotani, T. Sazuka, N. Miyajima, M. Sugiura, and S. Tabata.** 1995. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb region from map positions 64% to 92% of the genome. DNA Res. **2:**153–166.
47. **Karplus, P. A., K. A. Walsh, and J. R. Herriott.** 1984. Amino acid sequence of spinach ferredoxin:$NADP^+$ oxidoreductase. Biochemistry **23:**6576–6583.
48. **Khramtsov, N. V., D. S. Blunt, B. L. Monteleone, and S. J. Upton.** 1996. The putative acetyl-coenzyme A synthetase gene of *Cryptosporidium parvum* and a new conserved protein motif in acetyl-coenzyme A synthetases. J. Parasitol. **82:**423–427.
49. **Klotz, M. G., and A. J. Anderson.** 1995. Sequence of a gene encoding periplasmic *Pseudomonas syringae* ankyrin. Gene **164:**187–188.
50. **Kok, M., R. Oldenhuis, M. P. G. van der Linden, P. Raatjes, J. Kingma, P. H. van Lelyveld, and B. Witholt.** 1989. The *Pseudomonas oleovorans* alkane hydroxylase gene: sequence and expression. J. Biol. Chem. **262:**5435–5441.
51. **Leavitt, R. I.** 1967. Microbial oxidation of hydrocarbons. Oxidation of *p*-isopropyltoluene by a *Pseudomonas* sp. J. Gen. Microbiol. **49:**411–420.
52. **Madhyastha, K. M., and P. K. Bhattacharyya.** 1968. Microbiological transformations of terpenes. Part XII. Fermentation of *p*-cymene by a soil pseudomonad (PL-strain). Indian J. Biochem. **5:**102–111.
53. **Madhyastha, K. M., P. N. Rangachari, M. R. Rao, and P. K. Bhattacharyya.** 1968. Microbiological transformations of terpenes. Part XV. Enzyme systems in the catabolism of *p*-cymene in PL-strain. Indian J. Biochem. **5:**167–173.
54. **Margolin, W., D. Bramhill, and S. R. Long.** 1995. The *dnaA* gene of *Rhizobium meliloti* lies within an unusual gene arrangement. J. Bacteriol. **177:** 2892–2900.
55. **Martin, R. R., V. D. Marshall, J. R. Sokatch, and L. Unger.** 1973. Common enzymes of branched-chain amino acid catabolism in *Pseudomonas putida*. J. Bacteriol. **115:**198–204.
56. **Mason, J. R., and R. Cammack.** 1992. The electron-transport proteins of hydroxylating bacterial dioxygenases. Annu. Rev. Microbiol. **46:**277–305.
57. **Minami-Ishii, N., S. Taketani, T. Osumi, and T. Hashimoto.** 1989. Molecular cloning and sequence analysis of the cDNA for rat mitochondrial enoyl-coenzyme A hydratase. Structural and evolutionary relationships linked to the bifunctional enzyme of the peroxisomal beta-oxidation system. Eur. J. Biochem. **185:**73–78.
58. **Mitamura, T., R. V. Ebora, T. Nakai, Y. Makino, S. Negoro, I. Urabe, and H. Okada.** 1990. Structure of isozyme genes of glucose dehydrogenase from *Bacillus megaterium* IAM1030. J. Ferment. Bioeng. **70:**363–369.
59. **Miyachi, N., T. Tanaka, T. Suzuki, Y. Hotta, and T. Omori.** 1993. Microbial oxidation of dimethylnaphthalene isomers. Appl. Environ. Microbiol. **59:** 1504–1506.
60. **Munson, R., Jr., and S. Grass.** 1988. Purification, cloning, and sequencing of outer membrane protein P1 of *Haemophilus influenzae* type b. Infect. Immun. **56:**2235–2242.
61. **Nakagawa, J., H. Waldner, S. Meyer-Monard, J. Hofsteenge, P. Jeno, and C. Moroni.** 1995. AUH, a gene encoding an AU-specific RNA binding protein with intrinsic enoyl-coenzyme A hydratase activity. Proc. Natl. Acad. Sci. USA **92:**2051–2055.
62. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48:**443–453.
63. **Neidle, E. L., C. Hartnett, L. N. Ornston, A. Bairoch, M. Rekik, and S. Harayama.** 1991. Nucleotide sequences of the *Acinetobacter calcoaceticus benABC* genes for benzoate 1,2-dioxygenase reveal evolutionary relationships among multicomponent oxygenases. J. Bacteriol. **173:**5385–5395.
64. **Neumann-Spallart, C., M. Brandtner, M. Kraus, J. Jakowitsch, M. G. Bayer, T. L. Maier, H. E. Schenk, and W. Loffelhardt.** 1990. The *petFI* gene encoding ferredoxin I is located close to the *str* operon on the cyanelle genome of *Cyanophora paradoxa*. FEBS Lett. **268:**55–58.
65. **O'Connor, C. D., and G. O. Humphries.** 1982. Expression of the *Eco*RI restriction-modification system and the construction of positive-selection cloning vectors. Gene **20:**219–229.
66. **Oliver, D.** 1985. Protein secretion in *Escherichia coli*. Annu. Rev. Microbiol. **39:**615–648.
67. **Pan, W., and B. G. Spratt.** 1994. Regulation of the permeability of the gonococcal cell envelope by the *mtr* system. Mol. Microbiol. **11:**769–775.
68. **Persson, B., M. Krook, and H. Jörnvall.** 1991. Characteristics of short-chain alcohol dehydrogenases and related enzymes. Eur. J. Biochem. **200:**537–543.
69. **Porter, T. D., and C. B. Kasper.** 1986. NADPH-cytochrome P-450 oxidoreductase: flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins. Biochemistry **25:**1682–1687.
70. **Radke, M.** 1987. Organic geochemistry of aromatic hydrocarbons. Adv. Petroleum Geochem. **2:**141–207.
71. **Raymond, R. L., V. W. Jamison, and J. O. Hudson.** 1967. Microbial hydrocarbon co-oxidation. I. Oxidation of mono- and dicyclic hydrocarbons by soil isolates of the genus *Nocardia*. Appl. Microbiol. **15:**857–865.
72. **Reid, M. F., and C. A. Fewson.** 1994. Molecular characterization of microbial alcohol dehydrogenases. Crit. Rev. Microbiol. **20:**13–56.
73. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
74. **Saraste, M., P. R. Sibbald, and A. Wittinghofer.** 1990. The P-loop—a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. **15:**430–434.
75. **Shanklin, J., E. Whittle, and B. G. Fox.** 1994. Eight histidine residues are catalytically essential in a membrane-associated iron enzyme, stearoyl-coenzyme A desaturase, and are conserved in alkane hydroxylase and xylene monooxygenase. Biochemistry **33:**12787–12794.
76. **Shaw, J. P., and S. Harayama.** 1990. Purification and characterization of TOL plasmid-encoded benzyl alcohol dehydrogenase and benzaldehyde dehydrogenase of *Pseudomonas putida*. Eur. J. Biochem. **191:**705–714.
77. **Shaw, J. P., and S. Harayama.** 1992. Purification and characterization of the NADH:acceptor reductase component of xylene monooxygenase encoded by the TOL plasmid pWW0 of *Pseudomonas putida* mt-2. Eur. J. Biochem. **209:**51–61.
78. **Shaw, J. P., M. Rekik, F. Schwager, and S. Harayama.** 1993. Kinetic studies on benzyl alcohol dehydrogenase encoded by TOL plasmid pWW0. A member of the zinc-containing long chain alcohol dehydrogenase family. J. Biol. Chem. **268:**10842–10850.
79. **Shen, Z., and D. M. Byers.** 1996. Isolation of *Vibrio harveyi* acyl carrier protein and the *fabG*, *acpP*, and *fabF* genes involved in fatty acid biosynthesis. J. Bacteriol. **178:**571–573.
80. **Simon, M. J., T. D. Osslund, R. Saunders, B. D. Ensley, S. Suggs, A. Harcourt, W.-C. Suen, D. T. Gibson, and G. J. Zylstra.** 1993. Sequence of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4. Gene **127:**31–37.
81. **Suzuki, M., T. Hayakawa, J. P. Shaw, M. Rekik, and S. Harayama.** 1991. Primary structure of xylene monooxygenase: similarities to and differences from the alkane hydroxylation system. J. Bacteriol. **173:**1690–1695.
82. **Van Beilen, J. B., D. Penninga, and B. Witholt.** 1992. Topology of the membrane-bound alkane hydroxylase of *Pseudomonas oleovorans*. J. Biol. Chem. **267:**9194–9201.
83. **Wang, Y., M. Rawlings, D. T. Gibson, D. Labbé, H. Bergeron, R. Brousseau, and P. C. K. Lau.** 1995. Identification of a membrane protein and a truncated LysR-type regulator associated with the toluene degradation pathway in *Pseudomonas putida* F1. Mol. Gen. Genet. **246:**570–579.
84. **Wheelis, M. L.** 1975. The genetics of dissimilarity pathways in *Pseudomonas*. Annu. Rev. Microbiol. **29:**505–524.
85. **Wierenga, R. K., P. Terpstra, and W. G. J. Hol.** 1986. Prediction of the occurrence of the ADP-binding βαβ-fold in proteins, using an amino acid sequence fingerprint. J. Mol. Biol. **187:**101–107.
86. **Wigmore, G. J., A. J. Smith, and D. W. Ribbons.** 1977. Metabolism of *para*-substituted benzoates and toluenes with "2,3-dihydroxybenzoates" as ring cleavage substrates, abstr. K-188. *In* Abstracts of the 77th Annual Meeting of the American Society for Microbiology 1977. American Society for Microbiology, Washington, D.C.
87. **Wilbur, W. J., and D. J. Lipman.** 1983. Rapid similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA **80:**726–730.
88. **Wyndham, R. C., A. E. Cashore, C. H. Nakatsu, and M. Peel.** 1994. Catabolic transposons. Biodegradation **5:**323–342.
89. **Yamada, K., S. Horiguchi, and J. Takahashi.** 1965. Studies on the utilization of hydrocarbons by microorganisms. VI. Screening of aromatic hydrocarbon-utilizing microorganisms and cumic acid formation from *p*-cymene. Agric. Biol. Chem. **29:**943.
90. **Yanisch-Perron, C., J. Viera, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33:**103–119.
91. **Yeh, W. K., D. T. Gibson, and T.-N. Liu.** 1977. Toluene dioxygenase: a multicomponent enzyme system. Biochem. Biophys. Res. Commun. **78:**401–410.
92. **Zylstra, G. J., and D. T. Gibson.** 1989. Toluene degradation by *Pseudomonas putida* F1. Nucleotide sequence of the *todC1C1BADE* genes and their expression in *Escherichia coli*. J. Biol. Chem. **264:**14940–14946.
93. **Zylstra, G. J., W. R. McCombie, D. T. Gibson, and B. A. Finette.** 1988. Toluene degradation by *Pseudomonas putida* F1: genetic organization of the *tod* operon. Appl. Environ. Microbiol. **54:**1498–1503.