JOURNAL OF BACTERIOLOGY, June 1998, p. 2915–2923 Vol. 180, No. 11
0021-9193/98/$04.00+0

Characterization of the *hca* Cluster Encoding the Dioxygenolytic Pathway for Initial Catabolism of 3-Phenylpropionic Acid in *Escherichia coli* K-12

EDUARDO DÍAZ,* ABEL FERRÁNDEZ, AND JOSÉ L. GARCÍA

Department of Molecular Microbiology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain

Received 14 January 1998/Accepted 26 March 1998

We have identified, cloned, and sequenced the *hca* cluster encoding the dioxygenolytic pathway for initial catabolism of 3-phenylpropionic acid (PP) in *Escherichia coli* K-12. This cluster maps at min 57.5 of the chromosome and is composed of five catabolic genes arranged as a putative operon (*hcaA1A2CBD*) and two additional genes transcribed in the opposite direction that encode a potential permease (*hcaT*) and a regulator (*hcaR*). Sequence comparisons revealed that while *hcaA1A2CD* genes encode the four subunits of the 3-phenylpropionate dioxygenase, the *hcaB* gene codes for the corresponding *cis*-dihydrodiol dehydrogenase. This type of catabolic module is homologous to those encoding class IIB dioxygenases and becomes the first example of such a catabolic cluster in *E. coli*. The inducible expression of the *hca* genes requires the presence of the *hcaR* gene product, which acts as a transcriptional activator and shows significant sequence similarity to members of the LysR family of regulators. Interestingly, the HcaA1A2CD and HcaB enzymes are able to oxidize not only PP to 3-(2,3-dihydroxyphenyl)propionate (DHPP) but also cinnamic acid (CI) to its corresponding 2,3-dihydroxy derivative. Further catabolism of DHPP requires the *mhp*-encoded *meta* fission pathway for the mineralization of 3-hydroxyphenylpropionate (3HPP) (A. Ferrández, J. L. García, and E. Díaz, J. Bacteriol. 179:2573–2581, 1997). Expression in *Salmonella typhimurium* of the *mhp* genes alone or in combination with the *hca* cluster allowed the growth of the recombinant bacteria in 3-hydroxycinnamic acid (3HCI) and CI, respectively. Thus, the convergent *mhp*- and *hca*-encoded pathways are also functional in *S. typhimurium*, and they are responsible for the catabolism of different phenylpropanoid compounds (3HPP, 3HCI, PP, and CI) widely available in nature.

Phenylpropanoid compounds are widely available in natural environments, and they can originate from putrefaction of proteins in soil or as breakdown products of several constituents of plants, such as lignin, various oils, and resins (2, 6, 14, 20). Microbial catabolism of phenylpropanoid compounds plays an important role not only in the natural degradative cycle of these aromatic molecules but also in their industrial applications such as wine making, aging, and storage (13). In particular, degradation of cinnamic acid (CI), 3-phenylpropionic acid (PP), and their hydroxylated derivatives has been reported in several bacteria, including *Acinetobacter* sp. (14), *Pseudomonas* sp. (2, 51), *Arthrobacter* sp. (51), *Escherichia coli* (10), and *Rhodococcus globerulus* (6). Although most of the intermediates of these pathways are known, there has been little genetic characterization of these degradative routes, with the exception of the 3-(3-hydroxyphenyl)propionate (3HPP) catabolic pathways of *E. coli* K-12 (20) and *R. globerulus* PWD1 (6).

Biochemical studies and the isolation and characterization of mutants defective in the catabolism of PP and 3HPP (compounds I and IV in Fig. 1B, respectively) revealed that in *E. coli* the aerobic degradation of these compounds proceeds by two initially separate routes that converge into 3-(2,3-dihydroxyphenyl)propionate (DHPP) (compound III), which suffers an extradiol ring cleavage and is ultimately degraded to Krebs cycle intermediates (9–11) (Fig. 1B). The cloning, sequencing, and transcriptional regulation of the *meta* fission cluster for the catabolism of 3HPP in *E. coli* K-12 have been recently reported (20).

The catabolism of PP in *E. coli* is initiated by a dioxygenolytic pathway (10, 11) (Fig. 1B). The first step is catalyzed by a 3-phenylpropionate dioxygenase, which inserts both atoms of molecular oxygen into positions 2 and 3 of the phenyl ring of PP, yielding *cis*-3-(3-carboxyethyl)-3,5-cyclohexadiene-1,2-diol (PP-dihydrodiol; compound II), which is subsequently oxidized by the 3-phenylpropionate-dihydrodiol dehydrogenase to give DHPP (compound III) (10, 11) (Fig. 1B). Enzyme assays and respirometry showed that the syntheses of enzymes required to convert the two initial growth substrates, PP and 3HPP, into DHPP are inducible and under separate control (10, 11). Very recently, it has been shown that in batch cultures the utilization of PP was immediately repressed by glucose (30).

Here we present the cloning, genetic characterization, and mechanism of regulation of the *hca* genes encoding the complete dioxygenolytic pathway for the catabolism of PP in *E. coli* K-12. This work constitutes the first genetic characterization of such a pathway and represents the first report of a gene cluster encoding a phenyl ring hydroxylating dioxygenase from *E. coli*. Moreover, we provide experimental evidence that 3HPP and PP catabolic pathways are also responsible for the catabolism of 3-hydroxycinnamic acid (3HCI) and CI, respectively.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The *E. coli* K-12 strains used were MC1061 [F$^-$ *hsdR mcrB araD139* Δ(*araABC-leu*)*7679* Δ*lacX74 galU galK rpsL thi*] (46), DH5α [F$^-$ *endA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1 recA1 gyrA relA1* Δ(*argF-lac*)*U169 deoR* φ80*lacZ*ΔM15] (46), and MG1655 (F$^-$ λ$^-$) (4). *E. coli* ED1061 is an *hcaA1* mutant of *E. coli* MC1061 (this study). The other strain

* Corresponding author. Mailing address: Department of Molecular Microbiology, Centro de Investigaciones Biológicas, Velázquez 144, 28006 Madrid, Spain. Phone: 34-1-5611800. Fax: 34-1-5627518. E-mail: cibdf4f@fresno.csic.es.

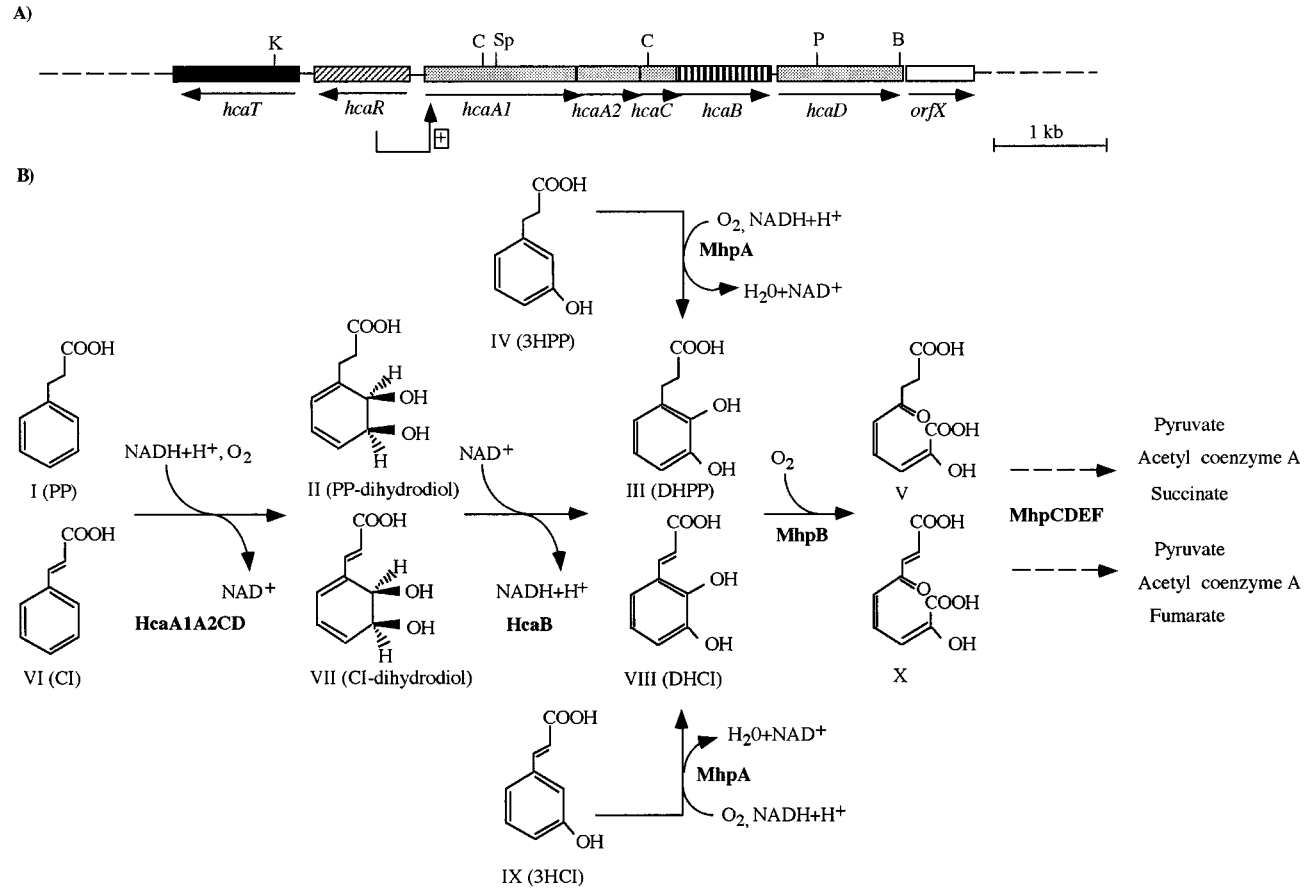

FIG. 1. Convergent pathways for the catabolism of PP (CI) and 3HPP (3HCI) in *E. coli*. (A) Physical and genetic map of the chromosomal *hca* region. The locations of the genes are shown relative to those of some relevant restriction endonuclease sites, i.e., *Bam*HI (B), *Cla*I (C), *Kpn*I (K), *Pst*I (P), and *Sph*I (Sp). Arrows indicate the directions of gene transcription. The boxed plus sign indicates stimulation of gene expression by the *hcaR* gene product in the presence of PP. Genes with similar shadings encode subunits of the same protein. (B) Proposed biochemistry of the PP (3HPP) and CI (3HCI) catabolic pathways. HcaA1A2CD and HcaB are the enzymes encoded by the corresponding *hca* structural genes. MhpA to MhpF are the enzymes for the catabolism of 3HPP (3HCI) and DHPP (DHCI). The metabolites are PP (compound I), *cis*-3-(3-carboxyethyl)-3,5-cyclohexadiene-1,2-diol (compound II), DHPP (compound III), 3HPP (compound IV), 2-hydroxy-6-ketononadienedioate (compound V), CI (compound VI), *cis*-3-(3-carboxyethenyl)-3,5-cyclohexadiene-1,2-diol (compound VII), DHCI (compound VIII), 3HCI (compound IX), and 2-hydroxy-6-ketononatrienedioate (compound X). Enzymes: HcaA1A2CD, 3-phenylpropionate dioxygenase; HcaB, 3-phenylpropionate-dihydrodiol dehydrogenase; MhpA, 3-(3-hydroxyphenyl)propionate hydroxylase; MhpB, 3-(2,3-dihydroxyphenyl)propionate 1,2-dioxygenase; MhpC, 2-hydroxy-6-ketonona-2,4-dienedioate hydrolase; MhpD, 2-keto-4-pentenoate hydratase; MhpE, 4-hydroxy-2-ketovalerate aldolase; MhpF, acetaldehyde dehydrogenase (acylating).

used in this study was *Salmonella typhimurium* LT-2 (20). For cloning and expression purposes we have used two chloramphenicol (CM) resistance low-copy-number cloning vectors, plasmids pCK01 (20) and pVTR-B (41), as well as the pUC18, pUC19 (46), and pUC18Not (20) vectors. Plasmid pPADR2 is an RSF1010-based promiscuous plasmid containing the complete *mhp* pathway (20). Plasmids pUC4K (Pharmacia) and pMAK700 (24) were used for insertional inactivation of the *hcaA1* gene (former *orfA*) and construction of the strain *E. coli* ED1061, respectively. Unless otherwise stated, bacteria were grown in Luria-Bertani (LB) medium (46) at 37°C. When used as carbon sources, aromatic acids were supplied at 1 mM (CI) or 5 mM (PP, 3HPP, and 3HCI) to M63 minimal medium (36), and the cultures were incubated at 30°C (*E. coli*) or 37°C (*S. typhimurium*). To perform the red color formation assay, cells were incubated at 30°C (*E. coli*) or 37°C (*S. typhimurium*) on LB medium containing 5 mM PP. Where appropriate, antibiotics were added at the following concentrations: ampicillin (AP), 100 μg/ml; CM, 35 μg/ml; kanamycin (KM), 50 μg/ml.

DNA manipulations and sequencing. Plasmid DNA was prepared by the rapid alkaline lysis method (46). Transformation of *E. coli* was carried out by the RbCl method (46). Electroporation (Gene Pulser, Bio-Rad) was used for *S. typhimurium*. DNA manipulations and other molecular biology techniques were essentially as described elsewhere (46). DNA fragments were purified by using low-melting-point agarose. 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 (47). 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 with a 377 automated DNA sequencer (Applied Biosystems Inc.). Sequences were extended by designing primers based on the already-determined sequence.

Sequence data analyses. Nucleotide sequence analyses were done with the DNA-Strider 1.2 program. Amino acid sequences were analyzed with Protein Analysis Tools on 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 (1) via the National Institute for Biotechnology Information server. Pairwise and multiple protein sequence alignments were made with the ALIGN (59) and CLUSTAL W (56) programs, respectively, on the Baylor College of Medicine Human Genome Center server. The *E. coli* database collection ECDC (31) was accessed via the Internet.

Insertional inactivation of the *hcaA1* gene (*orfA*) and construction of *E. coli* ED1061. The 3′-end-truncated *hcaA1* gene (former *orfA*) was PCR amplified from the chromosome of *E. coli* MC1061 by using oligonucleotides HCA5 (5′-CCGAATTCACATATTAGCAACCAACCAGC-3′ [the sequence corresponds to nucleotides 2384 to 2407 in Fig. 3; the engineered *Eco*RI site is underlined]) and HCA3 (5′-CCCTGCAGGTAAGCGGCGGTTTTATC-3′ [the sequence corresponds to nucleotides 3790 to 3815 in Fig. 3; the engineered *Pst*I site is underlined]) as primers. The 1.4-kb amplified fragment was digested with *Eco*RI-*Pst*I and cloned into the *Eco*RI-*Pst*I double-digested pUC19 vector to form pHCA. The *orfA* in pHCA was inactivated by the insertion at its *Cla*I restriction site of a 1.3-kb *Acc*I KM resistance cassette from plasmid pUC4K. The disrupted *orfA* was then subcloned as a 1.9-kb *Eco*RI-*Sph*I blunt-ended fragment into the blunt-ended-*Sph*I-digested pMAK700 plasmid, a pSC101-derived temperature-sensitive replicon (24), to form pHCA700. Because this plasmid replicates at 30°C but not at 44°C, it was possible to identify its integration through homol-

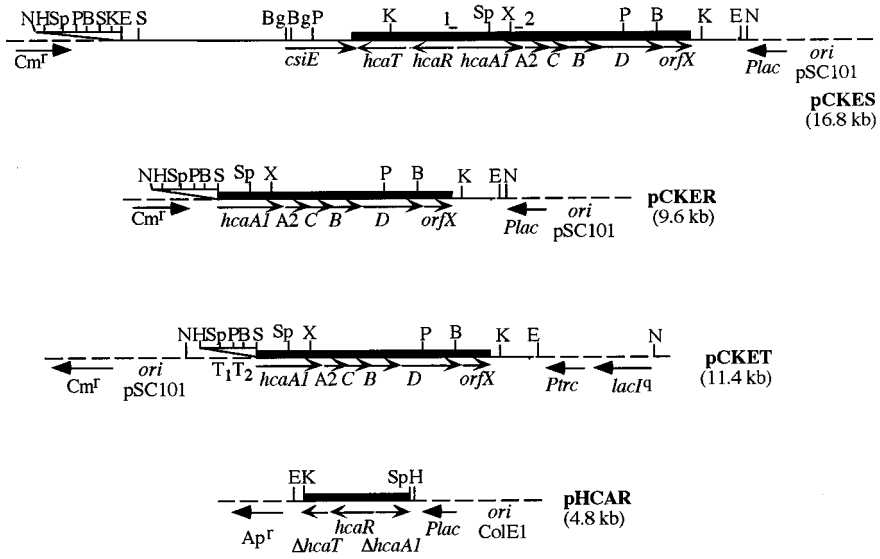

FIG. 2. Schematic representation of the subcloning and expression of the regulatory and catabolic *hca* genes. The subcloning strategies are described in detail in Materials and Methods. The relevant elements and restriction sites are indicated. The thick line represents the DNA fragment whose sequence is shown in Fig. 3. Vector-derived sequences are indicated by dashed lines. The *Plac* and the *Ptrc* promoters and direction of transcription are indicated (arrows). Δ, truncated gene. T_1 and T_2 are the transcriptional terminators of the *E. coli rrnB* operon (41). 1 and 2, oligonucleotides HCAR and HCA3, which were used as primers for the PCR to construct plasmid pCKER. The region encoding the replication (*ori*) function is also indicated. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; N, *Not*I; P, *Pst*I; S, *Sma*I; Sp, *Sph*I; X, *Xho*I. Ap^r and Cm^r, genes conferring resistance to AP and CM, respectively.

ogous recombination into the chromosome of *E. coli* MC1061 by selecting for KM resistance at 44°C. Since replication from the plasmid origin is deleterious to the host cell, when the cointegrates were subsequently grown at 30°C a second recombination event occurred, regenerating free plasmid in the cell and a disrupted *orfA* in the chromosome. The resident plasmid was then cured from the mutant strain by growing the cells at 44°C in LB medium without the antibiotic resistance marker of the vector, i.e., CM. A KM-resistant and CM-sensitive mutant strain was selected and named *E. coli* ED1061.

Construction of plasmids. To isolate the *hca* cluster, we constructed an *Eco*RI DNA library of *E. coli* MC1061 into pUC18 using as host the mutant strain ED1061, and then the transformants were screened for their ability to recover the dark red phenotype of *E. coli* MC1061 due to DHPP accumulation when growing on PP-containing LB medium. All red colonies harbored the plasmid pHCAES containing a 13.2-kb *Eco*RI insert. However, since plasmid pHCAES was shown to be highly unstable after several rounds of cultivation, the 13.2-kb *Eco*RI fragment was subcloned into the low-copy-number vector pCK01. The resulting plasmid pCKES (Fig. 2) also conferred on *E. coli* ED1061 the ability to produce the dark red color when growing on PP-containing LB medium, and this phenotype was stably maintained when the cells were grown in the presence of CM.

The *hcaR* gene was isolated from plasmid pCKES as a *Kpn*I-*Sph*I 2.1-kb fragment and subcloned into the *Kpn*I-*Sph*I double-digested pUC19 cloning vector to form pHCAR (Fig. 2). To delete genes *hcaRT* from the *hca* cluster, an *hcaA1* gene truncated at its 3′ end was PCR amplified from plasmid pCKES by using primers HCAR (5′-CCCCCGGGCCGTAGTTCCATCACCTTC-3′ [the sequence corresponds to nucleotides 2319 to 2338 in Fig. 3; the engineered *Sma*I restriction site is underlined]) and HCA3 (5′-CCCTGCAGGTAAGCGGCGG TTTTATC-3′ [the sequence corresponds to nucleotides 3790 to 3815 in Fig. 3). The 1.5-kb PCR product was digested with *Sma*I and *Xho*I, gel purified, and ligated to the *Sma*I-*Xho*I double-digested pCKES plasmid to form pCKER (Fig. 2). Since we have observed constitutive expression of the *hca* catabolic genes in *E. coli* ED1061(pCKER) cells, to avoid the possible transcriptional readthrough from the promoter of the CM resistance gene of the vector (Fig. 2), we subcloned the 6.0-kb *Eco*RI-*Hin*dIII fragment carrying the *hcaA1A2CBDorfX* genes into the *Eco*RI-*Hin*dIII double-digested low-copy-number pVTR-B vector. The resulting plasmid pCKET contained the catabolic *hca* genes together with its potential promoter region, downstream of the strong T_1T_2 transcriptional terminators of the *E. coli rrnB* operon (41) (Fig. 2), thus excluding additional expression signals from the vector.

Resting-cell reactions. Phenylpropionate dioxygenase and phenylpropionate-dihydrodiol dehydrogenase activities were checked by analyzing the formation of DHPP in resting-cell assays. Thus, cultures of *E. coli* or *S. typhimurium* were grown overnight in LB medium and then diluted into fresh medium in the presence or absence of 1 mM aromatic inducer (PP or CI) to an optical density at 600 nm of about 0.08. Growth was resumed at 30°C (*E. coli*) or 37°C (*S. typhimurium*) until the cultures reached an optical density at 600 nm of about 0.8. The cell cultures were then centrifuged at 3,000 × *g* for 10 min at 20°C, and cells were washed and resuspended in a 0.05 volume of M63 minimal medium. The resting-cell reaction was performed in a final volume of 5 ml containing 4.5 ml of M63 minimal medium supplemented with 1 mM glucose and 0.5 ml of the cell suspension. The reaction was started by the addition of 0.5 mM PP or CI, 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 for 5 min at 10,000 × *g* to remove the cells. Products accumulated in the supernatant were analyzed with Gilson high-pressure liquid chromatography (HPLC) equipment using a Lichrosphere 5 RP-8 column (150 by 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 22.3, 15.9, 6.20, and 4.63 min, corresponding to those of authentic standard CI, PP, 2,3-dihydroxicinnamic acid (DHCI), and DHPP, respectively, were monitored at 210 nm.

To confirm the formation of DHCI, ^1H nuclear magnetic resonance (NMR) spectra were recorded in CD_3OD at 30°C on a Varian Unity 500 spectrometer. ^1H chemical shifts were referenced to internal residual CHD_2OD.

Nucleotide sequence accession numbers. The nucleotide sequences reported in this study have been submitted to the GenBank/EMBL data bank under accession numbers Y11070 and Y11071.

RESULTS AND DISCUSSION

Cloning of the *hca* catabolic cluster of *E. coli* K-12. Although during the course of this work the complete genome sequence of *E. coli* K-12 was reported (8), at the beginning of this research the analysis of the current *E. coli* database collection (ECDC release 27) (31) revealed the existence of an unmapped 4.6-kb sequence (accession no. Z37966), containing a 1.3-kb open reading frame (ORF) (*orfA*) that coded for a product showing significant similarity to the large terminal subunit of some multicomponent aromatic-ring initial dioxygenases (12). Since it was reported that the catabolism of PP in *E. coli* proceeds via dioxygenolytic attack of the ring (10, 11) (Fig. 1B), we assumed that *orfA* could encode a component of the 3-phenylpropionic acid (hydrocinnamic acid) initial dioxygenase (Hca dioxygenase). To test this assumption, we constructed *E. coli* MC1061 mutants by the insertion of a KM resistance cassette within *orfA* (see Materials and Methods). The selection of the mutant strains was based on the previous observation that accumulation of DHPP on rich medium gen-

TABLE 1. *hca* expression in *S. typhimurium* LT-2 and *E. coli* ED1061

Strain[a]	Inducer	Conversion of PP to DHPP (mol%)[b]
LT-2	PP	BD
LT-2(pCKET, pHCAR)	None	2
LT-2(pCKET, pHCAR)	PP	10
ED1061	PP	BD
ED1061(pCKES)	None	2
ED1061(pCKES)	PP	80

[a] Plasmids pCKET, pHCAR, and pCKES are diagrammed in Fig. 2.
[b] Expression of the *hca* catabolic genes was monitored in resting-cell assays by measuring PP consumption and DHPP formation in HPLC. Cells were grown in LB medium in the presence or absence of 1 mM PP (inducer), and the resting-cell assays were performed for 60 min (*S. typhimurium* LT-2 cells) or 10 min (*E. coli* ED1061 cells) as described in Materials and Methods with 0.5 mM PP as substrate. BD, below detection limits.

erates a reddish-brown color due to autooxidation of this aromatic compound to the corresponding quinones and semiquinones (11). Thus, an *E. coli* strain such as MC1061, which contains a chromosomal deletion (Δ*lacX74*) spanning the initial genes of the *mhp* cluster responsible for the catabolism of DHPP (20), formed a dark red color when grown on PP-containing LB medium due to DHPP accumulation. Interestingly, the MC1061 KM-resistant mutants did not show the PP-dependent color reaction, indicating that the disrupted *orfA* was involved in the initial dioxygenation of PP. One of these mutants was selected, and it is referred to hereafter as *E. coli* ED1061.

To genetically characterize the complete Hca dioxygenase and the following enzyme of the dioxygenolytic pathway, i.e., the 3-phenylpropionate-dihydrodiol dehydrogenase (HcaB) (Fig. 1B), we constructed plasmid pCKES (Fig. 2) (see Materials and Methods), which contains a 13.2-kb *Eco*RI fragment from *E. coli* MC1061 that confers on *E. coli* ED1061 the ability to produce the dark red color during growth on PP-containing LB medium. Interestingly, *S. typhimurium* LT-2, which does not attack PP (see below) and lacks the *mhp* genes (20), showed the typical red color when transformed with pCKES on PP-containing LB medium. To directly assay the oxidation of PP to DHPP, resting-cell reactions were performed in the presence of PP, and the resultant products accumulated in the supernatants were analyzed by HPLC. *E. coli* mutant ED1061 and *S. typhimurium* LT-2 cells grown in LB medium in the presence of 1 mM PP were used as control bacteria since they did not attack this aromatic compound (Table 1). In contrast, resting cells of *S. typhimurium* LT-2(pCKES) and *E. coli* ED1061(pCKES) bacteria grown in LB medium in the presence of 1 mM PP rapidly consumed this aromatic compound and produced a metabolite that was eluted by HPLC as standard DHPP (Table 1). Therefore, all these data indicated that the 13.2-kb *Eco*RI DNA fragment encoded the complete dioxygenolytic pathway for the conversion of PP into DHPP; this pathway is also functional in *Salmonella.*

Structural analysis of the *hca* genes. The nucleotide sequence of a 7,259-bp DNA fragment that carries the *hca* cluster was determined (Fig. 3). Analyses of the ORFs and sequence comparisons (see below) suggested the existence of seven genes arranged as follows: (i) five genes encoding the 3-phenylpropionate dioxygenase (*hcaA1A2CD*) and 3-phenylpropionate-dihydrodiol dehydrogenase (*hcaB*), (ii) a regulatory gene (*hcaR*), and (iii) a gene (*hcaT*) that might encode a transporter. The *hca* genes are located on the chromosome in the order *hcaTRA1A2CBD* (Fig. 1A and 2). Downstream of the *hcaD* gene, the closely linked *orfX* (Fig. 1A and 3) could also be a member of the *hca* cluster. Interestingly, all five catabolic genes and *orfX* appear to be transcribed in the same direction. The Shine-Dalgarno sequences of *hcaA1*, *hcaA2*, *hcaC*, *hcaB*, and *hcaD* overlap the preceding ORFs (Fig. 3), suggesting that translational coupling occurs (23). Further-

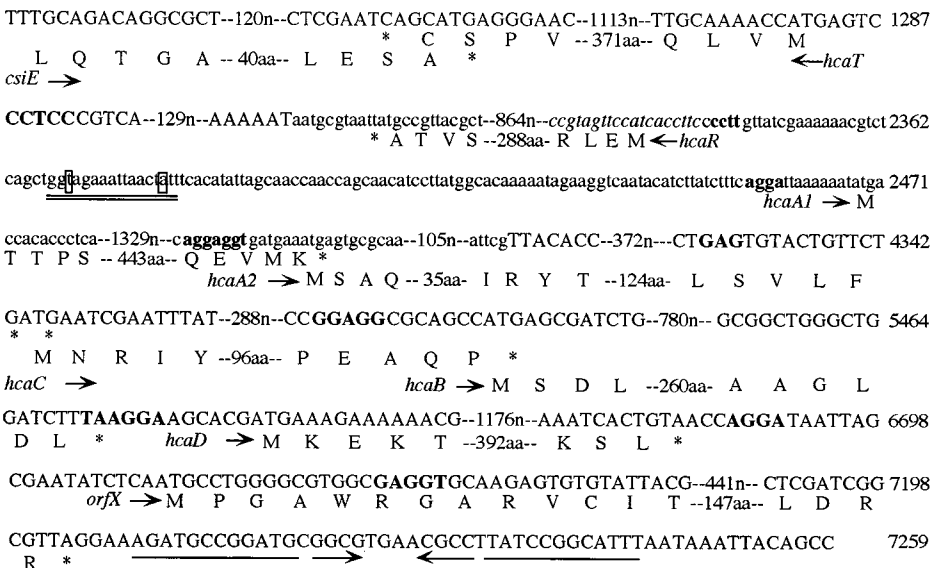

FIG. 3. Nucleotide and derived amino acid sequences of the PP dioxygenolytic pathway. The sequence data in uppercase letters appear in the GenBank/EMBL data bank under accession numbers Y11071 (nucleotides 1 to 1433) and Y11070 (nucleotides 3948 to 7259). The sequence from nucleotide 1434 to 3947 (lowercase letters) was taken from the GenBank/EMBL data bank (accession number AE000340). Only the sequences of the 5′- and 3′-end-coding regions of the *hca* genes and *orfX* are shown. The 3′ end of the *csiE* gene is also shown. Short arrows, direction of gene transcription; asterisks, stop codons. Potential Shine-Dalgarno sequences are boldfaced. Inverted repeats are marked with facing arrows underneath the sequence. The putative binding motif of LTTRs (49) is doubly underlined, and the characteristic T and A residues are boxed. The nucleotide sequence present in oligonucleotide HCAR, used for PCR amplification of *hcaA1*, is italicized.

TABLE 2. PP pathway genes, gene products, and identities with other proteins

Gene	% G+C content	Gene product	Deduced no. of residues (kDa)	% Identity with other gene products (no. of residues)[a]
hcaA1	52.9	Large terminal subunit of phenylpropionate dioxygenase	453 (51.1)	47.8, TcbAa (450); 47.6, IpbA1_BD2 (460); 47.4, BphA1_A1 (460); 47.3, TodC1 (450); 47.0, TecA1 (449); 46.8, BedC1 (450); 44.3, BphA1_P6 (461); 44.2, BphA1_KF (458); 44.2, BphA_B (457); 44.1, BpdC1 (461); 43.9, BphA_LB (459); 42.9, BnzA (448); 42.9, CumA1 (459); 42.9, IpbA1_JR (459); 42.9, BphA1_KKS (458); 27.5, CmtAb (434)
hcaA2	52.2	Small terminal subunit of phenylpropionate dioxygenase	172 (20.5)	37.7, BphE_B (186); 34.4, CumA2 (186); 34.4, IpbA2_JR (186); 34.2, BphE_LB (188); 32.8, BphA2_A1 (187); 32.3, IpbA2_BD2 (187); 32.0, BphA2_KKS (193); 31.2, BedC2 (187); 30.7, TecA2 (187); 30.5, BnzB (187); 30.5, TodC2 (187); 30.5, TcbAb (187); 30.2, BphA2_KF (213); 30.2, CmtAc (180); 29.6, BpdC2 (186); 29.6, BphA2_P6 (186)
hcaC	55.4	Ferredoxin subunit of phenylpropionate dioxygenase	106 (11.3)	48.1, BphA3_A1 (107); 47.2, IpbA3_BD2 (107); 44.5, CumA3 (109); 43.6, IpbA3_JR1 (109); 43.5, TcbAc (107); 43.1, BphA3_KKS (109); 42.6, TecA3 (107); 41.8, BphF_LB (109); 41.3, BedB (107); 40.0, BphA3_P6 (108); 40.0, BphA3_KF (108); 40.0, BpdB (108); 39.8, TodB (107); 39.0, CmtAd (118); 37.4, BnzC (91); 36.7, BphF_B (109)
hcaD	54.0	Ferredoxin reductase subunit of phenylpropionate dioxygenase	400 (43.9)	34.6, CmtAa (402); 31.5, IpbA4_JR (411); 31.5, CumA4 (411); 30.9, BphG_B (406); 30.8, BphG_LB (408); 30.8, BphA4_KF (408); 29.7, BpdA (412); 29.7, BphA4_P6 (412); 29.5, TecA4 (410); 29.5, TcbAd (410); 28.8, TodA (410); 28.7, BphA4_KKS (410); 28.5, IpbA4_BD2 (412); 27.5, BedA (410); 27.3, BphA4_A1 (413); 27.2, BnzD (409)
hcaB	55.9	2,3-Dihydroxy-2,3-dihydrophenylpropionate dehydrogenase	270 (28.5)	48.9, BphB_A1 (263); 46.2, BphB_KKS (276); 45.7, BphB_LB (277); 45.7, BphB_KF (277); 45.1, CumB (276); 44.3, BphB_B (281); 43.7, IpbB_JR (276); 43.1, TodD (275); 43.1, BnzE (275); 42.7, BpdD (280); 42.7, BphB_P6 (280); 41.7, TcbB (275); 22.2, CmtB (259); 17.4, BedD (365)
hcaR	49.8	Activator of *hca* cluster	296 (32.8)	35.6, CatR (289); 35.2, AlsR_E (297); 33.4, AlsR_B (302); 28.9, CatM (303); 28.9, BphR (314); 28.6, TfdR (295); 27.9, TcbR (294); 26.3, ClcR (294); 25.0, TfdT (228); 24.1, PcaQ (311); 19.0, NahR (374)
hcaT	56.3	Potential transporter	379 (41.6)	35.7, HI0308 (388); 23.6, MalA (394); 23.4, LacY (417); 21.6, CscB (415); 21.4, MhpT (403); 20.6, HppK (453); 20.4, HpaX (458); 20.1, MucK (413); 19.5, TfdK (460); 19.3, Pht1 (451); 19.3, MopB (449); 18.9, PcaT (429); 18.5, PcaK (448); 18.0, BenK (466); 17.7, PcaK_A (421)

[a] Sequences included in this analysis, with their accession numbers in parentheses, are as follows. Bed, benzene degradation of *P. putida* ML2 (L04642, U08463); Bnz, benzene degradation of *P. putida* 136-R3 (M17904); Bpd, biphenyl/chlorobiphenyl (PCB) degradation of *Rhodococcus* sp. strain M5 (U27591); Bph_A1, PCB degradation of *Rhodococcus* sp. strain RHA1 (D32142); Bph_B, PCB degradation of *C. testosteroni* B-356 (U47637, U47638); Bph_KF, PCB degradation of *P. pseudoalcaligenes* KF707 (M83673); Bph_KKS, PCB degradation of *Pseudomonas* sp. strain KKS102 (D17319); Bph_LB, PCB degradation of *Pseudomonas* sp. strain LB400 (M86348, X66122); Bph_P6, PCB degradation of *R. globerulus* P6 (X80041, X75633); Cmt, *p*-cumate degradation of *P. putida* F1 (U24215); Cum, cumene degradation of *P. fluorescens* IP01 (D37828); Ipb_BD2, isopropylbenzene degradation of *R. erythropolis* BD2 (U24277); Ipb_JR1, isopropylbenzene degradation of *Pseudomonas* sp. strain JR1 (U53507); Tcb, chlorobenzene degradation of *Pseudomonas* sp. strain P51 (U15298); Tec, tetrachlorobenzene degradation of *Burkholderia* sp. strain PS12 (U78099); Tod, toluene degradation of *P. putida* F1 (J04996); AlsR_B (L04470) and AlsR_E (D90801), regulators of the acetoin synthesis in *Bacillus subtilis* and *E. coli*, respectively; BphR, putative regulatory protein of biphenyl catabolism in *Pseudomonas* sp. strain KKS102 (D38633); CatR (M33817) and CatM (M76991), activators of the catechol degradation pathway of *P. putida* and *Acinetobacter calcoaceticus*, respectively; ClcR, activator of the 3-chlorocatechol degradation pathway of *P. putida* (L06464); NahR, activator of the naphthalene degradation pathway of *P. putida* (J04233); PcaQ, activator of the protocatechuate degradation pathway of *Agrobacterium tumefaciens* (U32867); TcbR, activator of the chlorobenzene degradation pathway of *Pseudomonas* sp. strain P51 (M80212); TfdT (AE16782) and TfdR (M98445), nonfunctional and functional activators of the chlorocatechol degradation pathway of *Ralstonia eutrophus* JMP134; HI0308, hypothetical protein from *Haemophilus influenzae* (U32716); MalA, maltose permease from *B. stearothermophilus* (L13418); LacY, lactose permease from *E. coli* (P02920); CscB, sucrose permease from *E. coli* (X63740); MhpT (D64043, X97543) and HppK (U89712), putative 3HPP transporters from *E. coli* and *R. globerulus* PWD1, respectively; HpaX, 4-hydroxyphenylacetic acid transporter from *E. coli* (Z37980); MucK (U87258) and BenK (AF009224), *cis*,*cis*-muconate and benzoate transporters, respectively, from *A. calcoaceticus* ADP1; TfdK, putative 2,4-dichlorophenoxyacetate transporter from *R. eutrophus* (U16782); Pht1, putative phthalate transporter from *P. putida* (D13229); MopB, 4-methylphthalate transporter from *Burkholderia cepacia* (U29532); PcaT, putative protocatechuate transporter from *P. putida* PRS2000 (U48776); PcaK, 4-hydroxybenzoate transporter from *P. putida* PRS2000 (U10895); PcaK_A, putative 4-hydroxybenzoate transporter from *A. calcoaceticus* ADP1 (L05770).

more, immediately downstream of *orfX* there is an inverted-repeat sequence (Fig. 3) predicted to form a hairpin loop (ΔG, −23.9 kcal/mol) that could act as a transcriptional terminator of a potential operon. Genes *hcaR* and *hcaT* are located upstream of *hcaA1A2CBD-orfX*, but they are transcribed in the opposite direction (Fig. 1A and 3). Although the intergenic spacing between genes *hcaR* and *hcaT* was 159 bp, we could not detect in this DNA fragment typical transcriptional terminator and promoter sequences. The G+C content of the *hca*-coding regions averaged 53.8%, a value close to the mean G+C content of the *E. coli* genomic DNA (51.5%) (37).

The *hca* cluster maps immediately downstream of gene *csiE*, which encodes the stationary-phase inducible protein CsiE, at min 57.5 of the *E. coli* chromosome (8) and therefore far from min 8, where the *mhp* cluster responsible for DHPP degradation is located (20). Interestingly, the third aromatic catabolic pathway characterized so far in *E. coli* at the molecular level, i.e., the *hpa* cluster for 4-hydroxyphenylacetate degradation, was shown to map at min 98 (43). Thus, while in some *Pseudomonas* and *Acinetobacter* species a supraoperonic clustering of the aromatic catabolic genes has been observed in a limited region of the chromosome (15, 16, 60), in *E. coli* the aromatic catabolic clusters are dispersed throughout the genome.

The deduced amino acid sequences of the *hca* gene products were compared with entries in the databases, and the ones showing the highest similarities were then retrieved and analyzed (Table 2).

***hca* catabolic genes.** Sequence comparison analyses of the *hcaA1*, *hcaA2*, *hcaC*, and *hcaD* gene products revealed significant similarities with the corresponding four protein subunits of the three-component class IIB ring-activating dioxygenases (12), mainly with the analogous *ipb*, *cum*, and *bph* gene products (Table 2). The *hcaA1A2CD* genes are suggested, therefore, to encode the HcaA1A2CD initial dioxygenase of the PP catabolic pathway (Fig. 1).

The *hcaA1* gene encodes a protein of 51,109 Da (453 amino acids) that shows significant similarity with the large (α) subunit of the terminal oxygenase component of multicomponent dioxygenases (Table 2). It is worth noting that although residues 85-CRHR*A*MRVSYADCGNTRAFTC*P*YH-108 in the HcaA1 protein match the binding site of a [2Fe-2S] Rieske-type iron-sulfur cluster (12) (the putative iron-sulfur ligands are underlined), the highly conserved G and S/T residues (12, 48) are replaced in HcaA1 by A and P (italicized), respectively. Residues 205-EQFASDQYHALFSH-218 in the HcaA1 primary structure match perfectly the mononuclear Fe(II) ligand at the site of oxygen activation (26). It should be noted that the C terminus of HcaA1 differs from that deduced from the reported sequence of *orfA* (accession no. Z37966), since *orfA* lacks a nucleotide leading to a change in the reading frame.

The *hcaA2* and *hcaC* genes encode proteins of 20,579 Da (172 amino acids) and 11,328 Da (106 amino acids) whose deduced amino acid sequences show significant identity to those of the small (β) subunit of the terminal oxygenase and the ferredoxin component of multicomponent dioxygenases, respectively (Table 2) (3). Residues 42-CSHGNASMSEGYLE DDATVEC*P*LH-65 in HcaC are likely to be involved in the coordination of a Rieske-type [2Fe-2S] cluster (the putative iron-sulfur ligands are underlined; the unusual proline residue which is also present in the Rieske-type cluster of HcaA1 is italicized).

The next gene of the *hca* cluster, *hcaB*, encodes a protein of 28,498 Da (270 amino acids) that shows significant identity with *cis*-dihydrodiol dehydrogenases that participate in pathways involving class IIB dioxygenases (Table 2) and convert the stable *cis*-dihydrodiols formed by the initial dioxygenases into the corresponding dihydroxy derivatives with regeneration of NADH (12). Therefore, HcaB is postulated to be the 3-phenylpropionate-dihydrodiol dehydrogenase (Fig. 1). The length of HcaB falls within the average of 270 amino acids for members of the short-chain alcohol dehydrogenase (type II) superfamily, which includes all dihydrodiol dehydrogenases in Table 2 with the exception of BedD (21). Residues 13-GGGSGLG-19 and 156-YTASKHAATGL-166 in HcaB fit the NAD^+-binding domain (21) and the consensus pattern for short-chain alcohol dehydrogenases (18, 53), respectively. The highly conserved aspartate 92 of HcaB has been also implicated in the catalytic activity of analogous enzymes (38).

The *hcaD* gene encodes a 43,978-Da protein (400 amino acids) that is homologous to the ferredoxin reductase subunit of other dioxygenases (Table 2). Multiple sequence alignments revealed the three conserved motifs in the same relative locations found in other reductase components of class IIB dioxygenases (12). Thus, residues 10-GGGQA*A*AMAAASLRQ QG-26 and 151-GAGTIGLELAASATQR*R*CKVTVIE-174 of HcaD match the consensus sequence postulated to be involved in binding of the ADP moiety of flavin adenine dinucleotide and NAD^+ (amino acids in italics indicate a replacement of a consensus residue) (12), and the sequence 265-TCDPAIFAG GD-275 fits with the consensus motif that has been postulated to bind the O-3 group of the ribityl chain of the flavin moiety of flavin adenine dinucleotide (12).

The gene organization within the *hca* catabolic cluster is similar to that of the analogous *ipb*, *cum*, *bph*, *tod*, *bed*, *tcb*, *bnz*, and *tec* clusters encoding class IIB dioxygenases and consisting of the large subunit of the terminal oxygenase, small subunit, ferredoxin, and reductase (57). However, the *hcaD* gene, although physically linked to the other three genes encoding the HcaA1A2CD dioxygenase, is separated from them by the *hcaB* gene (Fig. 1A and 3). An unusual location of the gene encoding the reductase component has been also observed for *bphA4* in *Pseudomonas* sp. strain KKS102 (28), *bphG* in *Comamonas testosteroni* B-356 (54), and *cmtAa* in *Pseudomonas putida* F1 (15). It has been reported that reductases have diverged more than the other components of the dioxygenases (57), and indeed the degrees of identity between HcaD and its orthologs in other clusters are lower than those observed for the other three HcaA1A2C subunits (Table 2). HcaD showed the highest level of identity to the CmtAa reductase component of the *p*-cumate dioxygenase (Table 2); HcaA1A2CD and CmtAabcd are the only class IIB dioxygenases described so far that attack carboxylated aryls.

At the 3′ end of the *hcaD* gene is located *orfX* (Fig. 1A and 3), which has two potential translational start codons at positions 6710 (ATG) and 6737 (GTG), although only the latter shows a putative Shine-Dalgarno sequence (GAGGT) at a reasonable distance (Fig. 3), and codes for a 155-amino-acid product of unknown function. It is worth noting that an additional ORF of unknown function has also been found in other gene clusters encoding biphenyl (19, 22, 55), isopropylbenzene (42), and benzoate (accession no. M76990) dioxygenases.

Regulation of the *hca* cluster. The 5′ end of the *hca* region contains two genes, *hcaR* and *hcaT*, that are oriented in the direction opposite to that of the other *hca* genes (Fig. 1A and 3). The *hcaR* gene encodes a protein of 32,838 Da (296 amino acids) that shows a size and an amino acid sequence similar to those of LysR-type transcriptional regulators (LTTRs) (49) (Table 2). The majority of the genes encoding LTTRs are transcribed divergently from the genes that they regulate (49). In this sense, *hcaR* is transcribed divergently from the catabolic genes *hcaA1A2CBD* (Fig. 1A and 3). LTTRs show a high degree of similarity in the N-terminal domain, where the helix-turn-helix DNA-binding region is located (49). Thus, HcaR possesses a sequence (18-FTRAAEKLHTSQPSLSSQIRDLE NCV-43) that matches the LTTR helix-turn-helix motif (Prosite signature PS00044) (5). The C-terminal domain of LTTRs seems to be involved in multimerization, and its consensus motif $(V/L)X_2GXG(V/I)XV(L/V)P$ (49) fits with the sequence (232-VGMGLGVTLIP-242) found in HcaR. Within the LysR family, HcaR shows the highest degrees of identity with the AlsR regulator of acetoin synthesis and with a select group of regulators from other biodegradative operons (Table 2). This group constitutes the Cat subfamily and includes the CatR, CatM, TfdR, TcbR, ClcR, and TfdT regulators that activate the genes encoding muconate- or chloromuconate-lactonizing enzymes and/or genes encoding oxygenases that act on catechol or chlorinated aromatic compounds from *Acinetobacter*, *Pseudomonas*, and *Ralstonia* species (33). Additionally, HcaR shows significant identity with the putative *bphR* gene product, which supposedly would be involved in regulation of biphenyl catabolism (27). Therefore, all of these observations strongly suggest that HcaR is the transcriptional regulator of the *hca* cluster of *E. coli.*

To study the regulation of the *hca* cluster, we performed complementation studies of strains lacking Hca activity. While resting-cell assays of *E. coli* ED1061(pCKES) bacteria grown in LB medium in the absence of PP did not show significant removal of this aromatic compound, resting cells of these bacteria grown in the presence of 1 mM PP revealed that removal

of PP was concomitant with the appearance of a product which cochromatographed with authentic DHPP in HPLC (Table 1). These data, therefore, confirmed that the *hca*-encoded pathway was inducible. To demonstrate that HcaR was required for *hca* expression and to determine its mechanism of action, the *hcaR* and the *hcaA1A2CBD-orfX* genes were independently expressed from the compatible plasmids pHCAR and pCKET (Fig. 2), respectively. Since *S. typhimurium* LT-2 is unable to attack PP (Table 1), we used this strain as host for studying the regulation of the *hca* genes. Resting cells of *S. typhimurium* LT-2(pCKET) bacteria grown in the presence of 1 mM PP did not reveal the formation of DHPP, and PP remained unaltered. However, when the gene *hcaR* was provided in *trans* from plasmid pHCAR, the resulting strain *S. typhimurium* LT-2(pCKET, pHCAR) showed a significant conversion of PP into DHPP in a resting-cell assay when the bacteria were grown in the presence of 1 mM PP (Table 1). Furthermore, while *S. typhimurium* LT-2 harboring simultaneously the compatible plasmids pPADR2 (containing the *mhp* genes for the catabolism of 3HPP) (20) and pCKES grew efficiently (doubling time, 5 h) on minimal medium containing 5 mM PP as the sole carbon and energy source, *S. typhimurium* LT-2(pPADR2, pCKET) cells did not grow in this aromatic compound. Thus, all these data indicated that HcaR fostered inducible expression of the *hca* catabolic genes, behaving as a transcriptional activator. It should be mentioned that although the formation of PP-dihydrodiol (compound II in Fig. 1B) is assumed by previous work and by analogy to other systems, the HcaB activity per se has not been demonstrated.

A 135-bp intergenic region is located between the potential translational start sites of the divergently transcribed *hcaR* and *hcaA1* genes (Fig. 3), suggesting that the *hcaR* promoter is located near or overlaps the regulated promoter of the putative *hca* catabolic operon. As it has been noted with other LysR-type regulatory targets (40), the A+T content (66%) of the *hcaR-hcaA1* intergenic region is higher than that of the *hca* genes (47%). LTTRs characteristically bind to a consensus T-N_{11}-A DNA binding motif, with the T and A being part of a short inverted repeat, positionally conserved upstream of the regulated promoter (49). In this sense, within the *hca* intergenic region and located 85 nucleotides upstream of the putative *hcaA1* translation start site, we have found the sequence TA<u>G</u>-N_7-<u>C</u>TA that matches the binding motif of LysR-type regulators (Fig. 3). Underlined are the guanine and cytosine of the dyad, which have been shown to be involved in the binding of some LTTRs (45). Demonstration of these assumptions and elucidation of the mechanisms of *hca* repression by glucose (30) and *hcaR* expression will require further research.

The *hcaT* gene encodes a protein of 41,619 Da (379 amino acids) that shows significant identity with several members of the major facilitator superfamily (MFS) of transport proteins (34) (Table 2). Interestingly, HcaT was smaller than other MFS members (about 400 amino acids), and some common amino acid sequences that characterize this superfamily (20) were not found in the primary structure of the *hcaT* gene product. However, analysis of the predicted secondary structure of HcaT revealed the characteristic 12 membrane-spanning α helices which are found in other MFS permeases and are believed to form a channel for transport through the membrane (34). Moreover, the hydrophilicity profile of HcaT showed that the protein could be divided by a central hydrophilic region into two halves, each containing six transmembrane domains (data not shown). Therefore, the putative HcaT protein might be involved in the uptake of PP in *E. coli* and could be another member of the rapidly expanding family of transporters for the catabolism of aromatic compounds (6).

It has been proposed that permeases for aromatic compounds are indirectly involved in the regulation of the catabolic pathways by bringing these substrates (inducers) inside the cell, leading to the induction of their respective regulatory proteins (43). This idea would agree with the close association between the *hcaR* and *hcaT* genes. Similar gene arrangements have been found for the permeases HpaX (4-hydroxyphenylacetate), PcaK (4-hydroxybenzoate), and HppK (3HPP) and the corresponding transcriptional regulators HpaA, PcaR, and HppR, respectively (6, 25, 39, 43, 44). Interestingly, the catabolism of 3HPP in *E. coli* also involves a gene encoding a putative 3HPP permease (MhpT) that shows similarity with HcaT (Table 2) and a gene encoding a transcriptional activator (MhpR) which belongs to a family of regulators different from that of HcaR; these two genes are not transcriptionally coupled in the *mhp* operon (20).

Catabolism of CI and 3HCI. It is known that *E. coli* K-12 is also able to grow with 3HCI (compound IX in Fig. 1B) as the sole carbon and energy source (10). Since growth with 3HCI induces the synthesis of enzymes MhpA and MhpB, responsible for initial attack upon 3HPP, it was suggested that the same enzymes are used for catabolizing these two compounds (10). To demonstrate this hypothesis, plasmid pPADR2 was introduced into *S. typhimurium* LT-2, a strain unable to grow on 3HPP (20) and 3HCI. The recombinant strain *S. typhimurium* LT-2(pPADR2) acquired the ability to grow not only on 3HPP (20) but also on minimal medium containing 3HCI as the sole carbon and energy source. Hence, these data provided experimental demonstration that the *mhp* genes are also responsible for the mineralization of 3HCI.

Although *E. coli* cannot grow on CI (compound VI in Fig. 1B) as the sole carbon source, whole cells grown on PP rapidly oxidized CI, suggesting that Hca enzymes are also able to attack this aromatic compound (10). To confirm this assumption, *E. coli* ED1061(pCKES) bacteria were grown in LB medium containing 1 mM PP for the induction of the *hca* genes, and then they were used in a resting-cell assay with 0.5 mM CI as substrate. After 30 min of incubation at 30°C, CI was enzymatically converted to a product which cochromatographed in HPLC with authentic DHCI (compound VIII in Fig. 1B). NMR spectroscopy of the purified product confirmed it as DHCI (data not shown). Interestingly, the mutant strain *E. coli* ED1061 (control cells) grown in PP-containing LB medium did not attack CI in a resting-cell assay. The conversion of CI to DHCI was also observed with *S. typhimurium* LT-2(pCKES) cells. Hence, we concluded that the HcaA1A2CD dioxygenase and HcaB dihydrodiol dehydrogenase were responsible for CI oxidation and that DHCI was the final product of these reactions.

To analyze whether CI is an inducer of the *hca* genes, we performed a resting-cell assay using PP as substrate and *E. coli* ED1061(pCKES) bacteria grown in CI-containing LB medium. Since conversion of PP into DHPP was observed, we concluded that CI can also induce the *hca* genes. Moreover, we were able to demonstrate that the *hca* and *mhp* genes are responsible for CI mineralization by showing that *S. typhimurium* LT-2(pCKES, pPADR2) cells grew on minimal medium containing 1 mM CI as the sole carbon and energy source. It is worth noting that while in some soil *Pseudomonas* species and in *Lactobacillus pastorianus* the catabolism of CI could be accomplished by an initial reduction of the double bond of the side chain with the formation of PP (2, 7, 58), the catabolism of CI by the Hca enzymes produces DHCI, which, via the *mhp*-encoded pathway (20), will be finally mineralized to pyruvate, acetyl coenzyme A, and fumarate (Fig. 1B).

Since CI can induce the *hca* genes and is converted to DHCI

by *E. coli* cells, it is difficult to explain the lack of growth of this bacterium on this aromatic compound. A possible explanation for this behavior could be that the DHCI generated in the reactions catalyzed by the Hca enzymes or other intermediates further down the *mhp*-encoded pathway accumulate to a toxic level that prevents the normal metabolic flux of the cell. When *E. coli* MG1655(pCKES, pPADR2) cells were grown on minimal medium containing 20 mM glycerol plus 1 mM CI and the culture supernatants were analyzed by HPLC, we observed that CI depletion was not accompanied by the accumulation of DHCI (data not shown). However, these supernatants acquired a yellow coloration that disappeared after acidification with HCl, thus suggesting accumulation of the ring fission product of DHCI. These data are in agreement with previous observations showing that whole cells of *E. coli* grown in PP or 3HPP were able to oxidize DHCI with the transient formation of a yellow compound having the typical characteristics of a ring fission product but differing from those of compound V derived from ring cleavage of DHPP (Fig. 1B) (10). Furthermore, it has been shown recently that although DHCI is a good substrate for the MhpB dioxygenase (50), the ring fission product of this compound is hydrolyzed by the MhpC enzyme 36-fold less efficiently than the ring fission product of DHPP (32). All these data taken together may suggest that CI cannot support the growth of *E. coli* because its oxidation to DHCI generates toxic levels of the corresponding ring fission product. Although the toxicity of the ring fission products has been reported (10), the possibility that CI can cause on *E. coli* toxic effects that are not directly related to its catabolism cannot be ruled out. Why this toxicity is not observed in *S. typhimurium* is still an open question.

In conclusion, the results presented here constitute the first genetic characterization of a dioxygenolytic pathway for the initial catabolism of PP and CI. It has been suggested that pathways for the catabolism of aromatic compounds widely available in nature, such as PP, HPP, phenylacetate, and hydroxyphenylacetate, are among the most ubiquitous aromatic-compound catabolic systems, and they are closer to central metabolism than those involved in the degradation of xenobiotic compounds (6). These pathways, which occupy central positions within secondary metabolism, may have been one of the most common sources for the initial recruitment of genes for many of the routes involved in the degradation of anthropogenic compounds which are more peripheral to the natural carbon cycle (6). Since the whole structure and organization of the *hca* cluster resemble those of clusters responsible for initial dioxygenation of the highly recalcitrant polychlorinated biphenyls (PCBs) (*bph*) and chlorinated benzenes (*tcb* and *tec*), it seems likely that these peripheral pathways have evolved from a central one, such as *hca*, through mutation, recombination, and gene transfer events. Interestingly, PCB degraders have been found to be associated with plant lignin degraders (29), and the breakdown of lignin is one of the major natural sources of phenylpropanoid compounds (2). Moreover, it has been postulated that in the PCB degrader *Rhodococcus* sp. strain RHA1, the *meta* cleavage pathway genes could have evolved from the same ancestor as hydroxyphenylacetate *meta* cleavage pathway genes (*hpa* and *hpc*) of *E. coli* (35). Thus, it is tempting to speculate that the catabolic pathways for the mineralization of PCBs may derive from aromatic-compound central pathways through the assembling of *hca*-like clusters with *mhp*- or *hpa*-like clusters. On the other hand, the characterization of the *hca*-encoded dioxygenolytic pathway of *E. coli* confirms that this bacterium is endowed with genetic systems (*hca*, *mhp*, and *hpa*) highly similar to those in environmentally relevant bacteria such as those of the genus *Pseudomonas*, and this fact should be taken into consideration when aromatic catabolic clusters are cloned and expressed in this enterobacterium. There are several reports on the cloning and expression of aromatic dioxygenases in *E. coli* claiming that equivalent enzymes from the host could explain partial activities observed when some of the subunits of the cloned dioxygenase were missing in the recombinant bacteria (17, 52). The expression of the *hca* genes in *E. coli* might explain these reported observations.

ACKNOWLEDGMENTS

We thank T. Bugg for providing DHPP and DHCI, M. Vicente and M. Aldea for the plasmid pMAK700 and strain MC1061, M. K. B. Berlyn for strain MG1655, and A. Díaz and G. Porras for assistance with the sequencing. The help of J. F. Espinosa with the NMR studies is gratefully acknowledged.

This work was supported by grants from CICYT (AMB94-1038-C02-02 and AMB97-063-C02-02). A. Ferrández was the recipient of a predoctoral fellowship from the Plan Nacional de Formación de Personal Investigador-MEC. E. Díaz was the recipient of a Contrato Temporal de Investigadores from the CSIC.

REFERENCES

1. **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.
2. **Andreoni, V., and G. Bestetti.** 1986. Comparative analysis of different *Pseudomonas* strains that degrade cinnamic acid. Appl. Environ. Microbiol. **52:**930–934.
3. **Asturias, J. A., E. Díaz, and K. N. Timmis.** 1995. The evolutionary relationships of biphenyl dioxygenase from Gram-positive *Rhodococcus globerulus* P6 to multicomponent dioxygenases from Gram-negative bacteria. Gene **156:**11–18.
4. **Bachmann, B. J.** 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 1190–1219. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella thyphimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
5. **Bairoch, A., P. Bucher, and K. Hofmann.** 1995. The PROSITE database, its status in 1995. Nucleic Acids Res. **24:**189–196.
6. **Barnes, M. R., W. A. Duetz, and P. A. Williams.** 1997. A 3-(3-hydroxyphenyl)propionic acid catabolic pathway in *Rhodococcus globerulus* PWD1: cloning and characterization of the *hpp* operon. J. Bacteriol. **179:**6145–6153.
7. **Blakley, E. R., and F. J. Simpson.** 1964. The microbial metabolism of cinnamic acid. Can. J. Microbiol. **10:**175–185.
8. **Blattner, F. R., G. Plunkett III, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453–1462.
9. **Bugg, T. D. H.** 1993. Overproduction, purification and properties of 2,3-dihydroxyphenylpropionate 1,2-dioxygenase from *Escherichia coli*. Biochim. Biophys. Acta **1202:**258–264.
10. **Burlingame, R., and P. J. Chapman.** 1983. Catabolism of phenylpropionic acid and its 3-hydroxy derivative by *Escherichia coli*. J. Bacteriol. **155:**113–121.
11. **Burlingame, R. P., L. Wyman, and P. J. Chapman.** 1986. Isolation and characterization of *Escherichia coli* mutants defective for phenylpropionate degradation. J. Bacteriol. **168:**55–64.
12. **Butler, C. S., and J. R. Mason.** 1997. Structure-function analysis of the bacterial aromatic ring-hydroxylating dioxygenases. Adv. Microb. Physiol. **38:**47–84.
13. **Cavin, J.-F., L. Barthelmebs, and C. Diviès.** 1997. Molecular characterization of an inducible *p*-coumaric acid decarboxylase from *Lactobacillus plantarum*: gene cloning, transcriptional analysis, overexpression in *Escherichia coli*, purification, and characterization. Appl. Environ. Microbiol. **63:**1939–1944.
14. **Dagley, S., P. J. Chapman, and D. T. Gibson.** 1965. The metabolism of β-phenylpropionic acid by an *Achromobacter*. Biochem. J. **97:**643–650.
15. **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.
16. **Eaton, R. W.** 1997. *p*-Cymene catabolic pathway in *Pseudomonas putida* F1: cloning and characterization of DNA encoding conversion of *p*-cymene to *p*-cumate. J. Bacteriol. **179:**3171–3180.
17. **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.

18. **Ensor, C. M., and H. H. Tai.** 1991. Site-directed mutagenesis of the conserved tyrosine 151 of human placental NAD^+-dependent 15-hydroxyprostaglandin dehydrogenase yields a catalytically inactive enzyme. Biochem. Biophys. Res. Commun. **176:**840–845.
19. **Erickson, B. D., and F. J. Mondello.** 1992. Nucleotide sequencing and transcriptional mapping of the genes encoding biphenyl dioxygenase, a multicomponent polychlorinated-biphenyl-degrading enzyme in *Pseudomonas* strain LB400. J. Bacteriol. **174:**2903–2912.
20. **Ferrández, A., J. L. García, and E. Díaz.** 1997. Genetic characterization and expression in heterologous hosts of the 3(3-hydroxyphenyl)propionate catabolic pathway of *Escherichia coli* K-12. J. Bacteriol. **179:**2573–2581.
21. **Fong, K. P. Y., C. B. H. Goh, and H.-M. Tan.** 1996. Characterization and expression of the plasmid-borne *bedD* gene from *Pseudomonas putida* ML2, which codes for a NAD^+-dependent *cis*-benzene dihydrodiol dehydrogenase. J. Bacteriol. **178:**5592–5601.
22. **Fukuda, M., Y. Yasukouchi, Y. Kikuchi, Y. Nagata, K. Kimbara, H. Horiuchi, M. Takagi, and K. Yano.** 1994. Identification of the *bphA* and *bphB* genes of *Pseudomonas* sp. strain KKS102 involved in degradation of biphenyl and polychlorinated biphenyls. Biochem. Biophys. Res. Commun. **202:**850–856.
23. **Gold, L.** 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. Annu. Rev. Biochem. **57:**199–233.
24. **Hamilton, C. M., M. Aldea, B. K. Washburn, P. Babitzke, and S. R. Kushner.** 1989. New method for generating deletions and gene replacements in *Escherichia coli*. J. Bacteriol. **171:**4617–4622.
25. **Harwood, C. S., N. N. Nichols, M.-K. Kim, J. L. Ditty, and R. E. Parales.** 1994. Identification of the *pcaRKF* gene cluster from *Pseudomonas putida*: involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. J. Bacteriol. **176:**6479–6488.
26. **Jiang, H., R. E. Parales, N. A. Lynch, and D. T. Gibson.** 1996. Site-directed mutagenesis of conserved amino acids in the alpha subunit of toluene dioxygenase: potential mononuclear non-heme iron coordination sites. J. Bacteriol. **178:**3133–3139.
27. **Kikuchi, Y.** GenBank accession no. D38633.
28. **Kikuchi, Y., Y. Nagata, M. Hinata, K. Kimbara, M. Fukuda, K. Yano, and M. Takagi.** 1994. Identification of the *bphA4* gene encoding ferredoxin reductase involved in biphenyl and polychlorinated biphenyl degradation in *Pseudomonas* sp. strain KKS102. J. Bacteriol. **176:**1689–1694.
29. **Kimura, N., A. Nishi, M. Goto, and K. Furukawa.** 1997. Functional analyses of a variety of chimeric dioxygenases constructed from two biphenyl dioxygenases that are similar structurally but different functionally. J. Bacteriol. **179:**3936–3943.
30. **Kovàrovà, K., A. Käch, A. J. B. Zehnder, and T. Egli.** 1997. Cultivation of *Escherichia coli* with mixtures of 3-phenylpropionic acid and glucose: steady-state growth kinetics. Appl. Environ. Microbiol. **63:**2619–2624.
31. **Kröger, M., and R. Wahl.** 1996. Compilation of DNA sequences of *Escherichia coli* K12 (ECD and ECDC; update 1995). Nucleic Acids Res. **24:**29–31.
32. **Lam, W. W. Y., and T. D. H. Bugg.** 1997. Purification, characterization, and stereochemical analysis of a C-C hydrolase: 2-hydroxy-6-keto-nona-2,4-diene-1,9-dioic acid 5,6-hydrolase. Biochemistry **36:**12242–12251.
33. **Leveau, J. H. J., and J. R. van der Meer.** 1996. The *tfdR* gene product can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. **178:**6824–6832.
34. **Marger, M. D., and M. H. Saier, Jr.** 1993. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem. Sci. **18:**13–20.
35. **Masai, E., K. Sugiyama, N. Iwashita, S. Shimizu, J. E. Hauschild, T. Hatta, K. Kimbara, K. Yano, and M. Fukuda.** 1997. The *bphDEF meta*-cleavage pathway genes involved in biphenyl/polychlorinated biphenyl degradation are located on a linear plasmid and separated from the initial *bphACB* genes in *Rhodococcus* sp. strain RHA 1. Gene **187:**141–149.
36. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
37. **Nakamura, Y., T. Gojobori, and T. Ikemura.** 1997. Codon usage tabulated from the international DNA sequence databases. Nucleic Acids Res. **25:**244–245.
38. **Nakatsu, C. H., M. Providenti, and R. C. Wyndham.** 1997. The *cis*-diol dehydrogenase *cbaC* gene of Tn*5271* is required for growth on 3-chlorobenzoate but not 3,4-dichlorobenzoate. Gene **196:**209–218.
39. **Nichols, N. N., and C. S. Harwood.** 1997. PcaK, a high-affinity permease for the aromatic compounds 4-hydroxybenzoate and protocatechuate from *Pseudomonas putida*. J. Bacteriol. **179:**5056–5061.
40. **Parke, D.** 1996. Characterization of PcaQ, a LysR-type transcriptional activator required for catabolism of phenolic compounds, from *Agrobacterium tumefaciens*. J. Bacteriol. **178:**266–272.
41. **Pérez-Martin, J., and V. de Lorenzo.** 1996. VTR expression cassettes for engineering conditional phenotypes in *Pseudomonas*: activity of the *Pu* promoter of the TOL plasmid under limiting concentrations of the XylR activator protein. Gene **172:**81–86.
42. **Pflugmacher, U., B. Averhoff, and G. Gottschalk.** 1996. Cloning, sequencing, and expression of isopropylbenzene degradation genes from *Pseudomonas* sp. strain JR1: identification of isopropylbenzene dioxygenase that mediates trichloroethene oxidation. Appl. Environ. Microbiol. **62:**3967–3977.
43. **Prieto, M. A., E. Díaz, and J. L. García.** 1996. Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli* W: engineering a mobile aromatic degradative cluster. J. Bacteriol. **178:**111–120.
44. **Prieto, M. A., and J. L. García.** 1997. Identification of the 4-hydroxyphenylacetate transport gene of *Escherichia coli* W: construction of a highly sensitive cellular biosensor. FEBS Lett. **414:**293–297.
45. **Romero-Arroyo, C. E., M. A. Schell, G. L. Gaines III, and E. L. Neidle.** 1995. *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. J. Bacteriol. **177:**5891–5898.
46. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
47. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
48. **Sato, S.-I., J.-W. Nam, K. Kasuga, H. Nojiri, H. Yamane, and T. Omori.** 1997. Identification and characterization of genes encoding carbazole 1,9a-dioxygenase in *Pseudomonas* sp. strain CA10. J. Bacteriol. **179:**4850–4858.
49. **Schell, M. A.** 1993. Molecular biology of the LysR family of transcriptional regulators. Annu. Rev. Microbiol. **47:**597–626.
50. **Spence, E. L., M. Kawamukai, J. Sanvoisin, H. Braven, and T. D. H. Bugg.** 1996. Catechol dioxygenases from *Escherichia coli* (MhpB) and *Alcaligenes eutrophus* (MpcI): sequence analysis and biochemical properties of a third family of extradiol dioxygenases. J. Bacteriol. **178:**5249–5256.
51. **Strickland, S., and V. Massey.** 1973. The purification and properties of the flavoprotein melilotate hydroxylase. J. Biol. Chem. **248:**2944–2952.
52. **Suen, W.-C., B. E. Haigler, and J. C. Spain.** 1996. 2,4-Dinitrotoluene dioxygenase from *Burkholderia* sp. strain DNT: similarity to naphthalene dioxygenase. J. Bacteriol. **178:**4926–4934.
53. **Sylvestre, M., Y. Hurtubise, D. Barriault, J. Bergeron, and D. Ahmad.** 1996. Characterization of active recombinant 2,3-dihydro-2,3-dihydroxybiphenyl dehydrogenase from *Comamonas testosteroni* B-356 and sequence of the encoding gene (*bphB*). Appl. Environ. Microbiol. **62:**2710–2715.
54. **Sylvestre, M., M. Sirois, Y. Hurtubise, J. Bergeron, D. Ahmad, F. Shareck, D. Barriault, I. Guillemette, and J. M. Juteau.** 1996. Sequencing of *Comamonas testosteroni* strain B-356-biphenyl/chlorobiphenyl dioxygenase genes: evolutionary relationships among Gram-negative bacterial biphenyl dioxygenases. Gene **174:**195–202.
55. **Taira, K., J. Hirose, S. Hayashida, and K. Furukawa.** 1992. Analysis of *bph* operon from the polychlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. J. Biol. Chem. **267:**4844–4853.
56. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22:**4673–4680.
57. **Werlen, C., H.-P. E. Kohler, and J. R. van der Meer.** 1996. The broad substrate chlorobenzene dioxygenase and *cis*-chlorobenzene dihydrodiol dehydrogenase of *Pseudomonas* sp. strain P51 are linked evolutionarily to the enzymes for benzene and toluene degradation. J. Biol. Chem. **271:**4009–4016.
58. **Whiting, G. C., and J. G. Carr.** 1959. Metabolism of cinnamic acid and hydroxycinnamic acids by *Lactobacillus pastorianus* var. *quinicus*. Nature (London) **184:**1427–1428.
59. **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.
60. **Williams, P. A., and L. E. Shaw.** 1997. *mucK*, a gene in *Acinetobacter calcoaceticus* ADP1 (BD413), encodes the ability to grow on exogenous *cis*,*cis*-muconate as the sole carbon source. J. Bacteriol. **179:**5935–5942.