JOURNAL OF BACTERIOLOGY, June 2005, p. 4050–4063 Vol. 187, No. 12
0021-9193/05/$08.00+0 doi:10.1128/JB.187.12.4050–4063.2005

Catabolism of Benzoate and Phthalate in *Rhodococcus* sp. Strain RHA1: Redundancies and Convergence

Marianna A. Patrauchan, Christine Florizone, Manisha Dosanjh, William W. Mohn, Julian Davies, and Lindsay D. Eltis*

Department of Microbiology and Immunology, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

Received 24 September 2004/Accepted 15 March 2005

Genomic and proteomic approaches were used to investigate phthalate and benzoate catabolism in *Rhodococcus* sp. strain RHA1, a polychlorinated biphenyl-degrading actinomycete. Sequence analyses identified genes involved in the catabolism of benzoate (*ben*) and phthalate (*pad*), the uptake of phthalate (*pat*), and two branches of the β-ketoadipate pathway (*catRABC* and *pcaJIHGBLFR*). The regulatory and structural *ben* genes are separated by genes encoding a cytochrome P450. The *pad* and *pat* genes are contained on a catabolic island that is duplicated on plasmids pRHL1 and pRHL2 and includes predicted terephthalate catabolic genes (*tpa*). Proteomic analyses demonstrated that the β-ketoadipate pathway is functionally convergent. Specifically, the *pad* and *pat* gene products were only detected in phthalate-grown cells. Similarly, the *ben* and *cat* gene products were only detected in benzoate-grown cells. However, *pca*-encoded enzymes were present under both growth conditions. Activity assays for key enzymes confirmed these results. Disruption of *pcaL*, which encodes a fusion enzyme, abolished growth on phthalate. In contrast, after a lag phase, growth of the mutant on benzoate was similar to that of the wild type. Proteomic analyses revealed 20 proteins in the mutant that were not detected in wild-type cells during growth on benzoate, including a CatD homolog that apparently compensated for loss of PcaL. Analysis of completed bacterial genomes indicates that the convergent β-ketoadipate pathway and some aspects of its genetic organization are characteristic of rhodococci and related actinomycetes. In contrast, the high redundancy of catabolic pathways and enzymes appears to be unique to RHA1 and may increase its potential to adapt to new carbon sources.

The genus *Rhodococcus* comprises aerobic, gram-positive, nonmotile soil bacteria that occur in a wide variety of environmental niches. Phylogenetically, the genus belongs to the suborder *Corynebacterineae*, a group of GC-rich, mycolic acid-producing bacteria within the order *Actinomycetales* that includes *Gordonia*, *Nocardia*, and *Mycobacterium* (34). Rhodococci assimilate an unusually broad range of organic compounds, particularly hydrophobic xenobiotics, thus playing a key role in the global carbon cycle (6). Their assimilatory abilities have been attributed to their diversity of enzymatic activities as well as their mycolic acid surfactants, which have been proposed to facilitate the uptake of hydrophobic compounds (83). The broad metabolic diversity of rhodococci makes them of great interest to the pharmaceutical, environmental, chemical, and energy industries (6, 81).

Rhodococcus sp. strain RHA1 was isolated from lindane-contaminated soil (69) for its exceptional ability to aerobically degrade polychlorinated biphenyls (PCBs), a class of toxic and persistent pollutants. As in other aerobic PCB-degrading bacteria, these pollutants are cometabolized by the *bph* pathway, which is responsible for the aerobic degradation of biphenyl. The *bph* pathway consists of four enzymatic activities which act sequentially to transform biphenyl to benzoate and 2-hydroxypenta-2,4-dienoate as recently reviewed by Furukawa (30). For each of these four steps, RHA1 appears to possess multiple isozymes, including at least three *bph*-type ring-hydroxylating dioxygenases (51) and at least seven different *bph*-type ring cleavage enzymes (67). While most of the genes of the upper *bph* pathway are located on two of three large linear plasmids (72), pRHL1 (1,100 kb) and pRHL2 (450 kb), genes encoding related isozymes are distributed throughout the 9.7-Mb genome. It is unclear which of these isozymes is involved in the catabolism of biphenyl or related compounds, how these different activities are regulated, and whether this apparent redundancy is a general characteristic of catabolic pathways in rhodococci.

Rhodococcus sp. strain RHA1 utilizes benzoate and phthalate as sole sources of carbon and energy. The catabolism of these compounds is initiated by ring-hydroxylating oxygenases encoded by the *ben* (50) and *pad* (Fukuda, personal communication) genes, respectively. In gram-negative bacteria, the initial phthalate dioxygenase mediates 4,5-dihydroxylation (5, 14). In contrast, the phthalate dioxygenase of gram-positive bacteria mediates 3,4-dihydroxylation (26). The involvement of the *ben* and *pad* genes implies that benzoate and phthalate are catabolized via catechol and protocatechuate, respectively, in RHA1.

In other bacteria, both protocatechuate and catechol can be further degraded via either *meta*- or *ortho*-cleavage pathways. The *ortho*-cleavage pathway, commonly known as the β-ketoadipate pathway, is widely conserved among diverse soil bacteria, and separate branches catabolize catechol and protocatechuate (37). Permutations of the pathway occur in different bacterial groups with respect to enzyme distribution (isozymes and points of convergence), regulation, and gene organization (37). For example, at least three arrangements of the branches

* Corresponding author. Mailing address: Dept. of Microbiology and Immunology, University of British Columbia, #300-6174 University Blvd., Vancouver, BC, V6T 1Z3, Canada. Phone: (604) 822-0042. Fax: (604) 822-6041. E-mail: leltis@interchange.ubc.ca.

TABLE 1. Strains and plasmids used in this study

Strain or plasmid	Relevant genotype/comments	Reference or source
Plasmids		
pKD46	λ-RED (*gam, bet, exo*) *araC rep101*(Ts) Amp^r	18
pIJ773	*aac(3)1V oriT*, source of Apra^r cassette	35
pUZ8002	*tra neo* RP4, Km^r	64
pUC-Hy	Source of *hyg* gene, Amp^r	56
RF00111bAO4	Fosmid clone carrying *pcaL*, Cm^r	84
RFMD1	RF00111bAO4 with *cat* replaced with *hyg*, Hm^r	This study
RFMD2	RFMD1 with *pcaL* replaced with Apra^r cassette, Hm^r Apra^r	This study
Strains		
E. coli BW25113	K-12 derivative: Δ*araBAD* Δ*rhaBAD* Amp^r	35
E. coli DH10B	Host for pUZ8002 and RFMD2, Km^r	35
Rhodococcus sp. strain RHA1	PCB degrader	69
Rhodococcus sp. strain Rha1_005	RHA1 Δ*pcaL* Apra^r	This study

have been reported in different genera: in *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*), the two branches converge at β-ketoadipate (44); in pseudomonads, they converge at an upstream enol-lactone intermediate (42); and in *Acinetobacter* sp. strain ADP1 (formerly *Acinetobacter calcoaceticus*), the branches do not converge at all (23).

As part of an effort to better understand the metabolism and physiology of rhodococci and related strains, we are sequencing and annotating the genome of *Rhodococcus* sp. strain RHA1 and studying this organism using a number of functional genomic approaches. One important methodology that has not been well developed for rhodococci is targeted gene deletion. The development of reliable methodologies has been hampered in part by the genetic instability and nonhomologous recombination typical of actinomycetes (53). Insertion mutagenesis has been used to disrupt genes by single crossover followed by genetic complementation (50, 57, 68, 78). However, this approach often generates polar effects on downstream genes. Using a counterselectable marker, Van der Geize et al. (82) constructed unmarked gene deletions in *Rhodococcus erythropolis* SQ1. Recently, Gust et al. described a strategy using λ Red-mediated double-crossover recombination to create in-frame, nonpolar gene deletions in *Streptomyces coelicolor* A3(2) (35). In this approach, the gene is first replaced in a cosmid carrying the genomic DNA of interest. The mutagenized cosmid is then introduced into the streptomycete to effect allelic exchange. The method enables the deletion of entire gene clusters as well as single and multiple deletions. In principle, a similar approach could be used in *Rhodococcus*.

The current study describes an investigation of the catabolism of phthalate and benzoate in *Rhodococcus* sp. strain RHA1. Analyses of the nearly completed genome assembly revealed the presence of several putative operons that are involved in the catabolism of these aromatic carboxylic acids. Proteomic analyses of cells grown on pyruvate, phthalate, and benzoate identified pathway genes and enzymes. The involvement of key enzymes was confirmed by activity assays and gene disruption. The latter was accomplished by adapting a gene replacement methodology developed for *Streptomyces* and should be of general utility to study *Rhodococcus* and other actinomycetes. To investigate aspects of pathway and gene organization that might be unique to actinomycetes, the deduced pathways in RHA1 were compared with those reported for other bacteria and what could be deduced from genomic sequence data. These studies provide insights into the origin of the catabolic diversity of rhodococci.

MATERIALS AND METHODS

Chemicals. Pharmalyte 3-10 and Immobiline DryStrips were purchased from Amersham Biosciences (Baie d'Urfé, Canada). Iodacetamide, 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) and Sypro Ruby were from Acros Organics (Belgium), Pierce (Rockford, IL), and Bio-Rad Laboratories Ltd. (Mississauga, Canada), respectively. Oligonucleotides were purchased from QIAGEN (Mississauga, Canada). All chemicals were of analytical grade and used without further purification.

Strains, media, and plasmids. The bacterial strains and plasmids used in this study are listed in Table 1. *Rhodococcus* sp. strains RHA1 and RHA1_005 were grown at 30°C on W medium (57) supplemented with 20 mM of an appropriate carbon source (benzoate, phthalate, or pyruvate). For proteomic studies and enzyme assays, rhodococcal strains were typically grown as 500-ml cultures in 2-liter Erlenmeyer flasks shaken at 200 rpm to mid-log phase as determined by optical density at 600 nm. Cells were harvested by centrifugation for 10 min at 16,887 × g at 25°C. The cell pellets were flash frozen in liquid nitrogen and stored at −80°C.

For gene replacement, *Escherichia coli* and *Rhodococcus* strains were grown using media and culture conditions described previously (35). *E. coli* BW25113 was used to propagate pKD46 and fosmid RF00111bAO4. *E. coli* DH10B/pUZ8002 was the fosmid donor strain for intergeneric conjugation. Strains were grown on LB broth supplemented with ampicillin (100 μg/ml), apramycin (Apra, 50 μg/ml), chloramphenicol (12.5 μg/ml), kanamycin (50 μg/ml), and hygromycin (50 and 150 μg/ml for RHA1 and *E. coli*, respectively) as required. To compare the growth rates of RHA1 and RHA1_005, cells were grown in 250-ml flasks containing 100 ml of W medium supplemented with 20 mM of pyruvate, benzoate, or phthalate. Cultures of RHA1_005 were analyzed by PCR using two sets of primers (PCALfor2/PCALrev2 and PCALfor3/PCALrev3 [Table 2]) to confirm the *pcaL* deletion and its stability.

Preparation of cell extracts. Cellular preparations were maintained at 4°C unless otherwise noted. For two-dimensional gel electrophoresis, the cell pellets were washed three times in saline (0.14 M NaCl), once in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0), and stored as aliquots at −80°C. The cells were disrupted by bead-beating. Briefly, a small amount of lysis buffer (8 M urea, 4% CHAPS, 30 mM Tris, pH 8.5, protease inhibitor cocktail [one tablet of Mini Complete per 10 ml solution; Roche]) was added to the cell pellet (1:100, vol/vol). The cells were beaten with 0.5 g of 0.1-mm zirconia/silica beads (BioSpec Products Inc., Bartlesvilles, OK) using a Fast Prep Bio 101 Thermo Savant bead beater for five cycles of 30 s, speed 6.0. Between each cycle, the tubes were cooled on ice. To remove unbroken cells and debris, the preparation was centrifuged at 34,180 × g for 30 min and the supernatant was recentrifuged at 16,100 × g for 10 min. The cell-free protein extract thus obtained was either stored at −80°C or used immediately for proteomic studies.

For enzyme assays, cell-free protein extracts were prepared in essentially the same manner except that the cell pellets were washed twice using 50 mM Tris-HCl, pH 7.0, and the extracts were used immediately. Protein concentration was determined using the 2D Quant kit (Amersham Biosciences).

TABLE 2. PCR primers

Primer	Sequence (5′-3′)	Comments
HYGfor1	GAGTTATCGAGATTT TCAGGAGCTAAGGAAGCTAAAATGGG GTACCAAGCCCTCGGCGA	Hm^r cassette amplification, forward
HYGrev1	GTAGCAACCAGGCGTTTAAGGGCACCAATAACTGCCTTAGG GGCGTCAGGCGCCGGG G	Hm^r cassette amplification, reverse
HYGfor2	TGATCGGCACGTAAG AGG	External *cat* primer, forward
HYGrev2	CATGTT TGACAGGTTATCATC G	External *cat* primer, reverse
HYGfor3	GT CAT CAA GCT GTT CGG	Internal *hyg* primer, forward
HYGrev3	GAA GGC GTT GAG ATG CAG	Internal *hyg* primer, reverse
PCALfor1	ACTGGAGACAGTGGCCCGCGGCCGGGAAGGAACGCGATGA TTCCGGGGATCCGTCGACC	Apra^r cassette amplification, forward
PCALrev1	GGGACGGGAGGATCTTCTCCGTCGGTTCGGTAGCGGTCATG TAGGCTGGAGCTGCTTC	Apra^r cassette amplification, reverse
PCALfor2	AAACGC TCACGGACCTGCTAC	External *pcaL* primer, forward
PCALrev2	CGACGAGGGTGA GCAGAAATC	External *pcaL* primer, reverse
APRAfor	CGAGAGCAGGATTCCCGTTGAG	Internal *apra* cassette primer, forward
APRArev	ATTGCACTCCAC CGCTGATGAC	Internal *apra* cassette primer, reverse
PCALfor3	GTCGGTCATGCAGATCCTCG	Internal *pcaL* primer, forward
PCALrev3	CTGTTTGCCGAGTGCGAAG	Internal *pcaL* primer, reverse

Two-dimensional gel electrophoresis. Protein two-dimensional gel electrophoresis was performed as previously described (31, 32), with the following modifications. The first-dimensional separation was carried out using nonlinear immobilized pH gradient (IPG) strips (24 cm, pH 3 to 7). The strips were rehydrated with the protein sample (90 μg of protein extract) in 400 μl rehydration solution (10 M urea, 2 M thiourea, 30 mM dithiothreitol, 3% CHAPS, Pharmalyte pH 3 to 10). Isoelectric focusing in the IPG strips was carried out for a total of 73.5 kVh at 20°C under mineral oil using ETTAN IPGphor (Amersham Biosciences). The IPG strips were then equilibrated and run into 12% sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) gels (24 by 20 cm) using the ETTAN DALTtwelve System (Amersham Biosciences). Broad-range molecular mass markers (Invitrogen) were run on each side of the gel. Protein was detected using silver nitrate (for screening purposes) or Sypro Ruby (for quantitative analysis). Silver- and Sypro Ruby-stained gels were imaged using a flat-bed image scanner and a variable mode imager Typhoon 9400 (excitation 488 nm, emission 610 nm; Amersham Biosciences), respectively.

Analysis of two-dimensional gels. Two-dimensional gels were differentially analyzed using Progenesis Workstation software (Nonlinear Dynamics, Durham, NC). Accordingly, each image was processed as follows: (i) sharp spikes were removed from the image as noise; (ii) background values were calculated using a mathematical surface model and subtracted; (iii) spots were detected; (iv) a signal intensity was assigned to each spot; and (v) the signal intensity of each spot was normalized against the total signal intensity of the gel. Processed gel images were matched using the combined warping and matching algorithm. Finally, the signal intensity of each spot was averaged over gels obtained from three biological replicates. In the current experiments, only spots with a minimum normalized volume of 0.002 or greater were analyzed. Molecular mass and isoelectric point values were assigned using the calibration standards. Protein spots whose intensities increased or decreased at least twofold versus the control (pyruvate-grown cells) were recorded as more or less abundant, respectively.

Protein identification. Proteins were identified based on peptide mass and/or peptide fragment mass fingerprint analyses (mass spectrometry [MS] and/or MS/MS). Spots of interest were excised from Sypro ruby-stained gels and digested in-gel using trypsin (49). Mass spectrometry analyses were performed using either a Voyager DESTR matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) or a Sciex linear ion trap quadrupole liquid chromatography (LC)-MS/MS (Applied Biosystems). Proteins were identified using the MASCOT search engine (www.matrixscience.com) and a database generated by in silico digestion of the total RHA1 proteome predicted from build 34.03 of the genome assembly (http://www.bcgsc.bc.ca/cgi-bin/rhodococcus/blast_rha1.pl).

Searches were performed without constraining protein molecular mass or isoelectric point and allowing for the following modifications: carbamidomethylation of cysteine, partial oxidation of methionine residues, and up to one missed trypsin cleavage. A protein was considered identified if the hit fulfilled four criteria: the hit was statistically significant (a MASCOT search score above 55 for the RHA1 database), the number of matched peptides was four or higher, the protein sequence coverage was above 20%, and predicted molecular mass and isoelectric point values were consistent with the experimentally determined ones. Theoretical molecular masses and isoelectric points of the proteins of interest were calculated using EXPASY tools (8).

Enzyme assays. Enzyme assays were performed using a Varian Cary IE UV-visible spectrophotometer equipped with a thermostatted cuvette holder. Protocatechuate 3,4-dioxygenase activity was determined by monitoring the transformation of protocatechuate to β-carboxy-*cis*,*cis*-muconic acid at 290 nm ($\varepsilon = 2.3$ mM^{-1} cm^{-1} (41)) in an assay mixture containing 50 mM Tris-HCl buffer (pH 8.8) and 160 μM protocatechuate. Catechol 1,2-dioxygenase activity was measured by monitoring the formation of *cis*,*cis*-muconate at 260 nm ($\varepsilon = 16.8$ mM^{-1} cm^{-1} (22)) in an assay mixture containing 50 mM Tris-HCl buffer (pH 8.8) and 200 μM catechol. The activity of β-ketoadipate:succinyl-coenzyme A (CoA) transferase was measured by monitoring the increase in absorbance at 305 nm ($\varepsilon = 16.3$ mM^{-1} cm^{-1}) using an assay mixture containing 35 μM Tris-HCl buffer (pH 8.0), 25 μM $MgCl_2$, 3.5 μM β-ketoadipate, and 0.15 μM succinyl-CoA (46). In all cases, 1 unit of enzyme activity was defined as the amount of enzyme required to produce 1 μmol of product per minute at 25°C.

Gene replacement on fosmids. Fosmid RF00111bAO4 (Fig. 1b), created as part of a fosmid library of the RHA1 genome (84), contains 46 kb of RHA1 genomic DNA, including *pcaL*. The parent vector, EpiFOS (Epicentre), carries a Cm^r gene, *cat*, that is ineffective in RHA1. Accordingly, the λ RED-based methodology was employed to replace the *cat* gene of RF00111bAO4 with an Hm^r gene, *hyg*, and *pcaL* with an Apra^r cassette.

The cassette used to replace the *cat* gene on RF00111bAO4 (Fig. 1c) was constructed by PCR-amplifying *hyg* using the primers HYGfor1 and HYGrev1 (Table 2). The template for amplification was a gel-purified 1,480-bp SmaI fragment of pUC-Hy (56). The 3′ ends of the PCR primers matched 20-nucleotide and 19-nucleotide extensions, respectively, of the *hyg* gene sequence. The 5′ ends contained 39 nucleotides flanking the target gene. For HYGfor1, these corresponded to the sense strand and ended in the start codon of the *cat* gene. For HYGrev1, these corresponded to the antisense strand and ended in the stop codon of *cat*.

Amplification was performed in a 50-μl reaction with 100 ng of template, 5% dimethyl sulfoxide, 50 pmol each primer, and 200 mM deoxynucleoside triphosphates as described (35). Electrocompetent cells of *E. coli* BW25113, containing the λ RED recombination plasmid pKD46 (Amp^r), were transformed with the fosmid and selected on ampicillin and chloramphenicol at 30°C to prevent the loss of pKD46. Transformants were used to prepare electrocompetent cells grown at 30°C in SOB (68a) containing ampicillin, chloramphenicol, 20 mM $MgSO_4$ and 10 mM L-arabinose. The latter induces the RED genes. Competent cells were electrotransformed with the PCR-extended Hm^r cassette and selected on LB containing hygromycin and ampicillin at 37°C to induce the loss of pKD46. Successful transformants (containing a copy of RFMD1, the Hm^r fosmid) were checked by PCR (reactions with 4% dimethyl sulfoxide for 30 cycles; 45 s at 94°C, 1 min at 60°C, and 90 s at 72°C in a 50-μl reaction mixture) using two sets of primers: HYGfor2 and HYGrev2 and HYGfor3 and HYGrev3 (Table 2).

The same approach was used to replace the *pcaL* gene on RFMD1 with an Apra^r gene flanked by FRT sequences (Fig. 1c). The template for amplification was a gel-purified 1,384-bp EcoRI/HindIII fragment of pIJ773. The 3′ ends of

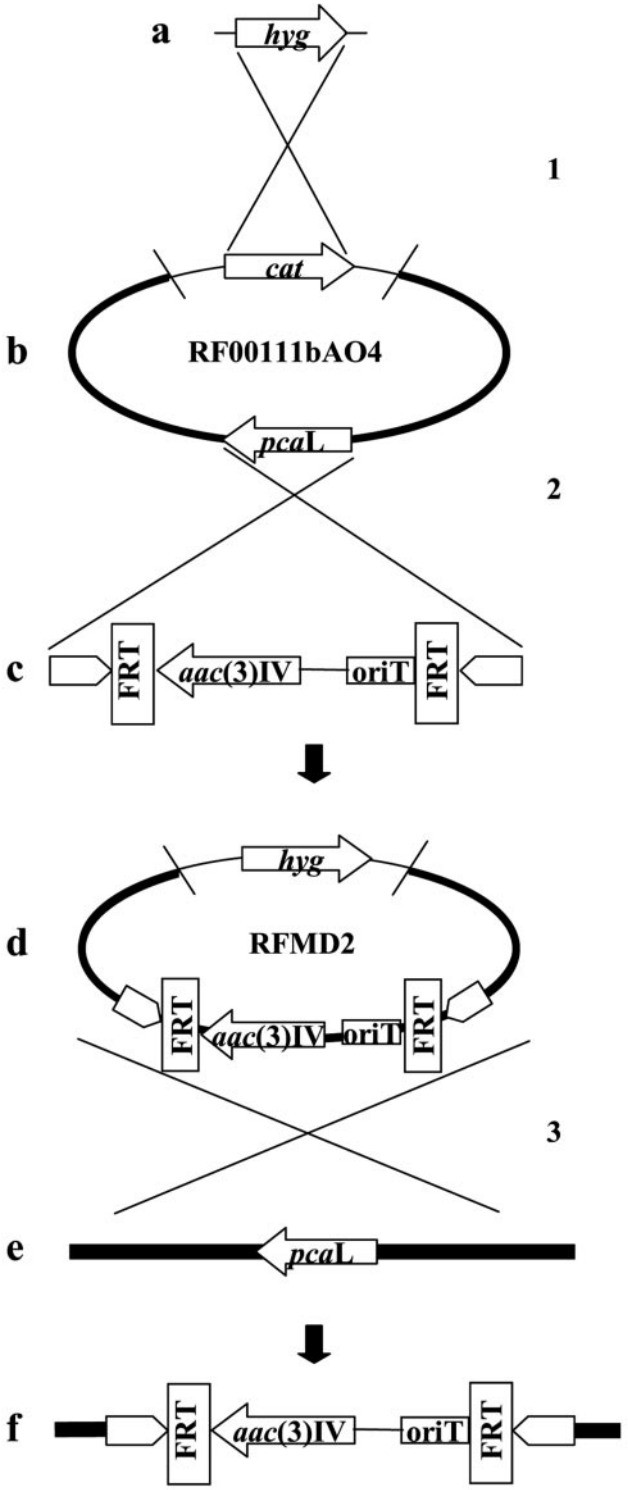


FIG. 1. Outline of the targeted gene disruption strategy. The method used to disrupt *pcaL* was adapted from the λ-Red-based system (35) and involves three major steps: (1) Replacement of the resistance gene in RF00111bAO4. The fosmid RF00111bAO4 (b) carries the chloramphenicol resistance gene (*cat*) and a 46-kb insert of RHA1 genomic DNA containing *pcaL*. An amplicon was generated containing the hygromycin resistance gene (*hyg*) flanked by the same 39-nucleotide sequences that flank the *cat* gene. The *cat* gene was replaced by electroporating the *hyg* amplicon into *E. coli* BW25113 carrying pKD46 and RF00111bAO4 and inducing the λ-Red system. (2) Replacement of *pcaL* on the fosmid. An amplicon (c) containing the apramycin resistance gene [*aac(3)IV*] flanked by the same 39-nucleotide sequences that flank *pcaL*. Allelic replacement in the fosmid was achieved as described in step 1, yielding RFMD2 (d). (3) Replacement of the *pcaL* gene on RHA1 chromosome. RFMD2, carrying the *hyg* gene and the disrupted *pca* cluster, was conjugated into RHA1 cells. Allelic exchange between the fosmid and the chromosome (e) resulted in *pcaL* replacement with Apra^r cassette (f). The latter was selected by screening for apramycin resistance and hygromycin sensitivity and verified by PCR. The Apra^r cassette contains *oriT* and FRT sites. *oriT* allows conjugal transfer into RHA1. The FRT sites, allowing FLP recombinase-mediated elimination of the disruption cassette, were not used in the current study. Additional details are provided in Materials and Methods.

the PCR primers (PCALfor1 and PCALrev1) matched 19-nucleotide and 20-nucleotide extensions, respectively, of the pUJ773 sequence flanking the FRT sequences. The 5′ end of PCALfor1 contained 39 nucleotides corresponding to the sense strand upstream of *pcaL* ending in the start codon. The 5′ end of PCALrev1 contained 39 nucleotides corresponding to the antisense strand ending in the stop codon. Transformants of *E. coli* BW25113 containing pKD46 and RFMD1 were selected on ampicillin and hygromycin. Cells transformed with the PCR-extended Apra^r cassette were selected on LB containing hygromycin plus apramycin at 37°C. Successful transformants containing a copy of the Hm^r- and Apra^r-mutagenized fosmid, RFMD2, were checked by PCR using three sets of oligonucleotides: PCALfor2 and PCALrev2; APRAfor and APRArev; and PCALfor2 and APRArev (Table 2).

Allelic exchange in RHA1. Fosmid RFMD2 was isolated, purified, and introduced into DH10b/pUZ8002 by electroporation. The fosmid was then transferred to *Rhodococcus* sp. strain RHA1 by intergeneric conjugation as described in (47). Apra^r Hm^s exconjugants were selected as double-crossover mutants. The mutants were analyzed by PCR with three sets of primers: PCALfor2 and PCAL rev2; PCALfor2 and APRArev; and PCALfor3 and PCALrev3 (Table 2). The frequency of double-crossover mutants among the exconjugants was ≈12%.

Bioinformatic analyses. Gene prediction and annotation were performed using integrated automated and manual approaches developed at Oak Ridge National Laboratories. Briefly, the automated step made use of three gene finders: Critica v.1.05 (3), Glimmer (20), and Generation. The annotation and locations of predicted open reading frames (ORFs) of interest were then evaluated using a variety of tools. ORF function and position were confirmed with BLASTP sequence alignments (1) to NCBI-nr and PFAM, TIGRfam, COGS, KEGG protein databases. Interproscan (2) was used with the ProfileScan, BLASTProDom, HMMPfam, HMMSMART, and ScanRegExp databases to search for conserved domains and motifs and to validate predicted gene function. PSI-BLAST alignments were used to identify ORF function that was not predicted by the above-mentioned searches. When BLAST alignments were performed, the global percent identity (over the full sequence length) was recorded. Finally, for the genes whose protein products were identified, the sequence information of signature peptides was used to verify gene coordinates.

RESULTS

Identification of phthalate and benzoate catabolic genes. *Rhodococcus* sp. strain RHA1 utilized benzoate and phthalate as sole sources of carbon and energy. In 500-ml cultures at 30°C, the strain grew faster on 20 mM benzoate ($\mu = 0.13 \pm 0.02\ h^{-1}$) than on 20 mM phthalate ($\mu = 0.07 \pm 0.02\ h^{-1}$) or pyruvate ($\mu = 0.05 \pm 0.01\ h^{-1}$).

A search of the current RHA1 genome assembly (www.rhodococcus.ca) revealed four clusters of genes which together could encode the pathways responsible for the catabolism of benzoate and phthalate in RHA1 (Fig. 2). The chromosomal *ben* genes encode a dioxygenase (*benABC*), dihydrodiol dehydrogenase (*benD*), and transporter (*benK*) as previously reported (50) (Fig. 3a). However, the sequence of *benA* reported here and supported by proteomic data (see below) differed to the published sequence (gi 16506124). It is

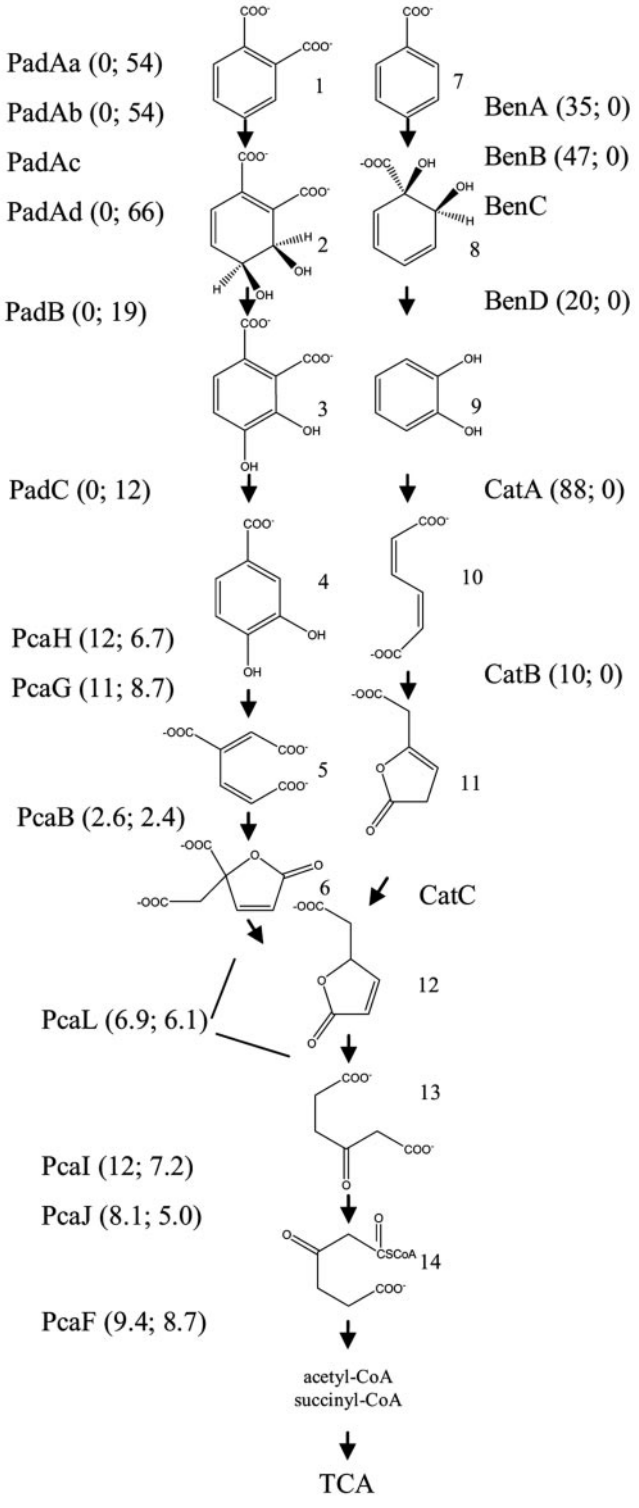


FIG. 2. Proposed pathways of benzoate and phthalate degradation in *Rhodococcus* sp. strain RHA1. Catabolites are numbered as follows: 1, phthalate; 2, phthalate 3,4-dihydrodiol; 3, 3,4-dihydroxyphthalate; 4, protocatechuate; 5, γ-carboxymuconate; 6, γ-carboxymuconolactone; 7, benzoate; 8, *cis*-1,6 dihydroxy-2,4-cyclohexadiene-1-carboxylic acid; 9, catechol; 10, *cis*, *cis*-muconate; 11, muconolactone; 12, β-ketoadipate enol-lactone; 13, β-ketoadipate. The molecular mass-corrected normalized volumes of each protein observed on benzoate and phthalate are indicated in parentheses. In this figure, a value of 0 indicates that the protein was not detected on the gel.

unclear whether these differences represent sequencing errors or spontaneous mutations.

Interestingly, the RHA1 genome sequence reveals that the putative *ben* operon includes genes predicted to encode a cytochrome P450 monooxygenase and its cognate reductase (Table 3) immediately upstream of *benA*. The reductase, encoded by *fpr254A1*, is predicted to harbor a 2Fe-2S cluster and flavin adenine dinucleotide/NAD domains. This is the second putative class V P450 system (66) found in RHA1 to date (84). The N-terminal 22 amino acids of the *cyp254A1*-encoded oxygenase shares 85% sequence identity with N-terminal peptide of the cytochrome $P450_{2EP}$ that catalyzes dealkylation of 2-ethoxyphenol (27), strongly suggesting that these enzymes have similar substrate specificities.

Two AraC-type regulatory genes are located upstream of the putative *ben* operon. The encoded proteins each share ≈30% sequence identity with BenR from *Pseudomonas putida* (17) and may regulate the transcription of the *ben* and *cyp* genes, respectively. However, we could not assign specific functions based on the current data. Inspection of the 183 bp between *fpr254A1* and *benA* revealed a 14-bp repeat centered at −94 bp with respect to the mapped transcription start (50). Two *benK* homologs and two additional *benR* homologs were found in the RHA1 genome (http://www.rhodococcus.ca/publications/supplementary/JBact05B.pdf).

Identical copies of a *pad* cluster were found on plasmids pRHL1 and pRHL2 (Fig. 3b). The cluster contains seven genes predicted to encode a regulatory protein (*padR*), and the enzymes that transform phthalate to protocatechuate: a 3,4-dioxygenase (*padAaAbAcAd*; in gene names, lowercase letters are used to designate subunits and numbers are used to designate isozymes), a dehydrogenase (*padB*), and a decarboxylase (*padC*) (Fig. 2, Table 3). These ORFs share 99% sequence identity with putative ORFs carried by pDK3, the 750-kb plasmid of *Rhodococcus* sp. DK17 (accession number AY502076).

Further analysis of the nucleotide sequence surrounding the *pad* genes revealed that they are part of a 32.1-kb duplication which contains 34 predicted ORFs organized in four major clusters flanked by transposase-encoding genes (Fig. 3b; Table 3). The *pad* cluster is at the 3′ end of this duplication. Two apparently related upstream clusters were named *pat* and *tpa* based on their predicted catabolic functions. The *pat* genes were predicted to encode four subunits of an ABC-type *p*htha*l*ate *t*ransporter system (*pa*tDACB) and a phthalate ester hydrolase (*patE*). The predicted gene products are 62 to 71% identical to the corresponding *ptr*–encoded proteins from *Arthrobacter keyseri* 12B (25). In this strain, the *ptr* cluster is flanked by operons involved in phthalate and protocatechuate catabolism.

The *tpa* genes were predicted to be involved in *t*ere*p*hthal*a*late catabolism. These include *tpaAaAb*, whose products share 68% sequence identity to the α and β subunits of terephthalate 1,2-dioxygenase from *Delftia tsuruhatensis* T7 (71), *tpaB*, which likely encodes the cognate reductase, *tpaC*, predicted to encode a dehydrogenase, and *tpaK*, predicted to encode an aromatic acid permease. The cluster also includes a divergently transcribed gene whose product shares 39% identity with PcaR, an IclR-type transcriptional regulator. The fourth cluster in the 32.1-kb duplication lies several kb upstream of the *tpa* cluster. Although several of the predicted gene products

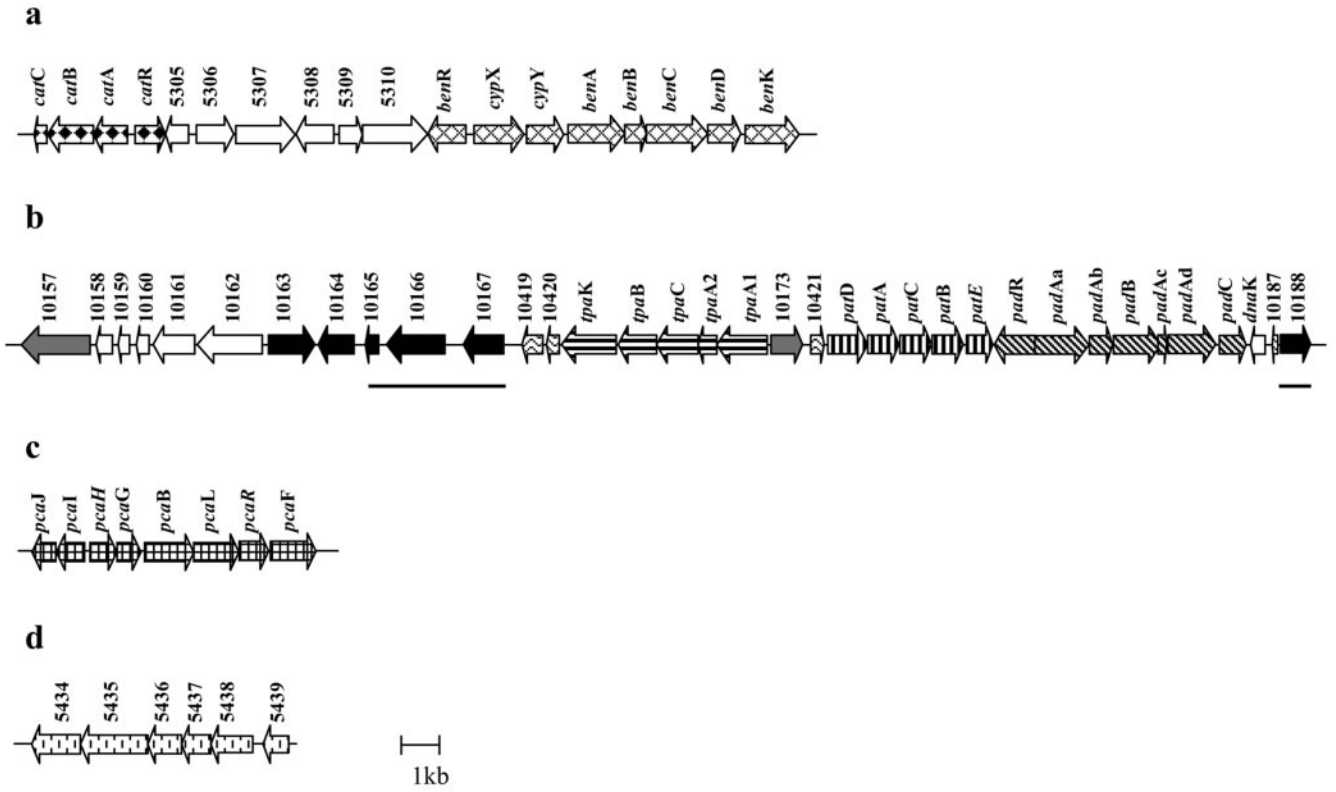


FIG. 3. Benzoate and phthalate catabolic genes in *Rhodococcus* sp. strain RHA1. (a) the chromosomal *cat-ben* cluster, (b) the *pad-pat-tpa* cluster found on pRHL1 and pRHL2, (c) the chromosomal *pca* cluster, and (d) a cluster of unknown function that contains a *catD* homolog. Arrows represent different types of genes as follows: diagonally striped, *pad* cluster; vertically striped, *pat* cluster; horizontally striped, *tpa* cluster; diamonds, *cat* cluster; hatched diamonds, *ben* cluster; squares, *pca* cluster; grey, transcriptional regulator; black, DNA recombination; white, miscellaneous; and waved, hypothetical. Numbers above the arrows indicate the ORF numbers (Table 1) from the RHA1 genome assembly (www.rhodococcus.ca). Solid lines below the *pad-pat-tpa* cluster indicate the two sections that are present in pRHL2 but not in pRHL1. The products of the depicted genes identified in the study are listed in Table 4.

show high similarity to catabolic enzymes such as cyclohexanone monooxygenase, the precise biological function of these genes is unknown.

The 32.1-kb duplication also includes genes predicted to encode an integrase (rha10163), 4 transposases (rha10188, rha10164, rha10165, and rha10167), and a reverse transcriptase (rha10166). Three of these genes are missing from the corresponding gene cluster on pRHL1 (solid bars in Fig. 3b). Moreover, the entire 32.1-kb region found on pRHL1 shares 99% sequence identity with a similar region of pDK3 of *Rhodococcus* sp. DK17. Indeed pDK3 and pRHL1 share two regions of 99% sequence identity: a 73-kb region that contains the *tpa*, *pat*, and *pad* clusters and a 22-kb region (data not shown). The similarity of the copy of the *tpa-pat-pad* island present on pRHL1 to that of pDK3 as well as the presence of additional transposase genes in the copy found on pRHL2 suggest that the copy on pRHL1 is the original in RHA1.

The RHA1 genome contains several genes potentially involved in catechol and protocatechuate catabolism. The chromosome contains genes predicted to encode two branches of the β-ketoadipate pathway: the *cat* and *pca* genes specify the catabolism of catechol and protocatechuate, respectively. The putative *catRABC* operon is located within 7 kb of the *ben* genes (Fig. 3a) and shares over 96% sequence identity and the same organization as the corresponding genes in *Rhodococcus opacus* 1CP (28) (Table 3). The three encoded enzymes, catechol 1,2-dioxygenase (*catA*), *cis*,*cis*-muconate lactonizing enzyme (*catB*), and muconolactone isomerase (*catC*), are predicted to transform catechol to β-ketoadipate enol-lactone but not to trichloroacetic acid (TCA) cycle intermediates. In contrast, the *pca* cluster (Fig. 3c) includes genes predicted to encode the enzymes (PcaJIHGBLF) required to convert protocatechuate to the TCA cycle intermediates (Fig. 2). The predicted products of these genes share high 97 to 99% sequence identity with characterized homologs from related actinomycetes (Table 3).

The organization of the *pca* genes in two putative divergently transcribed operons (*pcaJI* and *pcaHGBLRF)* is similar to their organization in *R. opacus* 1CP (29). Of particular note, *pcaL* in RHA1, *R. opacus* sp. 1CP, and *Streptomyces* sp. 2065 (41) appears to encode a bifunctional β-ketoadipate enol-lactone-hydrolyzing enzyme. The gene appears to have arisen from a fusion of *pcaD* and *pcaC*, which encode β-ketoadipate enol-lactone hydrolase and γ-carboxymuconolactone decarboxylase, respectively, in *Acinetobacter* sp. ADP1 and pseudomonads (reviewed in reference 43).

Paralogs of the *cat* and *pca* genes found on the RHA1 genome include two homologs of *catA*, two of *pcaIJ*, two of

TABLE 3. Gene annotation[a]

Gene	Gene Product	Length (aa)	Accession no. (reference)	Best hit (% id)	Organism
cyp254A1	Cytochrome P450, oxygenase	396	gi|36958773 (76)	35	*Rhizobium* sp. NGR234
fpr254A1	Cytochrome P450, reductase	353	gi|527552 (38)	38	*P. putida* H
benR	AraC-type transcriptional regulator	332	gi|576669 (59)	30	*R. erythropolis* NI86/21
catA	Catechol 1,2 dioxygenase	280	gi|2398777 (28)	98	*R. opacus* 1CP
catB	Muconate cycloisomerase	384	gi|5915882 (28)	97	*R. opacus* 1CP
catC	Muconolactone isomerase	93	gi|5915884 (28)	97	*R. opacus* 1CP
catR	Ic1R-type transcriptional regulator	256	gi|2398776 (28)	96	*R. opacus* 1CP
padAb	Phthalate 3,4-dioxygenase, β subunit	208	gi|37518569	73	*M. vanbaalenii* PYR-1
padAc	Phthalate 3,4-dioxygenase, ferredoxin	65	gi|13242056 (25)	73	*A. keyseri* 12B
padAd	Phthalate 3,4-dioxygenase, ferredoxin reductase	413	gi|37518573	54	*M. vanbaalenii* PYR-1
padB	Phthalate 3,4-dihydrodiol dehydrogenase	371	gi|37518571	73	*M. vanbaalenii* PYR-1
padC	3,4-Dihydroxyphthalate decarboxylase	243	gi|13242058 (25)	56	*A. keyseri* 12B
padR	Ic1R-type transcriptional regulator	268	gi|37518567	64	*M. vanbaalenii* PYR-1
pcaH	Protocatechuate 3,4-dioxygenase, α subunit	237	gi|2935025 (29)	97	*R. opacus* 1CP
pcaG	Protocatechuate 3,4-dioxygenase, β subunit	224	gi|2935024 (29)	55	*R. opacus* 1CP
pcaB	γ-Carboxy-*cis,cis*-muconate lactoning enzyme	453	gi|2935026 (29)	97	*R. opacus* 1CP
pcaL	β-Ketoadipate enol-lactone hydrolase/4-carboxy-mucono-lactone decarboxylase	400	gi|2935027 (29)	97	*R. opacus* 1CP
pcaI	β-Ketoadipate:succinyl-CoA transferase, α subunit	262	gi|15609641 (16)	75	*M. tuberculosis* H37Rv
pcaJ	β-Ketoadipate:succinyl-CoA transferase, β subunit	213	gi|15609640 (16)	69	*M. tuberculosis* H37Rv
pcaF	β-Ketoadipate:succinyl-CoA thiolase	408	gi|33945696	49	*Pseudomonas* sp.Y2
pcaR	Ic1R -type transcriptional regulator	265	gi|2935028 (29)	98	*R. opacus* 1CP
10157	σ^{54}-Dependent transcriptional regulator	586	gi|27657630 (12)	29	*Rhodococcus* sp. Phi2
10158	Acetate-CoA ligase	135	gi|27366018 (48)	13	*V. vulnificus* CMCP6
10159	Propionate-CoA ligase	100	gi|2120623 (74)	8	*P. aeruginosa*
10160	Ethyl *tert*-butyl ether (ETBE)-induced protein	103	gi|16551197 (15)	44	*R. ruber* IFP2001
10161	Homoserine *O*-acetyltransferase	353	gi|12230253 (9)	27	*L. meyeri*
10162	Cyclohexanone monooxygenase	546	gi|11356670 (58)	41	*R. rhodochrous*
10163	Integrase	384	gi|8118238 (19)	45	*P. aeruginosa* PAO17
10164	Transposase	289	gi|4741891	43	*S. meliloti* 220-12
10165	Transposase	122	gi|4838457 (54)	10	*R. erythropolis* SQ1
10166	RNA-directed DNA polymerase	490	gi|16263485 (4)	38	*S. meliloti* 1021
10167	Transposase	338	gi|4838457 (54)	27	*R. erythropolis* SQ1
10419	Conserved hypothetical	179	gi|40787275	100	*Rhodococcus* sp. DK17
10420	Hypothetical	104	No significant similarity		
*tpa*K	4-Hydroxybenzoate transporter	459	gi|6093655 (36)	34	*P. putida*
*tpa*B	Ferredoxin-NAD(+) reductase	336	gi|6580713 (73)	37	*Y. pseudotuberculosis* (type O:1b)
*tpa*C	4-Hydroxythreonine-4-phosphate dehydrogenase	338	gi|36958683 (76)	39	*Rhizobium* sp. NGR234
*tpa*A2	Terephthalate 1,2-dioxygenase β subunit	156	gi|28971831 (65)	36	*Sphingomonas* sp. P2
*tpa*A1	Terephthalate 1,2-dioxygenase α subunit	421	gi|22779305 (71)	68	*Delftia* sp. T7
10173	Ic1R-type transcriptional regulator	268	gi|2935028 (29)	38	*R. opacus* 1CP
10421	Hypothetical	74	gi|40787268	98	*Rhodococcus* sp. DK17
*pat*D	ABC transporter, substrate-binding component	317	gi|13242060 (25)	62	*A. keyseri* 12B
*pat*A	ABC transporter, ATPase	270	gi|13242049 (25)	71	*A. keyseri* 12B
*pat*C	ABC transporter, permease	269	gi|13242050 (25)	63	*A. keyseri* 12B
*pat*B	ABC transporter, permease	271	gi|13242051 (25)	64	*A. keyseri* 12B
*pat*E	Phthalate ester hydrolase	221	gi|13242052 (25)	76	*A. keyseri* 12B
*dna*K	Molecular chaperone	119	gi|19387023	15	*Actinomadura citrea* JCM 3295
10187	Membrane transport protein (fragment)	62	gi|40787254	100	*Rhodococcus* sp. DK17
10188	Transposase	236	gi|21734950 (80)	80	*R. erythropolis* MP50
5305	Intradiol dioxygenase	179	gi|2398775 (28)	94	*R. opacus* 1CP
5306	4-Hydroxybenzoyl-CoA thioesterase	315	gi|2398774 (28)	99	*R. opacus* 1CP
5307	Multicopper oxidase	494	gi|24306121 (39)	44	*C. diphtheriae*
5308	AraC-type transcriptional regulator	312	gi|576669 (21)	30	*R. erythropolis* NI86/21
5309	Phenol hydroxylase, reductase	186	gi|33317300	38	*G. thermoleovorans* A2
5310	Phenol hydroxylase, oxygenase	538	gi|3046914 (24)	62	*G. thermoleovorans* A2
5434	Acetyl-CoA C-acetyltransferase	411	gi|475715 (75)	44	*Clostridium acetobutylicum* ATCC 824
5435	*O*-Succinylbenzoate-CoA ligase	568	gi|34495234 (33)	29	*Rhodococcus* sp. NCIMB 9784
5436	3-Oxoadipate enol-lactonase hydrolase, *catD* homolog	282	gi|6166146 (70)	18	*Acinetobacter* sp. ADP1
5437	Hypothetical	249	No significant similarity		
5438	Butyryl-CoA dehydrogenase	397	gi|1703066 (10)	38	*C. acetobutylicum* ATCC

[a] The sequence and annotation of *padAa* and *benABCDK* were reported previously (50, 51) and so are not reported here. Only the experimentally confirmed hits with the highest percent identity were used for gene annotation. Percent identity was calculated over the entire amino acid (aa) sequence length and was based on alignments with BlastP hits from the nonredundant NCBI protein database.

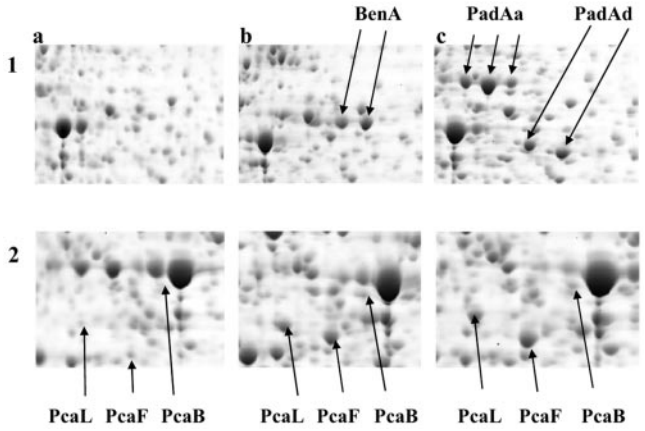


FIG. 4. Sections of two-dimensional gels showing parts of the *Rhodococcus* sp. strain RHA1 cellular proteome. Corresponding gel sections are shown from cells grown on pyruvate (a), benzoate (b), or phthalate (c). The protein spots of interest are shown with arrows. Panel 1 shows the proteins (BenA and PadAaAb) that were only observed in benzoate- or phthalate-grown cells. Panel 2 shows proteins (PcaLFB) that appeared in all the conditions tested but were more abundant in phthalate and benzoate samples.

pcaC encoding decarboxylases, and five of *pcaF* (http://www.rhodococcus.ca/publications/supplementary/JBact05B.pdf). The homologs appear to be functional as key catalytic residues are conserved, although their physiological roles in RHA1 are for the most part unclear. One of the *pcaF* homologs is involved in phenylacetate catabolism and was annotated as *paaE* (60). The annotation of the genes mentioned in this section is also available at www.rhodococcus.ca.

Identification of phthalate- and benzoate-catabolizing enzymes. To identify the pathways responsible for the catabolism of benzoate and phthalate, respectively, in *Rhodococcus* sp. strain RHA1, the cytosolic proteomes of cells grown on phthalate, benzoate or pyruvate as sole carbon source were quantitatively compared by two-dimensional gel electrophoresis. Approximately 1,500 protein spots were resolved per gel (see Fig. 4 for representative gel sections). Some of the proteins appeared on the gel as a horizontal series of spots suggesting that they represent multiple species differing in charge (e.g., PadAa in Fig. 3c). This may be caused by either posttranslational modification within the cell or chemical modification during sample preparation. Mass spectrometric analyses supports the latter for the proteins reported here as changes in the masses of peptides originating from "multiple" protein spots were consistent with carbamylation. Optimization of sample preparation, particularly with respect to temperature and buffer composition, minimized but did not completely eliminate carbamylation. For carbamylated proteins, the isoelectric point and molecular mass of the major spot in a series were recorded, and the expression difference was calculated based on the summed signal intensities of all the spots in the series.

Quantitative comparison of spot intensities on matched gels indicated that within the detected proteomes, 113 proteins were at least twofold more abundant during growth on benzoate versus pyruvate, and 77 were at least twice as abundant during growth on phthalate. Approximately 10% of these proteins were more abundant in both benzoate- and phthalate-grown cells. Of the proteins that were more abundant during growth on one of the aromatic acids, 59 were identified. Those probably involved in phthalate and benzoate catabolism are listed in Table 4 together with their averaged normalized signal intensity observed under each condition. Differential expression data are presented as the ratio (fold difference) between the averaged normalized signal intensity.

Five of the seven predicted *pad* gene products (PadAaAbAdBC) were identified in gels of phthalate-grown cells (Fig. 4c, Table 4). The two that were not found, PadAc and PadR, have predicted isoelectric points (3.69 and 9.72, respectively) that are not compatible with the two-dimensional gel electrophoresis protocols used. Each of the five identified proteins was detected only in phthalate-grown cells (Fig. 2). Two additional proteins that appear to be involved in the catabolism of phthalate were also much more abundant in phthalate-grown cells: PatE, a probable phthalate ester hydrolase, and PatA, the ABC transporter ATPase. PatA was only detected in phthalate-grown cells, strongly suggesting that PatDACB is responsible for phthalate transport.

Three of the four *ben* gene products (BenABD) and two of the four *cat* gene products (CatAB) were identified in benzoate-grown cells (Fig. 4b, Table 4). Each of these five proteins was only observed in benzoate-grown cells (Fig. 2). Failure to identify BenC, CatC, and CatR does not imply their absence in benzoate-grown cells, as ≈50% of the protein spots analyzed by mass spectrometry did not yield usable data. Moreover, CatC is unlikely to be detected using the current two-dimensional gel electrophoresis protocols due to its low molecular mass (10 kDa). The components of the type V cytochrome P450 were not identified on the gels. Moreover, spectrophotometric analyses failed to detect the presence of a cytochrome P450 in benzoate-grown cells (data not shown). Finally, transcriptomic data analysis confirmed that *cyp254A1* and *frp254A1* are not up-regulated in benzoate-grown RHA1 cells (H. Hara and W. W. Mohn, in preparation).

Seven of the nine predicted *pca*-encoded enzymes (PcaIJH GBLF) were identified. These were more abundant in phthalate- and benzoate-grown cells than in the pyruvate-grown controls (Fig. 4, Table 4). Protocatechuate 3,4-dioxygenase (PcaHG) was only detected in the phthalate and benzoate proteomes. The other *pca* products, including the regulatory protein PcaR, were present in low but detectable levels in pyruvate-grown cells (Fig. 2, Table 4).

Other than the enzymes that are involved in benzoate and phthalate catabolic pathways, we identified 31 other proteins that were more abundant in phthalate- and/or benzoate-grown cell samples (http://www.rhodococcus.ca/publications/supplementary/JBact05C.pdf). Importantly, these included none of the other *ben*, *pca*, and *cat* homologues identified in the RHA1 genome. Identified proteins of known physiological functions included TCA cycle enzymes as well as amino acid, fatty acid, and nucleotide biosynthetic enzymes. None of the products of the *tpa* genes, thought to be involved in terephthalate catabolism, were detected. Transcriptomic data indicate that these genes, but not the *pad* genes, are strongly up-regulated in terephthalate-grown cells (H. Hara and W. W. Mohn, in preparation). The more abundant TCA enzymes included citrate synthase, isocitrate dehydrogenase, succinate dehydrogenase, and malate dehydrogenase as well as the α-subunit of succinyl-CoA ligase, an

TABLE 4. Identification of the proteins involved in benzoate and phthalate catabolism.

Protein name[a]	Basis of identification	No. of peptides matched	Sequence coverage (%)	Signal intensity[b]				Source gel for protein identification[c]
				RHA1			RHA1_005, benzoate	
				Pyruvate	Phthalate	Benzoate		
BenA	MS[d]	14	44	ND	ND	1.8	1.5	B
BenB	MS[d]	6	31	ND	ND	0.94	0.76	B
BenD	MS[d]	8	39	ND	ND	0.56	0.60	B
CatA	MS	9	37	ND	ND	2.7	2.8	B
CatB	MS[d]	5	21	ND	ND	0.38	0.16	B
PadAa	MS[d]	16	39	ND	2.9	ND	ND	F
PadAb	MS[d]	9	44	ND	1.3	ND	ND	F
PadAd	MS[d]	11	31	ND	1.2	ND	ND	F
PadB	MSMS	5	21	ND	0.77	ND	ND	F
PadC	MS	7	28	ND	0.30	ND	ND	F
PcaH	MSMS	7	33	ND	0.18	0.31	0.70	B
PcaG	MS	9	35	ND	0.19	0.25	0.52	B
PcaB	MSMS	5	12	0.03	0.11	0.12	0.08	B, F
PcaL	MS	6	19	0.05	0.26	0.29	ND	B
PcaI	MSMS	8	47	ND	0.19	0.32	1.3	B, F
PcaJ	MS[d]	4	23	ND	0.11	0.18	0.66	B, P
PcaF	MS	13	41	0.10	0.37	0.40	0.26	B, F
PcaR	MS[d]	7	43	0.02	0.04	0.04	0.01	B
PatE	MS	9	48	0.01	0.81	0.02	0.07	F
PatA	MS	10	42	ND	1.2	ND	ND	F
CatD homolog	MS	13	34	ND	ND	ND	0.54	B
O-Succinylbenzoate-CoA ligase	MS	13	34	ND	ND	ND	0.39	B
Butyryl-CoA dehydrogenase	MS	19	54	ND	ND	ND	1.3	B
Acetyl-CoA acetyltransferase	MS	16	58	ND	ND	ND	1.1	B

[a] The encoding genes are shown in Figure 3.
[b] Spot signal intensities were normalized and averaged over three replica gels (each from an independent experiment).
[c] The gel from which each protein was picked for MS analyses is identified according to the growth substrate of the culture: benzoate (B), phthalate (F), or pyruvate (P). All proteins were identified using the wild-type strain except for the last four proteins in the table, which were identified using the *pcaL* mutant.
[d] Protein identity confirmed by MSMS analysis.

enzyme that couples the hydrolysis of succinyl-CoA to the synthesis of ATP.

Two enzymes of the central metabolic pathway were more abundant in pyruvate-grown cells: isocitrate lyase, an enzyme of the glyoxylate shunt; and acetate-CoA ligase, one of the two enzymes of the pyruvate oxidation pathway. The more abundant TCA cycle enzymes during growth on benzoate and phthalate is consistent with end products of the β-ketoadipate pathway, succinyl-CoA and acetyl-CoA, feeding directly into the cycle. In contrast, increased abundance of the glyoxylate shunt enzymes during growth on pyruvate is consistent with the general observation that these enzymes are required for net assimilation of carbon when the carbon source enters the TCA cycle solely at the level of acetyl-CoA (52).

Relative protein abundance. To better estimate the relative levels of protein abundance, the normalized volumes (NV) of the protein spots were corrected for their molecular mass in kDa (MWc): (MWc-NV = [NV/MW] × 1,000). For three of the four two-subunit enzymes of the studied pathways, the MWc-NV of the two subunits are in good agreement (Fig. 2). The lack of agreement for the fourth enzyme, PcaIJ, likely reflects the different intensities with which Sypro Ruby interacts with these two denatured polypeptides. Nevertheless, the data revealed several interesting features. First, the Pca proteins were consistently more abundant in benzoate-grown cells compared to phthalate-grown cells (Fig. 2, Table 4). This may reflect the higher growth rate of RHA1 on benzoate. Second, a subset of Pca proteins, PcaBLRF, were detected in pyruvate-grown cells, but PcaHG were not. This indicates that *pcaBLRF* expression may be regulated independently from *pcaHG*. Finally, the data indicate that the most abundant proteins in each pathway branch are the oxygenases BenAB, PadAaAb, CatA, and PcaHG. This may reflect the relatively low cytoplasmic concentration of O_2 in exponentially growing cells.

Enzyme activity. The RHA1 genome sequence and proteomic data suggested that the final steps of phthalate and benzoate catabolism in RHA1 are accomplished by the same set of *pca*-encoded enzymes (Fig. 2). To verify this conclusion, the activities of three enzymes representing different branches of the β-ketoadipate pathway were determined in whole-cell extracts of phthalate-, benzoate-, and pyruvate-grown cells. The specific activity of protocatechuate 3,4-dioxygenase (*pcaHG*), which cleaves protocatechuate to γ-carboxymuconate, was 10- and 6-fold higher in RHA1 cells grown in benzoate and phthalate, respectively, than in pyruvate (Table 5). Catechol 1,2-dioxygenase (*catA*) activity was 133-fold greater in benzoate- and only 1.3-fold higher in phthalate-grown cells than in pyruvate-grown cells. No protocatechuate or catechol *meta*-cleavage activities were detected in extracts of cells grown on phthalate or benzoate. However, catechol 2,3-dioxygenase activity was detected in pyruvate-grown cells (data not shown). β-Ketoadipate:succinyl-CoA transferase (*pcaIJ*) catalyzes the penultimate step in the β-ketoadipate pathway. This activity was detected in all the samples and was twofold higher in cells grown on an aromatic carboxylic acid.

Levels of catechol 1,2-dioxygenase activity agree well with

TABLE 5. Enzyme activities in cell lysates of RHA1 cells grown on different substrates[a]

Enzyme	Enzyme sp act (U mg^{-1}) on growth substrate:		
	Benzoate	Phthalate	Pyruvate
Catechol 1,2-dioxygenase	1,730 ± 70 [130]	17 ± 1 [1.3]	13 ± 1
Protocatechuate 3,4-dioxygenase	2,370 ± 280 [9.6]	1,560 ± 30 [6.3]	247 ± 7
β-Ketoadipate:succinyl-CoA transferase	10.6 ± 0.04 [2.0]	12 ± 1 [2.3]	5.2 ± 0.4

[a] Data are based on three replicates ± standard error of the mean. The fold difference versus pyruvate is shown in brackets.

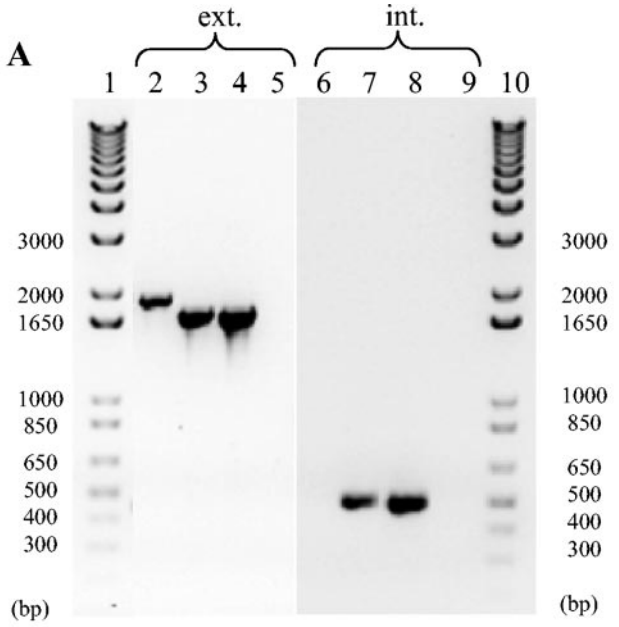
the corresponding abundance of CatA under the three growth conditions (Fig. 2, Table 4). In contrast, the levels of protocatechuate 3,4-dioxygenase and β-ketoadipate:succinyl-CoA transferase activities do not correspond well with the abundances of PcaHG and PcaIJ, respectively. Comparison of the data indicate that this is likely due to the presence of PcaHG and PcaIJ homologs in pyruvate-grown cells. The RHA1 genome contains at least two *pcaIJ* orthologs, as noted above, and seven orthologs of intradiol dioxygenases. Based on sequence identities, it is not possible to predict the substrate specificity of each of these orthologs.

Analysis of *pcaL* gene knockout. The first shared step in the catabolism of phthalate and benzoate in RHA1 is catalyzed by the bifunctional, *pcaL*-encoded β-ketoadipate enol-lactone hydrolase. The essential role of this enzyme in benzoate and phthalate catabolism was tested by investigating the phenotype of RHA1_005, a mutant of RHA1 in which *pcaL* was replaced with an Apra^r cassette using a λ RED-based methodology developed for *Streptomyces* spp. (35) (Fig. 1). RHA1_005 was resistant to apramycin but sensitive to hygromycin, consistent with allelic exchange and loss of the donor fosmid. PCR analysis of RHA1_005 genomic DNA using two sets of primers confirmed the presence of the Apra^r cassette and loss of *pcaL* (Fig. 5). More particularly, primers flanking the *pcaL* gene yielded a larger amplicon from RHA1_005 genomic DNA (lane 2) than from RHA1 genomic DNA (lane 3). PCR performed with primers internal to *pcaL* yielded the expected amplicon from RHA1 genomic DNA (lane 7) and no product from RHA1_005 DNA (lane 6).

The growth rate of RHA1_005 (100-ml cultures) on pyruvate as the carbon source ($\mu = 0.08 \pm 0.01$ h^{-1}) was essentially identical to that of RHA1 ($\mu = 0.09 \pm 0.01$ h^{-1}). However, deletion of *pcaL* completely abolished the ability of RHA1 to grow on phthalate. RHA1_005 was also impaired in its ability to grow on benzoate. However, the principal observed effect was on the duration of the lag phase: when 1 ml of mid-log-phase, pyruvate-grown RHA1_005 cells were used to inoculate 100 ml of W medium containing 20 mM benzoate, exponential growth was not detected until 3 days later. In contrast, wild-type cells grew exponentially within 24 h. Once growing, the growth rates of the two strains were very similar ($\mu = 0.12 \pm 0.005$ h^{-1}). Moreover, once RHA1_005 had been cultured on benzoate, it no longer displayed a prolonged lag phase.

We performed proteomic analyses to investigate the catabolism of benzoate in RHA1_005. Two-dimensional gel electrophoresis analysis confirmed the presence of all *ben*-, *cat*-, and *pca*-encoded proteins detected in benzoate-grown wild-type cells except for PcaL. In addition, 24 proteins were at least fivefold more abundant in benzoate-grown RHA1_005, of which 20 were not detected in benzoate-grown RHA1. Half of these 24 proteins were identified by peptide fingerprint analysis. One of the identified proteins, encoded by rha05436, shares greatest sequence identity (18%) with CatD from *Acinetobacter* sp. ADP1 (70) and 19% sequence identity with the N-terminal hydrolase domain of PcaL (Tables 3 and 4). It is therefore likely that this enzyme functions as a β-ketoadipate enol-lactone hydrolase, compensating for the loss of PcaL in RHA1_005.

Rha05436 resides in a putative operon containing five other genes (Fig. 3D), three of whose products were also more abun-



B	RHA1_005	RHA1	RF00111bAO4
ext	1882 bp	1702 bp	1702 bp
int	-	499 bp	499 bp

FIG. 5. PCR analysis of the *pcaL* replacement mutant RHA1_005. The results and expected fragment sizes are shown in panels A and B, respectively. Reactions were performed with primers that are external (ext., lanes 2 to 5, PCALfor2/PCALrev2) or internal to *pcaL* (int, lanes 6 to 9, PCALfor3/PCALrev3). Reactions contained RHA1_005 genomic DNA (lanes 2 and 6), RHA1 genomic DNA (lanes 3 and 7), fosmid RF00111bAO4 (lanes 4 and 8), or no template DNA (lanes 5 and 9). Lanes 1 and 10 were loaded with molecular mass markers (1-kb DNA ladder, Invitrogen).

dant in RHA1_005 (Table 4). These show sequence similarity to an acetyl-CoA acetyltransferase (CatF homolog), an *O*-succinylbenzoate-CoA ligase, and an acyl-CoA dehydrogenase, respectively (Table 3). None of these proteins were detected in wild-type RHA1 cells under any of the growth conditions studied. Four of the identified proteins appear to be involved in iron acquisition and storage, including two that appear to be involved in the biosynthesis of a catecholic siderophore, possibly an enterobactin-type molecule. Other proteins were identified in benzoate-grown RHA1_005 (http://www.rhodococcus.ca/publications/supplementary/JBact05B.pdf and http://www.rhodococcus.ca/publications/supplementary/JBact05C.pdf), and their current annotation is available at www.rhodococcus.ca.

DISCUSSION

The current study demonstrates that benzoate and phthalate are catabolized by a branched β-ketoadipate pathway in *Rhodococcus* sp. strain RHA1. The genomic sequence and proteomic and gene disruption data indicate that the catechol and protocatechuate branches of the β-ketoadipate pathway in RHA1 converge at β-ketoadipate enol-lactone (metabolite 12 of Fig. 2). The data further indicate that this metabolite is transformed by PcaL, a bifunctional enzyme that comprises a γ-carboxy-muconolactone decarboxylase and an enol-lactone hydrolase in separate domains. In other bacteria, including pseudomonads, these enzymes are encoded by *pcaC* and *pcaD*/*catD*, respectively (43). In RHA1, only the protocatechuate branch utilizes the decarboxylase activity of PcaL. The *pcaL* gene is not essential for growth of RHA1 on benzoate, apparently because the CatD homolog is expressed in the *pcaL* mutant grown on benzoate, compensating for loss of PcaL.

The current data also identify an ABC-type transporter that is likely involved in the uptake of phthalates. Genes encoding a related transporter were identified in *A. keyseri* 12B (25), but no functional data were obtained. The ATPase of the transporter (PatA) and a probable phthalate ester hydrolase (PatE) were highly abundant in the cytoplasm of phthalate-grown cells. Consistent with its cytoplasmic localization, PatE lacked a predicted secretion signal sequence. The localization of this enzyme in the cytoplasm suggests that phthalate esters are hydrolyzed in the cytoplasm after their uptake, either by the *pat*-encoded ABC transporter or by another route. In this respect, symporter-type phthalate permeases are inhibited by substituted phthalates but not structurally related compounds, such as 2-Cl benzoate, that lack one of the carboxylates, suggesting that vicinal carboxylate is an important substrate-binding determinant (13). We are currently investigating the role of PatDACB in the transport of phthalates, terephthalates, and related compounds.

Our comparison of the β-ketoadipate pathway of RHA1 to that of other bacteria in which it occurs revealed several features of the pathway that appear to be unique to rhodococci and other actinomycetes with respect to component enzymes, gene organization, and regulation. Thus, analyses of genome sequences indicate that the same branched β-ketoadipate pathway also occurs in *Streptomyces avermitilis* MA-4680, *Streptomyces coelicolor* A3(2), and *Corynebacterium glutamicum* (45). Moreover a bifunctional PcaL has been reported in *R. opacus* sp. strain 1CP (29) and *Streptomyces* sp. strain 2065 (41). In contrast, the pathway branches do not converge at all in *Acinetobacter* sp. strain ADP1 (23) and converge at a different point, β-ketoadipate, in *Ralstonia eutropha* (44).

The pathway in pseudomonads appears to be similar to that in actinomycetes in that the branches converge at the enol-lactone (42, 61). However, the pseudomonads do not use a bifunctional enzyme at the point of convergence: the γ-carboxymuconolactone decarboxylase and the enol-lactone hydrolase are encoded by *pcaC* and *pcaD*, respectively (43). Nonetheless, the bifunctional PcaL is not unique to actinomycetes: it also occurs in *Ralstonia metallidurans* and *Caulobacter crescentus* (43). Finally, the β-ketoadipate pathways of pseudomonads and actinomycetes appear to differ by the presence of a β-ketoadipate transporter (PcaT) in the former (43).

An analysis of 22 complete or partial bacterial genomes containing the *pca* and *cat* genes revealed that the latter appear to be organized in a fashion that is unique to rhodococci and most similar to that of the closely related corynebacteria. The *pca* genes are organized in a single cluster in all actinomycetes in which they have been found as well as in *Caulobacter crescentus* (62), an α-proteobacterium, and *Acinetobacter* sp. ADP1 (11), a γ-proteobacterium. Actinomycetes containing the *pca* genes include *R. opacus* 1CP (29), *C. glutamicum* sp. ATCC 13032 (BX927155) (45), *Streptomyces* sp. 2065 (41), *S. coelicolor* A3(2) (SC0939128) (7), and *S. avermitilis* MA-4680 (AP005027.1) (63). In contrast, the *pca* genes can be arranged in up to three clusters in pseudomonads (42). Multiple *pca* clusters also occur in β-proteobacteria such as *Burkholderia pseudomallei* and *R. metallidurans* (43).

Nevertheless, the organization of the *pca* genes in a single cluster of two divergently transcribed operons with the gene order of RHA1 appears to be unique to rhodococci. In *C. glutamicum* the gene order is different, and in streptomycetes, the genes appear to be arranged in a single operon. In all bacteria, the *cat* genes are usually organized in a single cluster (reviewed in reference 43). However, permutations occur with respect to gene order (e.g., *catRBAC* in the streptomycetes sequenced to date) and the presence of additional genes in the transcriptional unit (e.g., in *Acinetobacter*, there are six genes in the *cat* operon). The order of the genes in RHA1, *catRABC*, is seen in *R. opacus* 1CP (29), *R. erythropolis* AN-13 (D83237), and *R. erythropolis* CCM2595 (AJ605581). In *C. glutamicum* ATCC 13032, *catR* is not adjacent to the other genes.

There are insufficient data to compare the regulation of the β-ketoadipate pathway genes in rhodococci to that in other organisms. It is nonetheless interesting to note the proteomic data which suggest that *pcaBLRF* are independently regulated to *pcaHG* despite their organization in an apparent operon (Fig. 3c). Consistent with this possibility, the gap between *pcaG* and *pcaB*, 28 bp, is larger than that between the other genes, which typically overlap 1 or 4 bp. This same spacing is seen in the *pcaHGBLRF* cluster of *R. opacus* 1CP (29). Further related to regulation, we note that PcaR, BenR, and PadR of RHA1 all belong to the same family of regulators as the identically named regulators in other bacteria. In contrast, the CatRs of RHA1 and *R. opacus* 1CP (28) belong to the IclR family, not to the LysR family to which the CatRs of other strains often belong (79). Finally, not all β-ketoadipate pathway enzymes that are expressed in RHA1 during growth on benzoate are utilized (e.g., PcaHG and PcaB), as observed in

pseudomonads (37). The physiological relevance of this apparent inefficiency is unclear, particularly given the apparent differential regulation of *pcaHG* and *pcaBLRF*: it is possible that these bacteria only encounter mixtures of compounds degraded via benzoate and protocatechuate in their natural environments.

The emerging data, including those from the current study, suggest that the catabolism of aromatic compounds in rhodococci is organized in a fashion similar to that found in the better-studied pseudomonads (42): a large number of "peripheral" pathways funnel a range of natural and xenobiotic compounds into a restricted number of "central" pathways. The latter, exemplified by the β-ketoadipate pathway, complete the transformation of these compounds to TCA cycle intermediates. Analyses of the genomic sequences of four pseudomonad strains, *P. putida* KT2440, *P. fluorescens* Pf0-1, *P. aeruginosa* PAO1, *and P. syringae* pv. *tomato* DC3000 (42, 61), together with functional studies have identified at least 38 peripheral pathways, some of which are strain specific, and five conserved central pathways. The current genomic and proteomic data suggest that in RHA1, three peripheral pathways funnel phthalate, terephthalate, and 4-hydroxybenzoate to the β-ketoadipate pathway via protocatechuate, while three others funnel benzoate, phenol, and 2-ethoxyphenol to the β-ketoadipate pathway via catechol (M. A. Patrauchan, H. Hara, and L. D. Eltis, unpublished data).

The duplication of the peripheral pathways responsible for phthalate and terephthalate catabolism (*tpa-pat-pad*) establishes that catabolic redundancy in RHA1 is not confined to the *bph*, *etb*, and *ebd* genes, which specify pathways involved in biphenyl and ethylbenzene catabolism (51, 55, 67). While redundancy in catabolic genes has been cited as a trait of *Rhodococcus* (53, 81), the cited examples appear to involve paralogs involved in distinct, nonredundant physiological processes (40, 77). The duplication of catabolic pathways in RHA1 may increase the organism's potential to adapt to new carbon sources. This hypothesis is supported by the growth of the *pcaL* deletion mutant on benzoate: in this case, adaptation apparently involved recruitment of a *catD* ortholog. Nevertheless, it is unclear whether such duplication is shared by divergent rhodococcal strains and whether this redundancy is associated primarily with catabolic genes.

The physiological role of the gene cluster containing the *catD* ortholog, rha05436, remains unclear. Considering the likely enzymatic activities of the encoded proteins (Table 3), as well as the simultaneous expression of a probable catecholic siderophore biosynthetic pathway (Tables 3 and 4), it is tempting to speculate that the cluster containing the *cat*D homolog is involved in the degradation of the catecholic siderophore and that the two gene clusters are under the control of a common regulatory mechanism. Interestingly, the RHA1 genome contains 21 genes whose products share as much or greater sequence identity with the enol-lactone hydrolase domain of PcaL than the rha0543-encoded ortholog does (results not shown). Nevertheless, it is unclear whether all of these enzymes could function in the β-ketoadipate pathway. Moreover, although the *pcaL* mutant did not adapt to phthalate, the RHA1 genome also contains two *pcaC* homologs whose products share 32 and 53% sequence identity with the decarboxylase domain of PcaL. Finally, the growth of the *pcaL* mutant on benzoate illustrates the limitation of assigning gene function based on gene deletion studies.

In conclusion, this study establishes the particular configuration of the β-ketoadipate pathway in RHA1 and indicates that the organization of the genes encoding this pathway is characteristic of rhodococci and related actinomycetes. Moreover, our analyses indicate that the overall organization of the catabolism of aromatic compounds in rhodococci is similar to that described in pseudomonads, with multiple peripheral pathways feeding into a limited number of central pathways. The redundancy of the peripheral phthalate and terephthalate pathways in RHA1 stands in marked contrast to the convergent nature of the β-ketoadipate pathway. It is possible that this configuration augments the ability of RHA1 to adapt to newly encountered carbon sources, much as the *pcaL* mutant adapted to benzoate by recruiting an ortholog of the deleted gene. Finally, several lines of evidence provided by the current study, including the arrangement of the catabolic islands and the expression of *pcaHG* during growth on benzoate, indicate that RHA1 might normally grow on mixtures of aromatic compounds in the environment. We are currently investigating the regulation of these different pathways on single and multiple carbon sources.

ACKNOWLEDGMENTS

This work was supported by grants from Genome Canada and Genome BC.

Masao Fukuda is thanked for helpful discussions. Robert Olafson, Derek Smith, and other members of the Proteomics Centre, University of Victoria, are thanked for their assistance with the mass spectrometry. Ritesh Patel, Matthew J. Myhre, Clinton Fernandez, and Michael McLeod are thanked for their assistance in bioinformatic analyses. Hirofumi Hara is thanked for valuable assistance with the enzyme assays and for sharing microarray data.

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. **Apweiler, R., T. K. Attwood, A. Bairoch, A. Bateman, E. Birney, M. Biswas, P. Bucher, L. Cerutti, F. Corpet, M. D. Croning, R. Durbin, L. Falquet, W. Fleischmann, J. Gouzy, H. Hermjakob, N. Hulo, I. Jonassen, D. Kahn, A. Kanapin, Y. Karavidopoulou, R. Lopez, B. Marx, N. J. Mulder, T. M. Oinn, M. Pagni, F. Servant, C. J. Sigrist, and E. M. Zdobnov.** 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. Nucleic Acids Res. **29:**37–40.
3. **Badger, J. H., and G. J. Olsen.** 1999. CRITICA: coding region identification tool invoking comparative analysis. Mol. Biol. Evol. **16:**512–524.
4. **Barnett, M. J., R. F. Fisher, T. Jones, C. Komp, A. P. Abola, F. Barloy-Hubler, L. Bowser, D. Capela, F. Galibert, J. Gouzy, M. Gurjal, A. Hong, L. Huizar, R. W. Hyman, D. Kahn, M. L. Kahn, S. Kalman, D. H. Keating, C. Palm, M. C. Peck, R. Surzycki, D. H. Wells, K. C. Yeh, R. W. Davis, N. A. Federspiel, and S. R. Long.** 2001. Nucleotide sequence and predicted functions of the entire *Sinorhizobium meliloti* pSymA megaplasmid. Proc. Natl. Acad. Sci. USA **98:**9883–9888.
5. **Batie, C. J., E. LaHaie, and D. P. Ballou.** 1987. Purification and characterization of phthalate oxygenase and phthalate oxygenase reductase from *Pseudomonas cepacia*. J. Biol. Chem. **262:**1510–1518.
6. **Bell, K. S., J. C. Philp, D. W. Aw, and N. Christofi.** 1998. The genus *Rhodococcus*. J. Appl. Microbiol. **85:**195–210.
7. **Bentley, S. D., K. F. Chater, A. M. Cerdeno-Tarraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C. H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M. A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, and D. A. Hopwood.** 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature **417:**141–147.
8. **Bjellqvist, B., G. J. Hughes, C. Pasquali, N. Paquet, F. Ravier, J. C. Sanchez, S. Frutiger, and D. Hochstrasser.** 1993. The focusing positions of polypep-

tides in immobilized pH gradients can be predicted from their amino acid sequences. Electrophoresis **14:**1023–1031.
9. **Bourhy, P., A. Martel, D. Margarita, I. Saint Girons, and J. Belfaiza.** 1997. Homoserine O-acetyltransferase, involved in the *Leptospira meyeri* methionine biosynthetic pathway, is not feedback inhibited. J. Bacteriol. **179:**4396–4398.
10. **Boynton, Z. L., G. N. Bennet, and F. B. Rudolph.** 1996. Cloning, sequencing, and expression of clustered genes encoding beta-hydroxybutyryl-coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-CoA dehydrogenase from *Clostridium acetobutylicum* ATCC 824. J. Bacteriol. **178:**3015–3024.
11. **Brzostowicz, P. C., A. B. Reams, T. J. Clark, and E. L. Neidle.** 2003. Transcriptional cross-regulation of the catechol and protocatechuate branches of the beta-ketoadipate pathway contributes to carbon source-dependent expression of the *Acinetobacter* sp. strain ADP1 *pob*A gene. Appl. Environ. Microbiol. **69:**1598–1606.
12. **Brzostowicz, P. C., D. M. Walters, S. M. Thomas, V. Nagarajan, and P. E. Rouviere.** 2003. mRNA differential display in a microbial enrichment culture: simultaneous identification of three cyclohexanone monooxygenases from three species. Appl. Environ. Microbiol. **69:**334–342.
13. **Chang, H. K., and G. J. Zylstra.** 1999. Characterization of the phthalate permease OphD from *Burkholderia cepacia* ATCC 17616. J. Bacteriol. **181:**6197–6199.
14. **Chang, H. K., and G. J. Zylstra.** 1998. Novel organization of the genes for phthalate degradation from *Burkholderia cepacia* DBO1. J. Bacteriol. **180:**6529–6537.
15. **Chauvaux, S., F. Chevalier, C. Le Dantec, F. Fayolle, I. Miras, F. Kunst, and P. Beguin.** 2001. Cloning of a genetically unstable cytochrome P-450 gene cluster involved in degradation of the pollutant ethyl tert-butyl ether by *Rhodococcus ruber*. J. Bacteriol. **183:**6551–6557.
16. **Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eiglmeier, S. Gas, C. E. Barry, 3rd, F. Tekaia, K. Badcock, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. Davies, K. Devlin, T. Feltwell, S. Gentles, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, B. G. Barrell, et al.** 1998. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature **393:**537–544.
17. **Cowles, C. E., N. N. Nichols, and C. S. Harwood.** 2000. BenR, a XylS homologue, regulates three different pathways of aromatic acid degradation in *Pseudomonas putida*. J. Bacteriol. **182:**6339–6346.
18. **Datsenko, K. A., and B. L. Wanner.** 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc. Natl. Acad. Sci. USA **97:**6640–6645.
19. **Dean, C. R., and J. B. Goldberg.** 2000. The *wbp*M gene in *Pseudomonas aeruginosa* serogroup O17 resides on a cryptic copy of the serogroup O11 O antigen gene locus. FEMS Microbiol. Lett. **187:**59–63.
20. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic. Acids Res. **27:**4636–4641.
21. **De Mot, R., I. Nagy, G. Schoofs, and J. Vanderleyden.** 1994. Sequences of the cobalamin biosynthetic genes *cob*K, cobL and *cob*M from *Rhodococcus* sp. NI86/21. Gene **143:**91–93.
22. **Dorn, E., and H. J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Substituent effects on 1,2-dioxygenation of catechol. Biochem. J. **174:**85–94.
23. **Doten, R. C., K. L. Ngai, D. J. Mitchell, and L. N. Ornston.** 1987. Cloning and genetic organization of the *pca* gene cluster from *Acinetobacter calcoaceticus*. J. Bacteriol. **169:**3168–3174.
24. **Duffner, F. M., and R. Muller.** 1998. A novel phenol hydroxylase and catechol 2,3-dioxygenase from the thermophilic *Bacillus thermoleovorans* strain A2: nucleotide sequence and analysis of the genes. FEMS Microbiol. Lett. **161:**37–45.
25. **Eaton, R. W.** 2001. Plasmid-encoded phthalate catabolic pathway in *Arthrobacter keyseri* 12B. J. Bacteriol. **183:**3689–3703.
26. **Eaton, R. W., and D. W. Ribbons.** 1982. Metabolism of dibutylphthalate and phthalate by *Micrococcus* sp. strain 12B. J. Bacteriol. **151:**48–57.
27. **Eltis, L. D., U. Karlson, and K. N. Timmis.** 1993. Purification and characterization of cytochrome P450RR1 from *Rhodococcus rhodochrous*. Eur. J. Biochem. **213:**211–216.
28. **Eulberg, D., L. A. Golovleva, and M. Schlomann.** 1997. Characterization of catechol catabolic genes from *Rhodococcus erythropolis* 1CP. J. Bacteriol. **179:**370–381.
29. **Eulberg, D., S. Lakner, L. A. Golovleva, and M. Schlomann.** 1998. Characterization of a protocatechuate catabolic gene cluster from *Rhodococcus opacus* 1CP: evidence for a merged enzyme with 4-carboxymuconolactone-decarboxylating and 3-oxoadipate enol-lactone-hydrolyzing activity. J. Bacteriol. **180:**1072–1081.
30. **Furukawa, K., H. Suenaga, and M. Goto.** 2004. Biphenyl dioxygenases: functional versatilities and directed evolution. J. Bacteriol. **186:**5189–5196.
31. **Gorg, A., C. Obermaier, G. Boguth, A. Harder, B. Scheibe, R. Wildgruber, and W. Weiss.** 2000. The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis **21:**1037–1053.
32. **Gorg, A., C. Obermaier, G. Boguth, and W. Weiss.** 1999. Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures. Electrophoresis **20:**712–717.
33. **Grogan, G., G. A. Roberts, D. Bougioukou, N. J. Turner, and S. L. Flitsch.** 2001. The desymmetrization of bicyclic beta-diketones by an enzymatic retro-Claisen reaction. A new reaction of the crotonase superfamily. J. Biol. Chem. **276:**12565–12572.
34. **Gurtler, V., B. C. Mayall, and R. Seviour.** 2004. Can whole genome analysis refine the taxonomy of the genus *Rhodococcus*? FEMS Microbiol. Rev. **28:**377–403.
35. **Gust, B., G. L. Challis, K. Fowler, T. Kieser, and K. F. Chater.** 2003. PCR-targeted *Streptomyces* gene replacement identifies a protein domain needed for biosynthesis of the sesquiterpene soil odor geosmin. Proc. Natl. Acad. Sci. USA **100:**1541–1546.
36. **Harwood, C. S., N. N. Nichols, M. K. Kim, J. L. Ditty, and R. E. Parales.** 1994. Identification of the *pca*RKF gene cluster from *Pseudomonas putida*: involvement in chemotaxis, biodegradation, and transport of 4-hydroxybenzoate. J. Bacteriol. **176:**6479–6488.
37. **Harwood, C. S., and R. E. Parales.** 1996. The beta-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. **50:**553–590.
38. **Herrmann, H., C. Muller, 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.
39. **Huston, W. M., M. P. Jennings, and A. G. McEwan.** 2002. The multicopper oxidase of *Pseudomonas aeruginosa* is a ferroxidase with a central role in iron acquisition. Mol. Microbiol. **45:**1741–1750.
40. **Irvine, V. A., L. A. Kulakov, and M. J. Larkin.** 2000. The diversity of extradiol dioxygenase (*edo*) genes in cresol degrading rhodococci from a creosote-contaminated site that express a wide range of degradative abilities. Antonie van Leeuwenhoek **78:**341–352.
41. **Iwagami, S. G., K. Yang, and J. Davies.** 2000. Characterization of the protocatechuic acid catabolic gene cluster from *Streptomyces* sp. strain 2065. Appl. Environ. Microbiol. **66:**1499–1508.
42. **Jimenez, J. I., and J. L. Garcia.** 2004. Genomic insights in the metabolism of aromatic compounds in *Pseudomonas*, p. 425–461. *In* J.-L. Ramos (ed.), *Pseudomonas*, vol. 3. Kluwer Academic/Plenum Publishers, New York, N.Y.
43. **Jimenez, J. I., B. Minambres, J. L. Garcia, and E. Diaz.** 2002. Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. Environ. Microbiol. **4:**824–841.
44. **Johnson, B. F., and R. Y. Stanier.** 1971. Regulation of the-ketoadipate pathway in *Alcaligenes eutrophus*. J. Bacteriol. **107:**476–485.
45. **Kalinowski, J., B. Bathe, D. Bartels, N. Bischoff, M. Bott, A. Burkovski, N. Dusch, L. Eggeling, B. J. Eikmanns, L. Gaigalat, A. Goesmann, M. Hartmann, K. Huthmacher, R. Kramer, B. Linke, A. C. McHardy, F. Meyer, B. Mockel, W. Pfefferle, A. Puhler, D. A. Rey, C. Ruckert, O. Rupp, H. Sahm, V. F. Wendisch, I. Wiegrabe, and A. Tauch.** 2003. The complete *Corynebacterium glutamicum* ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins. J. Biotechnol. **104:**5–25.
46. **Kaschabek, S. R., B. Kuhn, D. Muller, E. Schmidt, and W. Reineke.** 2002. Degradation of aromatics and chloroaromatics by *Pseudomonas* sp. strain B13: purification and characterization of 3-oxoadipate:succinyl-coenzyme A (CoA) transferase and 3-oxoadipyl-CoA thiolase. J. Bacteriol. **184:**207–215.
47. **Kieser, T., M. J. Bibb, M. J. Buttner, K. F. Chater, and D. A. Hopwood.** 2000. Practical *Streptomyces* genetics. The John Innes Foundation, Norwich, United Kingdom.
48. **Kim, Y. R., S. E. Lee, C. M. Kim, S. Y. Kim, E. K. Shin, D. H. Shin, S. S. Chung, H. E. Choy, A. Progulske-Fox, J. D. Hillman, M. Handfield, and J. H. Rhee.** 2003. Characterization and pathogenic significance of *Vibrio vulnificus* antigens preferentially expressed in septicemic patients. Infect. Immun. **71:**5461–5471.
49. **Kinter, M. T., and N. E. Sherman.** 2000. Protein sequencing and identification using tandem mass spectrometry. John Wiley & Sons Inc., New York, N.Y.
50. **Kitagawa, W., K. Miyauchi, E. Masai, and M. Fukuda.** 2001. Cloning and characterization of benzoate catabolic genes in the gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. J. Bacteriol. **183:**6598–6606.
51. **Kitagawa, W., A. Suzuki, T. Hoaki, E. Masai, and M. Fukuda.** 2001. Multiplicity of aromatic ring hydroxylation dioxygenase genes in a strong PCB degrader, Rhodococcus sp. strain RHA1 demonstrated by denaturing gradient gel electrophoresis. Biosci. Biotechnol. Biochem. **65:**1907–1911.
52. **Kornberg, H. L.** 1966. The role and control of the glyoxylate cycle in *Escherichia coli*. Biochem. J. **99:**1–11.
53. **Kulakov, L. A., and M. J. Larkin.** 2002. Genetic organization of *Rhodococcus*, p. 15–46. *In* A. Danchin (ed.), Genomics of GC-rich gram-positive bacteria. Caister Academic Press, Wymondham, United Kingdom.
54. **Lessard, P. A., X. M. O'Brien, N. A. Ahlgren, S. A. Ribich, and A. J. Sinskey.** 1999. Characterization of IS1676 from *Rhodococcus erythropolis* SQ1. Appl. Microbiol. Biotechnol. **52:**811–819.
55. **Maeda, M., S. Y. Chung, E. Song, and T. Kudo.** 1995. Multiple genes encoding 2,3-dihydroxybiphenyl 1,2-dioxygenase in the gram-positive poly-

chlorinated biphenyl-degrading bacterium *Rhodococcus erythropolis* TA421, isolated from a termite ecosystem. Appl. Environ. Microbiol. **61:**549–555.

56. **Mahenthiralingam, E., B. I. Marklund, L. A. Brooks, D. A. Smith, G. J. Bancroft, and R. W. Stokes.** 1998. Site-directed mutagenesis of the 19-kilodalton lipoprotein antigen reveals No essential role for the protein in the growth and virulence of *Mycobacterium intracellular*. Infect. Immun. **66:** 3626–3634.
57. **Masai, E., A. Yamada, J. M. Healy, T. Hatta, K. Kimbara, M. Fukuda, and K. Yano.** 1995. Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol. **61:**2079–2085.
58. **Morii, S., S. Sawamoto, Y. Yamauchi, M. Miyamoto, M. Iwami, and E. Itagaki.** 1999. Steroid monooxygenase of *Rhodococcus rhodochrous*: sequencing of the genomic DNA, and hyperexpression, purification, and characterization of the recombinant enzyme. J. Biochem. (Tokyo) **126:**624–631.
59. **Nagy, I., G. Schoofs, F. Compernolle, P. Proost, J. Vanderleyden, and R. de Mot.** 1995. Degradation of the thiocarbamate herbicide EPTC (S-ethyl dipropylcarbamothioate) and biosafening by *Rhodococcus* sp. strain NI86/21 involve an inducible cytochrome P-450 system and aldehyde dehydrogenase. J. Bacteriol. **177:**676–687.
60. **Navarro-Llorens, J. M., M. A. Patrauchan, G. R. Stewart, J. E. Davies, L. D. Eltis, and W. W. Mohn.** Phenylacetate catabolism in *Rhodococcus* sp. strain RHA1: a central pathway for degradation of aromatic compounds. J. Bacteriol., in press.
61. **Nichols, N. N., and C. S. Harwood.** 1995. Repression of 4-hydroxybenzoate transport and degradation by benzoate: a new layer of regulatory control in the *Pseudomonas putida* beta-ketoadipate pathway. J. Bacteriol. **177:**7033–7040.
62. **Nierman, W. C., T. V. Feldblyum, M. T. Laub, I. T. Paulsen, K. E. Nelson, J. A. Eisen, J. F. Heidelberg, M. R. Alley, N. Ohta, J. R. Maddock, I. Potocka, W. C. Nelson, A. Newton, C. Stephens, N. D. Phadke, B. Ely, R. T. DeBoy, R. J. Dodson, A. S. Durkin, M. L. Gwinn, D. H. Haft, J. F. Kolonay, J. Smit, M. B. Craven, H. Khouri, J. Shetty, K. Berry, T. Utterback, K. Tran, A. Wolf, J. Vamathevan, M. Ermolaeva, O. White, S. L. Salzberg, J. C. Venter, L. Shapiro, C. M. Fraser, and J. Eisen.** 2001. Complete genome sequence of *Caulobacter crescentus*. Proc. Natl. Acad. Sci. USA **98:**4136–4141.
63. **Omura, S., H. Ikeda, J. Ishikawa, A. Hanamoto, C. Takahashi, M. Shinose, Y. Takahashi, H. Horikawa, H. Nakazawa, T. Osonoe, H. Kikuchi, T. Shiba, Y. Sakaki, and M. Hattori.** 2001. Genome sequence of an industrial microorganism *Streptomyces avermitilis*: deducing the ability of producing secondary metabolites. Proc. Natl. Acad. Sci. USA **98:**12215–12220.
64. **Paget, M. S., L. Chamberlin, A. Atrih, S. J. Foster, and M. J. Buttner.** 1999. Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). J. Bacteriol. **181:**204–211.
65. **Pinyakong, O., H. Habe, T. Yoshida, H. Nojiri, and T. Omori.** 2003. Identification of three novel salicylate 1-hydroxylases involved in the phenanthrene degradation of *Sphingobium* sp. strain P2. Biochem. Biophys. Res. Commun. **301:**350–357.
66. **Roberts, G. A., G. Grogan, A. Greter, S. L. Flitsch, and N. J. Turner.** 2002. Identification of a new class of cytochrome P450 from a *Rhodococcus* sp. J. Bacteriol. **184:**3898–3908.
67. **Sakai, M., E. Masai, H. Asami, K. Sugiyama, K. Kimbara, and M. Fukuda.** 2002. Diversity of 2,3-dihydroxybiphenyl dioxygenase genes in a strong PCB degrader, Rhodococcus sp strain RHA1. J. Biosci. Bioeng. **93:**421–427.
68. **Sakai, M., K. Miyauchi, N. Kato, E. Masai, and M. Fukuda.** 2003. 2-Hydroxypenta-2,4-dienoate metabolic pathway genes in a strong polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol. **69:**427–433.

68a. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning: a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

69. **Seto, M., K. Kimbara, M. Shimura, T. Hatta, M. Fukuda, and K. Yano.** 1995. A novel transformation of polychlorinated biphenyls by *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol. **61:**3353–3358.
70. **Shanley, M. S., A. Harrison, R. E. Parales, G. Kowalchuk, D. J. Mitchell, and L. N. Ornston.** 1994. Unusual G + C content and codon usage in catIJF, a segment of the ben-cat supra-operonic cluster in the *Acinetobacter calcoaceticus* chromosome. Gene **138:**59–65.
71. **Shigematsu, T., K. Yumihara, Y. Ueda, S. Morimura, and K. Kida.** 2003. Purification and gene cloning of the oxygenase component of the terephthalate 1,2-dioxygenase system from *Delftia tsuruhatensis* strain T7. FEMS Microbiol. Lett. **220:**255–260.
72. **Shimizu, S., H. Kobayashi, E. Masai, and M. Fukuda.** 2001. Characterization of the 450-kb linear plasmid in a polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. Appl. Environ. Microbiol. **67:**2021–2028.
73. **Skurnik, M., A. Peippo, and E. Ervela.** 2000. Characterization of the O-antigen gene clusters of Yersinia pseudotuberculosis and the cryptic O-antigen gene cluster of *Yersinia pestis* shows that the plague bacillus is most closely related to and has evolved from *Y. pseudotuberculosis* serotype O:1b. Mol. Microbiol. **37:**316–330.
74. **Steele, M. I., D. Lorenz, K. Hatter, A. Park, and J. R. Sokatch.** 1992. Characterization of the *mmsAB* operon of *Pseudomonas aeruginosa* PAO encoding methylmalonate-semialdehyde dehydrogenase and 3-hydroxyisobutyrate dehydrogenase. J. Biol. Chem. **267:**13585–13592.
75. **Stim-Herndon, K. P., D. J. Petersen, and G. N. Bennett.** 1995. Characterization of an acetyl-CoA C-acetyltransferase (thiolase) gene from *Clostridium acetobutylicum* ATCC 824. Gene **154:**81–85.
76. **Streit, W. R., R. A. Schmitz, X. Perret, C. Staehelin, W. J. Deakin, C. Raasch, H. Liesegang, and W. J. Broughton.** 2004. An evolutionary hot spot: the pNGR234b replicon of *Rhizobium* sp. strain NGR234. J. Bacteriol. **186:**535–542.
77. **Taguchi, K., M. Motoyama, and T. Kudo.** 2004. Multiplicity of 2,3-dihydroxybiphenyl dioxygenase genes in the Gram-positive polychlorinated biphenyl degrading bacterium *Rhodococcus rhodochrous* K37. Biosci. Biotechnol. Biochem. **68:**787–795.
78. **Takeda, H., A. Yamada, K. Miyauchi, E. Masai, and M. Fukuda.** 2004. Characterization of transcriptional regulatory genes for biphenyl degradation in *Rhodococcus* sp. strain RHA1. J. Bacteriol. **186:**2134–2146.
79. **Tropel, D., and J. R. Van Der Meer.** 2004. Bacterial transcriptional regulators for degradation pathways of aromatic compounds. Microbiol. Mol. Biol. Rev. **68:**474–500.
80. **Trott, S., S. Burger, C. Calaminus, and A. Stolz.** 2002. Cloning and heterologous expression of an enantioselective amidase from *Rhodococcus erythropolis* strain MP50. Appl. Environ. Microbiol. **68:**3279–3286.
81. **van der Geize, R., and L. Dijkhuizen.** 2004. Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr. Opin. Microbiol. **7:**255–261.
82. **van der Geize, R., G. I. Hessels, R. van Gerwen, P. van der Meijden, and L. Dijkhuizen.** 2001. Unmarked gene deletion mutagenesis of *kst*D, encoding 3-ketosteroid Delta1-dehydrogenase, in *Rhodococcus erythropolis* SQ1 using *sac*B as counter-selectable marker. FEMS Microbiol. Lett. **205:**197–202.
83. **van Hylckama Vlieg, J. E., and D. B. Janssen.** 2001. Formation and detoxification of reactive intermediates in the metabolism of chlorinated ethenes. J. Biotechnol. **85:**81–102.
84. **Warren, R., W. W. L. Hsiao, H. Kudo, M. Myhre, M. Dosanjh, A. Petrescu, H. Kobayashi, S. Shimizu, K. Miyauchi, E. Masai, G. Yang, J. M. Stott, J. E. Schein, H. Shin, J. Khattra, D. Smailus, Y. S. Butterfield, A. Siddiqui, R. A. Holt, M. A. Marra, S. J. M. Jones, W. W. Mohn, F. S. Brinkman, M. Fukuda, J. Davies, and L. D. Eltis.** 2004. Functional characterization of a catabolic plasmid from polychlorinated biphenyl-degrading *Rhodococcus* sp. RHA1. J. Bacteriol. **186:**7783–7795.