Proc. Natl. Acad. Sci. USA
Vol. 94, pp. 6484–6489, June 1997
Microbiology

A cluster of bacterial genes for anaerobic benzene ring biodegradation

PAUL G. EGLAND*, DALE A. PELLETIER*, MARILYN DISPENSA*, JANE GIBSON†, AND CAROLINE S. HARWOOD*‡

*Department of Microbiology, and Center for Biocatalysis and Bioprocessing, University of Iowa, Iowa City, IA 52242; and †Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853

Communicated by Ralph S. Wolfe, University of Illinois, Urbana, IL, April 14, 1997 (received for review December 31, 1996)

ABSTRACT **A reductive benzoate pathway is the central conduit for the anaerobic biodegradation of aromatic pollutants and lignin monomers. Benzene ring reduction requires a large input of energy and this metabolic capability has, so far, been reported only in bacteria. To determine the molecular basis for this environmentally important process, we cloned and analyzed genes required for the anaerobic degradation of benzoate and related compounds from the phototrophic bacterium, *Rhodopseudomonas palustris*. A cluster of 24 genes was identified that includes twelve genes likely to be involved in anaerobic benzoate degradation and additional genes that convert the related compounds 4-hydroxybenzoate and cyclohexanecarboxylate to benzoyl-CoA. Genes encoding benzoyl-CoA reductase, a novel enzyme able to overcome the resonance stability of the aromatic ring, were identified by directed mutagenesis. The gene encoding the ring-cleavage enzyme, 2-ketocyclohexanecarboxyl-CoA hydrolase, was identified by assaying the enzymatic activity of the protein expressed in *Escherichia coli*. Physiological data and DNA sequence analyses indicate that the benzoate pathway consists of unusual enzymes for ring reduction and cleavage interposed among enzymes homologous to those catalyzing fatty acid degradation. The cloned genes should be useful as probes to identify benzoate degradation genes from other metabolically distinct groups of anaerobic bacteria, such as denitrifying bacteria and sulfate-reducing bacteria.**

The anaerobic degradation of aromatic rings is an important metabolic process carried out by bacteria that does not appear to be catalyzed by eukaryotes. It is important for environmental bioremediation because substantial amounts of toxic aromatic hydrocarbons, produced industrially as solvents and constituents of gasoline, make their way into anoxic groundwaters and sediments. It is also critical to the recycling of biomass on a global scale because aromatic compounds are the monomeric constituents of lignin, a major plant polymer. Although the biodegradation of aromatic rings under anaerobic conditions has great practical significance, remarkably little is known about the biochemical and molecular details of this process.

As a general rule, anaerobic bacteria metabolize structurally diverse aromatic compounds including aromatic hydrocarbons, phenols, halogenated aromatics, and phenylpropanoids, to generate benzoate, usually as its CoA thioester. The subsequent steps of the benzoate pathway make up the central route used for aromatic ring reduction and cleavage by anaerobes (1–4) (Fig. 1). The high resonance energy of the benzene ring makes controlled chemical reduction difficult to achieve in the laboratory and this is also true in biological systems. In fact, the reductive aspect of anaerobic aromatic degradation is distinctive, and radically different from the well-studied oxidative pathways employed by microbes to degrade aromatic compounds in air (15). The anaerobic benzoate pathway has not been defined in detail, but is proposed to include at least six enzymes, just three of which have been purified and partially characterized (7–10). These studies have provided important, but incomplete, information about key features of the pathway. For example, the reaction catalyzed by benzoyl-CoA reductase, an iron-sulfur flavoprotein recently purified from the denitrifying species *Thauera aromatica*, requires a stoichiometric consumption of ATP (9). This is a feature shared by nitrogenase, a uniquely prokaryotic enzyme that also catalyzes a challenging reduction reaction. To develop a comprehensive view of anaerobic aromatic compound metabolism, we cloned and analyzed a fragment of DNA from the phototrophic bacterium *Rhodopseudomonas palustris* that appears to include all the genes required for anaerobic benzoate and 4-hydroxybenzoate (4-OHBen) degradation, and at least one gene specific for cyclohexanecarboxylate degradation.

MATERIALS AND METHODS

Bacterial Strains and Culture Conditions. *R. palustris* cultures were grown anaerobically in light, and aerobically with shaking, in defined mineral medium at 30°C as described (16). *Escherichia coli* strains were grown in Luria broth at 37°C. Antibiotic concentrations were as described (17). The cyclohexadienecarboxylates were prepared as described (18). The 2,5- and 1,4-cyclohexadienecarboxylates were gifts from Katharine Gibson (DuPont). Compounds were tested as growth substrates at final concentrations of 3 mM.

Cloning, DNA Manipulations, and Mutant Construction. Standard protocols were used for DNA cloning and transformation and for plasmid DNA purification (19). DNA fragments were purified from agarose gels by using the GENECLEAN spin kit from Bio 101. Most recombinant plasmids were generated in pUC vectors and were maintained in *E. coli* DH5α. Chromosomal DNA, purified as described (20), was partially digested with *Sau*3A1, dephosphorylated with alkaline phosphatase, and ligated into the unique *Bam*HI site of cosmid pHC79 (GIBCO/BRL). The ligation mix was packaged into lambda phage using MaxPlax packaging extract from Epicentre Technologies (Madison, WI) and infected into *E. coli* strain JM109 (21) as described by the MaxPlax protocol. Cosmid clones pPE302 and pPE304 were identified as described in *Results*. An *R. palustris* mutant with a *'lacZ*-kanamycin resistance (Km^r) cassette insertion in *badE* was constructed by inserting the 5.2-kb promoterless *lacZ*-Km^r cassette derived from pUTminiTn*5lacZ* (22) into the unique *Bgl*II site present on a *Sac*I fragment that was subcloned from

Abbreviations: 4-OHBen, 4-hydroxybenzoate; Km^r, kanamycin resistance; Gm^r, gentamycin resistance.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. L42322, U02033, U65440, U75363, and U75364).

‡To whom reprint requests should be addressed. e-mail: caroline-harwood@uiowa.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

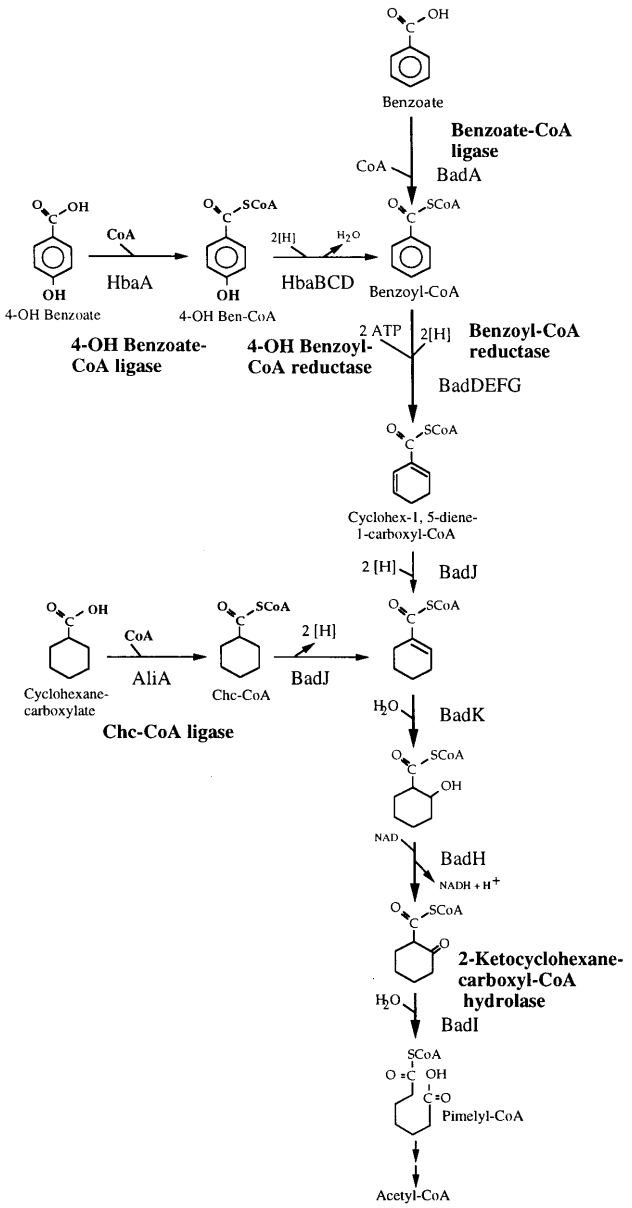

FIG. 1. The anaerobic benzoate degradation pathway and reactions involved in the anaerobic conversion of cyclohexanecarboxylate and 4-hydroxybenzoate to intermediates of benzoate degradation. A similar sequence of reactions, but with intermediates that have slightly different structures has been proposed by Koch *et al.* (5) and is not shown. The product of benzoyl-CoA reduction is uncertain, but is probably one or more of three cyclohexadienecarboxyl-CoA isomers (5, 6), any of which could be subsequently reduced as shown. For simplicity, only one of these possible isomers (cyclohex-1, 5-diene-1-carboxyl-CoA) is shown. The names of enzymes that have been purified from either *R. palustris* or a denitrifying species (*Azoarcus evansii* or *Thauera aromatica*) (7–14) are shown. Assignments of products from genes shown in Fig. 2 to specific steps are indicated.

pPE304 and that encompassed the 3′ half of *badD*, all of *badE* and the 5′ half of *badF*. A *badF*:: Km^r mutant was constructed by inserting the 1.3-kb Km^r GenBlock cassette (Pharmacia) into the unique *Msc*I site on the same subclone. Constructions were cloned into the suicide vector, pJQ200KS (23) and then introduced into *R. palustris* by conjugation from *E. coli* S17–1. Recombinants were selected as described (20, 24). The *badCDEFGA*::Gm^r*badB* (Gm^r, gentamycin resistance) genes were subcloned on an 11.4-kb *Xma*I fragment from the cosmid pPE304 into the broad-host-range vector pBBR1MCS (25) for use in complementation experiments.

DNA Sequencing and Computer Analysis. DNA sequences were determined with an Applied Biosystems model 373A stretch fluorescent automated sequencer at the University of Iowa DNA core facility. Custom synthesized primers were used. DNA templates included overlapping cosmid clones pPE302 and pPE304 and subclones derived from these cosmids. The DNA sequence was analyzed using GENE INSPECTOR, 1.0.1 (Textco, West Lebanon, NH). Similar sequences were identified from SWISS-PROT 26, and GENPEP 78.0, databases using the BLAST network service at the National Center for Biotechnology Information (Bethesda, MD). The GAP program from the Genetics Computer Group (Madison, WI) software package, version 7.0, was used to make sequence comparisons and alignments.

Expression of BadI and BadK in *E. coli*. The BadI protein was expressed from the plasmid pDP105, which contained a 1.4-kb fragment of the cosmid clone pPE304 in pUC18. The subcloned DNA included the 782-bp *badI* gene, as well as the last 230-bp of *badH* and the first 373 bp of *badJ*. BadK was expressed from the plasmid pDP111, which consisted of a 1.3-kb fragment of pPE304 in pUC18. The subcloned DNA included the 681-bp *badK* gene, the last 235 bp of *aliA*, and 287 bp of DNA downstream of *badK*. Expression of the proteins was driven from the *lac* promoter of pUC18. Expression was induced with 0.1 mM isopropyl β-D-thiogalactoside when cells reached the mid-logarithmic phase of growth.

Enzyme Assays. Cell extracts were prepared and 2-ketocyclohexanecarboxyl-CoA hydrolase activity determined by measuring the disappearance of a magnesium-enolate complex of the substrate as previously described (26). Cyclohex-1-enecarboxyl-CoA hydratase activity was assayed by measuring the conversion of cyclohex-1-enecarboxyl-CoA to 2-hydroxycyclohexanecarboxyl-CoA. The reaction mixture contained 40 mM Tris buffer (pH 7.5), 2 mM dithiothreitol, and 1 mM cyclohex-1-enecarboxyl-CoA. Reactions were initiated by the addition of cell extract and stopped by addition of 100 mM (final concentration) perchloric acid. The reaction product was separated and collected by HPLC (Waters model 501) using a Ultrasphere ODS-C_{18} reverse-phase (4.6 mm × 25 cm) column (Beckman). The solvent system used was 20 mM ammonium acetate (pH 6.0) and methanol as solvents A and B, respectively. The column was equilibrated at 20% solvent B and elution was by a linear gradient of 20–80% B in 30 min. The effluent was monitored by absorbance at 210–260 nm using a photodiode array detector (Waters). Peaks were collected, lyophilized, and resuspended in Milli-Q water. The collected product had a molecular mass that matched that of 2-hydroxycyclohexanecarboxyl-CoA as determined by electrospray mass spectrometry at the University of Iowa High Resolution Mass Spectrometry Research Facility.

β-Galactosidase activity was measured by a variation of the method of Miller (27). Cells in the logarithmic phase of growth were harvested, washed in Z buffer and sonicated. Cell extract and Z buffer were combined to a volume of 1 ml and 0.2 ml of a 4 mg/ml solution of *O*-nitrophenylgalactopyranoside was added to start the reactions. The rate of increase in absorbance at 420 nm due to *O*-nitrophenol formation was measured spectrophotometrically. Activity was calculated using millimolar extinction coefficient of 4.5 for *O*-nitrophenol at 420 nm.

Protein was determined using the Bio-Rad protein assay kit.

Other Procedures. Primer extension analysis was carried out using the AMV reverse transcriptase primer extension system according to the protocol supplied by the manufacturer (Promega). The access reverse transcription–PCR system (Promega) was used to determine the transcriptional organization of the *badDEF* and *G* genes. With this technique, reverse transcriptase and an appropriate primer are used to synthesize cDNA from a region of interest. Polymerase chain reactions with primers flanking the intergenic regions of the *badDEF* and *G* genes were then used to amplify intergenic DNA from

the cDNA. Negative control reactions were done in which reverse transcriptase was omitted to ensure that the DNA products resulted from amplification of cDNA. RNA was prepared using Tri-Reagent from Molecular Research Center (Cincinnati).

Cyclohexanecarboxylate-CoA ligase was purified as described (11) and blotted to a polyvinylidene difluoride membrane (Immobilon P; Millipore). Its N-terminal sequence was determined by the Cornell Biotechnology Analytical/Synthesis Facility (Ithaca, NY) using a Waters Pico-Tag system. Immunoblot analysis of cell extracts with benzoate-CoA ligase antiserum was performed as described (10). 2-Ketocyclohexanecarboxyl-CoA and cyclohex-1-enecarboxyl-CoA were synthesized as described (26).

RESULTS AND DISCUSSION

A cosmid clone bank was constructed with chromosomal DNA from an *R. palustris* mutant (CGA604) that had antibiotic cassette disruptions in genes (*badA* and *hbaA*) encoding two CoA ligases; one for initiating benzoate degradation and one for initiating 4-OHBen degradation (12, 20). *E. coli* cells infected with packaged cosmid DNA were screened for growth in the presence of Gm or Km, or both, to identify clones that included the *badA*::Gm^r or *hbaA*::Km^r regions of the chromosome. Two overlapping cosmid clones (pPE302 and pPE304) were identified that included the 25.4 kb of DNA shown in Fig. 2. The nucleotide sequences of DNA subcloned from these cosmids were determined. The sequenced DNA includes 24 predicted genes that encode at least 13 proteins, some of which have multiple polypeptide components. The genes encoding the two enzymes required for the conversion of 4-OHBen to benzoyl-CoA (*hbaABCD*), and the benzoate-CoA ligase gene (*badA*) have been described (12, 17, 20). Here, we have used physiological observations and sequence information to ascribe functions to many of the other genes in the cluster (Figs. 1 and 2).

The genes *badDEF* and *G* are transcribed as an operon, the expression of which is induced by benzoate. A transcription start site was determined by primer extension analysis to be present 71 bases upstream of the predicted translation initiation codon of *badD*. Reverse transcription–PCR amplification of the intergenic regions between *badD* and *E*, *badE* and *F*, and *badF* and *G*, indicated that the four genes are cotranscribed. Measurement of β-galactosidase activity in a *badE::'lacZ* strain showed that expression of these genes is induced fourfold by anaerobic growth in the presence of benzoate (Fig. 3).

Mutants with disruptions in *badE* or *badF* were constructed by homologous recombination between the wild-type *R. palustris* chromosome and plasmids containing cloned DNA interrupted with a promoterless *lacZ*-Km^r cassette or a Km^r cassette, respectively. Neither mutant was able to grow anaerobically with benzoate or 4-OHBen as a sole carbon source, but both grew at wild-type rates on cyclohexa-1,5-, -1,4-, and -2,5-diene-1-carboxylates, the free acid derivatives of three compounds that have been proposed as the first two-equivalent reduction products of benzoyl-CoA reduction (5, 6). The *badE* and *badF* mutants accumulated radiolabeled benzoyl-CoA intracellularly when whole cells were provided with ^{14}C benzoate (Fig. 4). The benzoate growth defect of the mutants is not due to an effect of the mutations on genes downstream of *badG* because immunoblot analysis of cell extracts prepared from the *badE* and *badF* strains showed that the product of *badA*, the gene immediately downstream of *badG*, is expressed. We were also able to complement the benzoate growth defect in trans by introducing a clone containing *badDEFGA*::Gm^r*badB* into the *badE* and *badF* mutant strains.

These data suggest that *badDEFG* encode the enzyme benzoyl-CoA reductase (Fig. 1). In support of this, the sizes of the predicted protein products of *badDEFG* (44.0, 47.6, 47.8, and 29.6 kDa) are similar to those of the four subunits of the benzoyl-CoA reductase purified from *T. aromatica* (38, 45, 48, and 32 kDa) (9). Moreover, the experimentally determined N-terminal amino acid sequences of the benzoyl-CoA reductase subunits from *T. aromatica* are similar to the deduced N-terminal amino acid sequences of the corresponding Bad proteins from *R. palustris*. When the sequences of the first 30 N-terminal amino acids of the α, β, γ, and δ subunits from the *T. aromatica* benzoyl-CoA reductase were compared with those of BadF, BadE, BadD and BadG, respectively, from *R. palustris*, amino acid identities of 76%, 33%, 78%, and 68%

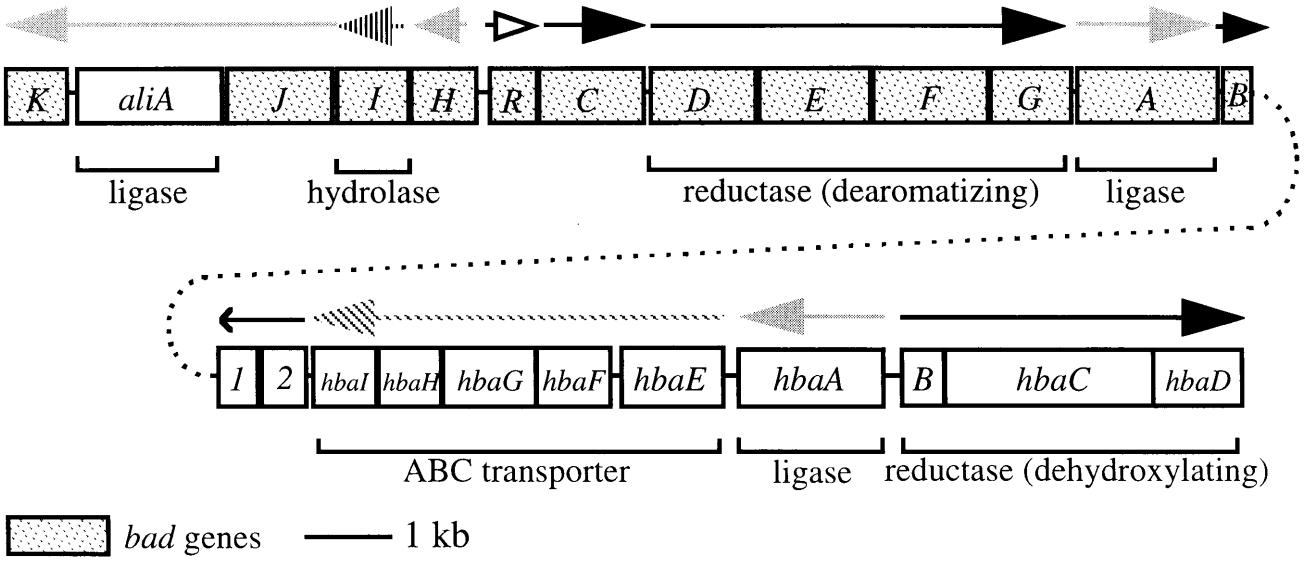

FIG. 2. Organization of *bad* (benzoic acid degradation), *ali* (alicyclic acid degradation), and *hba* (4-hydroxybenzoic acid degradation) genes from *R. palustris*. All genes are contiguous on the chromosome. ORF 1 lies directly downstream of *badB*. Horizontal arrows indicate direction of transcription, but not necessarily transcriptional units. With the exception of the *badDEFG* operon, these have not yet been determined. No function has been proposed for the ORFs designated 1 and 2. Solid gray arrows indicate genes encoding proteins proposed to carry out reactions homologous to those of fatty acid degradation. Large black arrows indicate genes encoding proteins proposed to participate in reduction reactions involving transfer of low potential electrons. The hydrolase gene is indicated by vertical hatch marks, ABC transporter genes by slanted hatch marks and a putative regulatory gene by an open arrow.

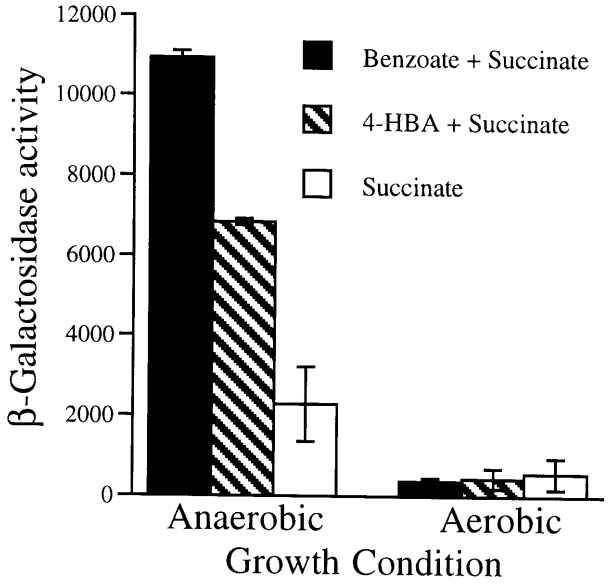

FIG. 3. Expression of β-galactosidase by a *badE*::*'lacZ* fusion strain of *R. palustris* grown under different conditions. Activities are expressed as nmoles of product formed per min per mg protein. All cultures were grown on 10 mM succinate in the presence or absence of 3 mM benzoate or 3 mM 4-OHBen. Values are the averages of data obtained from three or more separately prepared cell extracts.

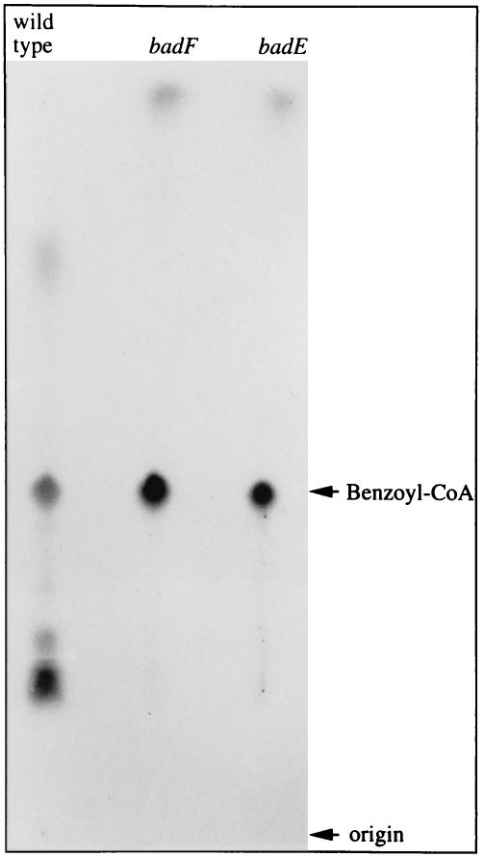

FIG. 4. Autoradiogram comparing the intracellular accumulation of radioactive intermediates by whole cells of wild-type *R. palustris* and the *badE* and *badF* mutants during a 2 min anaerobic incubation with 7-^{14}C benzoate (50 μM). Cells were grown anaerobically with 3 mM cinnamate as carbon source. Following anaerobic incubation with labeled benzoate, cells were suspended in boiling water to release soluble intracellular contents, which were concentrated and then separated by thin layer chromatography before exposure to x-ray film (30). 7-^{14}C-benzoyl-CoA was identified by co-migration with an unlabeled standard purchased from Sigma. Cyclohex-1,4-dienecarboxyl-CoA, prepared enzymatically with cyclohexanecarboxylate-CoA ligase (11) had a lower R_f than benzoyl-CoA (26).

were seen in regions of overlap that included 17, 15, 18, and 25 amino acid residues for the four pairwise comparisons (G. Fuchs, personal communication). This level of identity suggests that there may be a high degree of conservation between the benzoyl-CoA reductases of *T. aromatica*, a member of the β subdivision of the proteobacteria (28), and *R. palustris*, which is a member of the α subdivision.

The deduced amino acid sequences of the BadD, E, F, and G proteins are notable for their lack of similarity to all characterized proteins in genome databases except one: 2-hydroxyglutaryl-CoA dehydratase, an enzyme responsible for the reversible dehydration of (R)-2-hydroxyglutaryl-CoA to glutaconyl-CoA in the pathway of glutamate fermentation by the bacterium *Acidaminococcus fermentans* (29, 31). The reaction catalyzed is unusual because it involves removal of hydrogen atoms from an unactivated carbon atom. Buckel and Keese (32, 33) have proposed that this is accomplished by a one-electron oxidation/reduction by a novel reaction mechanism in which a ketyl radical anion is the first reactive intermediate. Buckel and Keese have noted that a similar mechanism could, in principle, also account for benzoyl-CoA reduction by two successive one-electron transfers (33). 2-Hydroxyglutaryl-CoA dehydratase is comprised of two catalytic subunits, HgdA and HgdB, which share amino acid sequence similarity with BadE and BadD, respectively, and HgdC, an activator subunit that is similar in sequence to both BadF and BadG (Fig. 5). Although the homologous pairs of proteins align along their entire lengths, neither shows any large stretches of identity. The BadD and BadE subunits are each predicted to contain several cysteine residues positioned near aromatic amino acids, some of which align with cysteines in the HgdAB sequences. These may be involved in coordinating the iron-sulfur centers that have been observed in purified 2-hydroxyglutaryl-CoA dehydratase and in benzoyl-CoA reductase.

BadF and BadG are similar in amino acid sequence to each other, as well as to HgdC (Fig. 5). Only catalytic amounts of HgdC and ATP are needed to energize 2-hydroxyglutaryl-CoA dehydration. Presumably, ATP hydrolysis is required to generate a low potential electron donor that, once made, can function through many catalytic cycles of oxidation and reduction (32). The observation that the *T. aromatica* benzoyl-CoA reductase is composed of equimolar amounts of each of its four protein components, is consistent with the finding that stoichiometric, rather than catalytic, amounts of ATP are required to catalyze the reduction reaction. The hydrolysis of ATP that, by analogy with the *T. aromatica* enzyme, is expected to occur during benzoyl-CoA reduction by *R. palustris*, is probably catalyzed by the two subunits BadF and BadG. When considered as a unit, the predicted amino acid sequence of these polypeptides have an ATPase domain as defined by a pattern of amino acid properties (34). This includes two phosphate binding motifs (one present on each polypeptide) and an adenosine binding domain (on BadF). The energy of ATP hydrolysis may serve to generate an electron donor of sufficiently low potential to overcome the high energy barrier to ring reduction.

Neighboring genes, *badC* and *badB*, may encode proteins involved in donating electrons to benzoyl-CoA reductase. The *badC* sequence predicts a protein with ≈30% amino acid sequence identity to NADPH:quinone oxidoreductases and the deduced amino acid sequence of BadB has a high degree of amino acid identity (≈50%) to clostridial-type ferredoxins. The natural electron donor for benzoyl-CoA reductase has not been reported, but ferredoxins are strong candidates because of their capacity to transfer very low potential electrons.

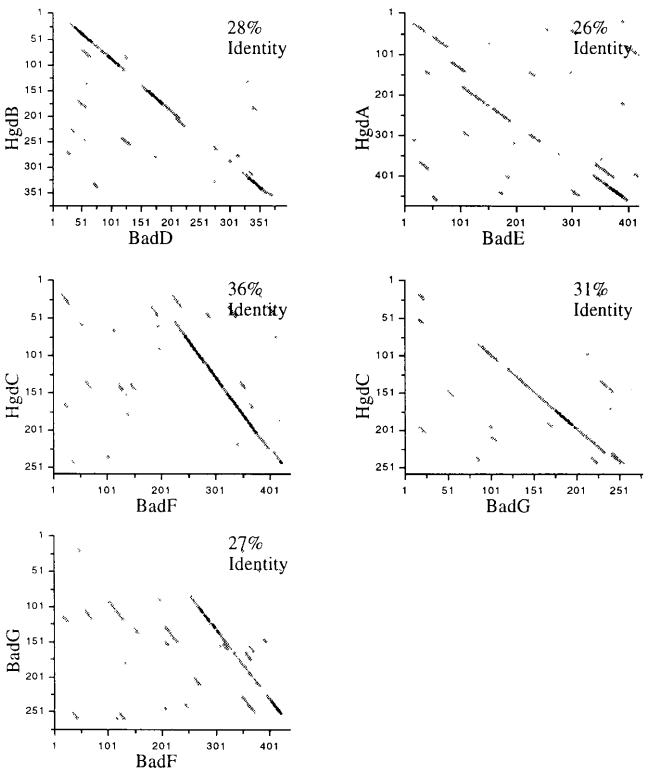

FIG. 5. Dot-matrix analysis. Comparisons were made between the predicted amino acid sequences of benzoyl-CoA reductase subunits (BadDEFG) and the homologous sequences from *A. fermentans* 2-hydroxyglutaryl-CoA dehydratase (HgdABC), and between BadF and BadG. The sequences compared in each graph are printed on the axes, with axes values corresponding to residue positions for each sequence. Diagonal lines indicate regions of eight or more matches over a 35-amino acid window. The analysis was done using the dot matrix program from GENE INSPECTOR, 1.0.1 (Textco). The lengths of HgdA, -B, and -C are 476, 378, and 260 amino acids respectively, and those of BadD, -E, -F, and -G are 394, 436, 437, and 277 amino acids. The overall level of identity for a pair, based on GAP analysis, is displayed in the upper right corner of each graph.

The three reactions in the benzoate pathway that convert cyclohexadienecarboxyl-CoA to 2-ketocyclohexanecarboxyl-CoA, appear to correspond to the three successive oxidation reactions that typify the classical cycle of fatty acid β-oxidation as studied in bacteria and mammals (35) (Fig. 1). Consistent with this, the *badJ* and *badK* gene products have predicted amino acid sequences that are, respectively, ≈35% identical to flavin-containing dehydrogenases and enoyl-CoA hydratases from mammalian mitochondria and bacteria. The *badH* product has the signature sequence motifs of a short-chain alcohol dehydrogenase (36) and shares 30–40% amino acid sequence identity with NADH-dependent dehydrogenases involved in the biosynthesis of bacterial fatty acids and poly-β-hydroxyalkanoates. Extracts of *E. coli* cells expressing *badK* catalyzed the hydration of cyclohex-1-enecarboxyl-CoA to 2-hydroxycyclohexanecarboxyl-CoA, confirming the predicted enzymatic function of this gene product as an enoyl-CoA hydratase. BadH and BadJ have tentatively been assigned the enzymatic functions shown in Fig. 1 based on their amino acid sequence similarities to fatty acid β-oxidation enzymes.

CoA ligases are hallmark enzymes of fatty acid degradation, and three of the genes in the cluster, *badA*, *hbaA*, and *aliA*, encode ligases that catalyze the initial reactions of anaerobic benzoate, 4-OHBen, and cyclohexanecarboxylate degradation, respectively. The *badA* and *hbaA* genes were originally identified by immunoscreening with antisera specific to the purified proteins (12, 20). We assigned a function to *aliA* by determining the sequence of the first 14 N-terminal amino acids of cyclohexanecarboxylate-CoA ligase (11) and finding that it matched the deduced amino acid sequence of the N terminus of the gene product exactly. Each of the aromatic and alicyclic acid CoA ligases is 20–30% identical at the amino acid level to ligases from both prokaryotic and eukaryotic sources that activate fatty acids with CoA.

The ring opening step of anaerobic benzoate degradation involves a hydrolytic cleavage, rather than the thiolytic cleavage that is typical of fatty acid degradation. In *R. palustris* this is accomplished by a previously unknown enzyme, 2-ketocyclohexanecarboxyl-CoA hydrolase (10, 26). We identified the gene encoding this enzyme, *badI*, by showing that the encoded protein catalyzed the cleavage of 2-ketocyclohexanecarboxyl-CoA when expressed in *E. coli* from pUC18. The specific activities measured in *E. coli* cell extracts were ≈five-fold higher than those measured in extracts of benzoate-grown *R. palustris* cells (26). The BadI amino acid sequence has 32% amino acid identity to BadK (Fig. 1), proposed to function as cyclohex-1-ene-1-carboxyl-CoA hydratase, and it is ≈45% identical to dihydroxynaphthoate synthase from *Bacillus subtilis* and *E. coli*, a ring closure enzyme involved in the biosynthesis of menaquinone, a low potential H carrier.

In addition to structural genes, the benzoate degradation gene cluster includes a possible regulatory gene, *badR*. The predicted BadR protein has 28% amino acid identity to an *E. coli* protein termed HpcR (37), that regulates expression of genes for aerobic homoprotocatechuate degradation. Neither BadR or HpcR show any striking amino acid similarities to other characterized regulators. A cluster of five genes downstream of *hbaA* are predicted to encode a binding protein-dependent permease belonging to the general category of ABC transporters (38). Although it is tempting to speculate that such a permease may catalyze the uptake of benzoate or 4-OHBen into cells, this has not yet been shown. Two ORFs (designated 1 and 2) whose predicted products do not resemble characterized proteins in sequence databases lie downstream of this gene cluster. A regulatory protein of the Crp/Fnr class that is essential for anaerobic growth of *R. palustris* on 4-OHBen, and that is required for optimal growth on benzoate, has been described (39). The location of the gene encoding this protein, called *aadR*, for anaerobic aromatic degradation regulator, relative to the *bad-hba* gene cluster, is not known, but this is the only known relevant gene located outside the cluster.

Nucleotide sequencing indicates that genes involved in the biosynthesis of flagella lie immediately downstream of *badK*. We cannot exclude the possibility of additional genes for aromatic compound degradation at the other end of the cluster, downstream of *hbaD*. However, the structural genes described here can account for all the reactions known or hypothesized to be required for the conversion of benzoate and 4-OHBen to the ring cleavage product, pimelyl-CoA. This includes reactions proposed by Koch *et al.* (5) that differ somewhat from those shown in Fig. 1 and that involve conversion of the product of benzoyl-CoA reduction, cyclohexadienecarboxyl-CoA, to a dihydroxylated intermediate (2,6-dihydroxycyclohexanecarboxyl-CoA) by two hydration steps, followed by a dehydrogenation to generate 2-keto-6-hydroxycylohexanecarboxyl-CoA. Cleavage of this substrate would then generate 3-hydroxypimelyl-CoA rather than pimelyl-CoA. It is possible that the physiological role of the BadK, BadH, and BadI proteins is to catalyze this alternative series of reactions. This would require that these enzymes be active with substrates that have slightly different structures than those with which they have been shown to be active so far.

CONCLUSIONS

The results described here suggest that the anaerobic benzoate degradation pathway consists of enzymes specifying the com-

mon metabolic theme of fatty acid degradation, mixed with a reductive enzyme that appears to have very restricted metabolic application, and an enzyme (2-ketocyclohexanecarboxyl-CoA hydrolase) that may have been recruited from a biosynthetic sequence. It is also noteworthy that genes for the degradation of benzoate and 4-OHBen, two compounds that are key aromatic intermediates formed during the anaerobic degradation of many structurally diverse aromatic compounds (3, 4), lie adjacent to each other. The nucleotide sequences of the genes shown in Fig. 2 have been useful for posing hypotheses about gene function that can be tested by genetic and biochemical means. Here, we have presented experimental evidence in support of gene assignments for four of the six enzymes of the benzoate pathway and for one of the two enzymes of cyclohexanecarboxylate degradation. Evidence for the genes encoding the two enzymes specific to 4-OHBen degradation has been presented previously (12, 17). These results provide a framework for additional studies that will be required in order for firm assignments to be made to each of the 24 genes in the *bad–hba* cluster. We anticipate that the sequences reported here will stimulate studies on the properties of the enzymes of anaerobic benzoate and 4-OHBen degradation. For example, the cloned *bad–hba* genes will permit site-directed mutagenesis as an approach to probing enzyme mechanisms. This will be especially useful in studies of the benzoyl-CoA and 4-hydroxybenzoyl-CoA reductases. It may be that the rate-limiting step in the overall degradation of many aromatic compounds, benzoyl-CoA reduction, can be relieved through genetic engineering. The *R. palustris* benzoate degradation genes will also be useful in comparative studies with other physiological types of anaerobic bacteria, and possibly, for identifying homologous genes from other groups of anaerobes. It is generally assumed that the essential biochemistry of attack on the benzene ring is the same in fermentative bacteria, phototrophic bacteria (e.g., *R. palustris*), and different groups of anaerobic respiring bacteria (nitrate-, sulfate-, and iron-respirers), but the possibility that there may be major or minor differences in the enzymes of benzoate degradation among different physiological types of anaerobes needs to be explored for both fundamental and practical reasons.

This work was supported by the Department of Energy, Division of Energy Biosciences (DE-FG02-86ER13495 to J.G. and DE-FG02-95ER20184 to C.S.H.), and by the U.S. Army Research Office (DAAH04-95-0124 and -0315 to C.S.H.).

1. Dutton, P. L. & Evans, W. C. (1969) *Biochem. J.* **113,** 525–536.
2. Gibson, J. & Harwood, C. S. (1995) in *Anoxygenic Photosynthetic Bacteria*, eds. Blankenship, R. E., Madigan, M. T. & Bauer, C. E. (Kluwer, Dordrecht, The Netherlands), pp. 991–1003.
3. Harwood, C. S. & Gibson, J. (1997) *J. Bacteriol.* **179,** 301–309.
4. Heider, J. & Fuchs, G. (1997) *Eur. J. Biochem.* **243,** 577–596.
5. Koch, J., Eisenreich, W., Bacher, A. & Fuchs, G. (1993) *Eur. J. Biochem.* **221,** 649–661.
6. Gibson, K. J. & Gibson, J. (1992) *Appl. Environ. Microbiol.* **58,** 696–698.
7. Geissler, J. F., Harwood, C. S. & Gibson, J. (1988) *J. Bacteriol.* **170,** 1709–1714.
8. Altenschmidt, U., Oswald, B. & Fuchs, G. (1991) *J. Bacteriol.* **173,** 5494–5501.
9. Boll, M. & Fuchs, G. (1995) *Eur. J. Biochem.* **234,** 921–933.
10. Pelletier, D. A. & Harwood, C. S. (1996) *Abstr. Annu. Meet. Am. Soc. Microbiol.* **96,** Q-161.
11. Küver, J., Xue, J. Y. & Gibson, J. (1995) *Arch. Microbiol.* **164,** 337–345.
12. Gibson, J., Dispensa, M., Fogg, G. C., Evans, D. T. & Harwood, C. S. (1994) *J. Bacteriol.* **176,** 634–641.
13. Biegert, T., U. Altenschmidt, Eckerskorn, C. & Fuchs, G. (1993) *Eur. J. Biochem.* **213,** 555–561.
14. Brackmann, R. & Fuchs, G. (1993) *Eur. J. Biochem.* **213,** 563–571.
15. Harayama, S., Kok, M. & Neidle, E. L. (1992) *Annu. Rev. Microbiol.* **46,** 565–601.
16. Kim, M.-K. & Harwood, C. S. (1991) *FEMS Microbiol. Lett.* **83,** 199–204.
17. Gibson, J., Dispensa, M. & Harwood, C. S. (1997) *J. Bacteriol.* **179,** 634–642.
18. Reynolds, K. A., Wang, P., Fox, K. M. & Floss, H. G. (1992) *J. Antibiot.* **45,** 411–419.
19. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1990) *Current Protocols in Molecular Biology* (Greene, New York).
20. Egland, P. G., Gibson, J. & Harwood, C. S. (1995) *J. Bacteriol.* **177,** 6545–6551.
21. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
22. De Lorenzo, V., Cases, I., Herrero, M. & Timmis, K. N. (1993) *J. Bacteriol.* **175,** 6902–6907.
23. Quandt, J. & Hynes, M. F. (1993) *Gene* **127,** 15–21.
24. Parales, R. E. & Harwood, C. S. (1993) *J. Bacteriol.* **175,** 5829–5838.
25. Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., II & Peterson, K. M. (1994) *BioTechniques* **16,** 800–802.
26. Perrotta, J. A. & Harwood, C. S. (1994) *Appl. Environ. Microbiol.* **60,** 1775–1782.
27. Miller, J. H. (1992) *Experiments in Molecular Genetics* (Cold Spring Harbor Lab. Press, Plainview, NY).
28. Anders, H.-J., Kaetzke, A., Kämpfer, P., Ludwig, W. & Fuchs, G. (1995) *Int. J. Syst. Bacteriol.* **45,** 327–333.
29. Dutscho, R., Wohlfarth, G., Buckel, P. & Buckel, W. (1989) *Eur. J. Biochem.* **181,** 741–746.
30. Merkel, S. M., Eberhard, A. E., Gibson, J. & Harwood, C. S. (1989) *J. Bacteriol.* **171,** 1–7.
31. Müller, U. & Buckel, W. (1995) *Eur. J. Biochem.* **230,** 698–704.
32. Buckel, W. (1996) *FEBS Lett.* **389,** 20–24.
33. Buckel, W. & Keese, R. (1995) *Angew. Chem. Int. Ed. Engl.* **34,** 1502–1506.
34. Bork, P., Sander, C. & Valencia, A. (1992.) *Proc. Natl. Acad. Sci. USA* **89,** 7290–7294.
35. Nunn, W. D. (1986) *Microbiol. Rev.* **50,** 179–192.
36. Persson, B., Krook, M. & Jörnvall, H. (1991) *Eur. J. Biochem.* **200,** 537–543.
37. Roper, D. I., Fawcett, T. & Cooper, R. A. (1993) *Mol. Gen. Genet.* **237,** 241–250.
38. Higgins, C. F. (1992) *Annu. Rev. Cell Biol.* **8,** 67–113.
39. Dispensa, M., Thomas, C. T., Kim, M.-K., Perrotta, J. A., Gibson, J. & Harwood, C. S. (1992) *J. Bacteriol.* **174,** 5803–5813.