The fluorene catabolic linear plasmid in *Terrabacter* sp. strain DBF63 carries the β -ketoadipate pathway genes, *pcaRHGBDCFIJ*, also found in proteobacteria

Hiroshi Habe,¹ Jin-Sung Chung,¹ Ayako Ishida,² Kano Kasuga,³ Kazuki Ide,⁴ Tetsuo Takemura,² Hideaki Nojiri,¹ Hisakazu Yamane¹ and Toshio Omori^{1†}

Correspondence

Hiroshi Habe

uhhabe@mail.ecc.u-tokyo.ac.jp

¹Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

²Department of Chemistry, Faculty of Sciences, Science University of Tokyo, 1-3 Kagurazaka, Shinjuku-ku, Tokyo 162-8601, Japan

³Department of Biotechnology, Faculty of Bioresource Sciences, Akita Prefectural University, 241-7 Kaidobata-nishi, Shimoshinjo-nakano, Akita 010-0195, Japan

⁴Geo and Water Environmental Engineering Department, Obayashi Corporation, 4-640 Shimokiyoto, Kiyose-shi, Tokyo 204-0011, Japan

Terrabacter sp. strain DBF63 is capable of degrading fluorene (FN) to tricarboxylic acid cycle intermediates via phthalate and protocatechuate. Genes were identified for the protocatechuate branch of the β -ketoadipate pathway (*pcaR*, *pcaHGBDCFIJ*) by sequence analysis of a 70 kb DNA region of the FN-catabolic linear plasmid pDBF1. RT-PCR analysis of RNA from DBF63 cells grown with FN, dibenzofuran, and protocatechuate indicated that the *pcaHGBDCFIJ* operon was expressed during both FN and protocatechuate degradation in strain DBF63. The gene encoding β -ketoadipate enol-lactone hydrolase (*pcaD*) was not fused to the next gene, which encodes γ -carboxymuconolactone decarboxylase (*pcaC*), in strain DBF63, even though the presence of the *pcaL* gene (the fusion of *pcaD* and *pcaC*) within a *pca* gene cluster has been thought to be a Gram-positive trait. Quantitative RT-PCR analysis revealed that *pcaD* mRNA levels increased sharply in response to protocatechuate, and a biotransformation experiment with *cis,cis*-muconate using *Escherichia coli* carrying both *catBC* and *pcaD* indicated that PcaD exhibited β -ketoadipate enol-lactone hydrolase activity. The location of the *pca* gene cluster on the linear plasmid, and the insertion sequences around the *pca* gene cluster suggest that the ecologically important β -ketoadipate pathway genes, usually located chromosomally, may be spread widely among bacterial species via horizontal transfer or transposition events.

Received 24 May 2005

Revised 28 July 2005

Accepted 11 August 2005

INTRODUCTION

Polycyclic aromatic hydrocarbons (PAHs) are a ubiquitous class of hydrophobic organic compounds commonly found

in the environment because of industrial activities, such as coal gasification, wood preservation, liquefaction and fuel refinement. Given that PAHs have toxic, mutagenic and carcinogenic properties, their persistence in the environment is of major concern. Therefore, the degradation of PAHs by diverse aerobic bacteria has been extensively studied in the last two decades (Cerniglia, 1992; Habe & Omori, 2003; Kanaly & Harayama, 2000; Pinyakong *et al.*, 2003; Shuttleworth & Cerniglia, 1995). The microbial degradation of fluorene (FN), classified by the US Environmental Protection Agency as one of sixteen priority PAHs (Keith & Telliard, 1979), has been studied because FN is a major constituent of coal derivatives (e.g. about 7.6 % of creosote PAHs; Grifoll *et al.*, 1995a). Furthermore, FN and a number of its alkyl derivatives have been detected in vehicle

[†]Present address: Department of Industrial Chemistry, Faculty of Engineering, Shibaura Institute of Technology, Minato-ku, Tokyo 108-8548, Japan.

Abbreviations: DF, dibenzofuran; FN, fluorene; IR, inverted repeat; IS, insertion sequence; PAH, polycyclic aromatic hydrocarbon.

The GenBank/EMBL/DDBJ accession number for the sequence reported in this paper is AP008980.

A figure showing RT-PCR and expression analysis of *pca* genes, and a table of annotated ORFs are available as supplementary material with the online version of this paper.

exhaust, air particles and sediments (Grifoll *et al.*, 1995b; Jensen & Hites, 1983). Nevertheless, little is known about bacterial FN-catabolic genes. The genes involved in the degradation of FN to protocatechuate via phthalate have only been identified in *Terrabacter* sp. strain DBF63 (Habe *et al.*, 2003, 2004; Kasuga *et al.*, 2001).

Strain DBF63 was originally isolated as a bacterium capable of utilizing both dibenzofuran (DF) and FN (Monna *et al.*, 1993). Initially, *dbfA1* and *dbfA2* were isolated from strain DBF63 as genes encoding novel terminal oxygenase components of the angular dioxygenase for DF degradation (Kasuga *et al.*, 2001), and later, the genes were revealed to be located within an FN-catabolic gene cluster (Habe *et al.*, 2004). The FN-catabolic genes, designated *dbf-fln* (involved with the degradation of FN to phthalate) and *pht* (involved with the degradation of phthalate to protocatechuate), were located on linear plasmids pDBF1 and pDBF2 (approx. 160 and 190 kb, respectively), and pDBF2 is thought to be derived from pDBF1 (Nojiri *et al.*, 2002a). The presence of protocatechuate catabolic genes in strain DBF63, which are required for the complete degradation of FN to TCA cycle intermediates, is supported by the fact that strain DBF63 can grow with protocatechuate as its sole carbon and energy source.

Two major catabolic pathways for protocatechuate have been proposed. The first is the *meta*-cleavage pathway initiated by protocatechuate 4,5-dioxygenase, and the second is the β -ketoadipate pathway initiated by protocatechuate 3,4-dioxygenase (Dagley *et al.*, 1960; Fig. 1). The β -ketoadipate pathway, in particular, is considered to be a good model system for studying mechanisms of evolution in an ecologically important catabolic pathway (Buchan *et al.*, 2004; Parke *et al.*, 2000). Following initial *ortho*-cleavage by protocatechuate 3,4-dioxygenase (PcaHG), five additional enzymes, i.e. β -carboxy-*cis,cis*-muconate cycloisomerase (PcaB), γ -carboxymuconolactone decarboxylase (PcaC), β -ketoadipate enol-lactone hydrolase (PcaD), β -ketoadipate succinyl CoA transferase (PcaIJ) and β -ketoadipyl CoA thiolase (PcaF), convert the ring cleavage product to TCA cycle intermediates (Fig. 1). In the case of the proteobacteria, this pathway is biochemically conserved among phylogenetically different strains, but operon organization, regulatory proteins, coinducer molecules and transport proteins for the pathway are remarkably diverse, most likely reflecting subtle aspects of niche adaptation (Buchan *et al.*, 2004; Parke, 1997; Parke *et al.*, 2000). In contrast, there is still little information available concerning protocatechuate catabolic gene clusters from actinobacteria (Eaton, 2001; Eulberg *et al.*, 1998; Iwagami *et al.*, 2000). The protocatechuate branch of the β -ketoadipate pathway (involving *pca* gene products) has only been functionally analysed in two systems, i.e. *Rhodococcus* sp. (Eulberg *et al.*, 1998; Patrauchan *et al.*, 2005) and *Streptomyces* sp. (Iwagami *et al.*, 2000). Analysis of the *pca* genes in *Terrabacter* may, therefore, provide a better perspective on the evolution of the β -ketoadipate pathway in actinobacteria, if this strain has such genes.

In this study, to elucidate all of the FN-degradative genes in *Terrabacter* sp. strain DBF63, we sequenced the 70 kb DNA region containing the already identified *dbf-fln* and *pht* genes. Interestingly, the *pca* genes, usually located chromosomally, were located on the linear plasmid. This is believed to be the first description of *pca* genes from actinobacteria, in which the β -ketoadipate enol-lactone-hydrolase and γ -carboxymuconolactone decarboxylase enzymes are encoded by separate genes, *pcaD* and *pcaC*, respectively.

METHODS

Bacterial strains, plasmids and culture conditions. The bacterial strains and plasmids used in this study are listed in Table 1. *Terrabacter* sp. strain DBF63 was routinely cultivated as previously reported (Kasuga *et al.*, 1997). Briefly, this strain was cultivated in a carbon-free mineral medium (CFMM) supplemented with each substrate at a final concentration 0.1% (w/v) at 30 °C. The CFMM contained, per litre of distilled water: 2.2 g Na₂HPO₄, 0.8 g KH₂PO₄, 3.0 g NH₄NO₃, 0.2 g MgSO₄·7H₂O, 50 mg CaCl₂·2H₂O, 50 mg FeCl₃·6H₂O. Strains of *Escherichia coli* were cultivated using LB broth, 2 × YT medium or Terrific broth, as described by Sambrook *et al.* (1989). Ampicillin sodium salt, chloramphenicol and IPTG were added to the media at final concentrations of 100 µg ml⁻¹, 30 µg ml⁻¹ and 1 mM, respectively, when necessary.

DNA manipulations. Preparation of the total DNA of strain DBF63 and recombinant DNA manipulations were performed as reported previously (Kasuga *et al.*, 1997). DNA fragments were purified using an EZNA gel extraction kit (Omega Bio-tek) according to the manufacturer's instructions. Other DNA manipulations were carried out as described by Sambrook *et al.* (1989), and commercially available enzymes and kits were used as indicated by their manufacturers.

Sequencing and annotation of the 70 kb DNA region of the linear plasmid. Shotgun sequencing of (i) two previously isolated cosmid clones (designated pCC4 and pCC19) that contain the *phtA1* gene (Kasuga *et al.*, 2001), and (ii) one cosmid clone (designated pCC31) that contains the flanking region of the pCC4 insert (this study) was performed by Dragon Genomics. Fragmented, blunt-ended DNA was ligated into the *Sma*I site of pUC118. Determination of the nucleotide sequence was carried out by the chain-termination method. To identify ORFs, the nucleotide sequence was analysed with DNASIS-Mac software (version 3.7; Hitachi Software Engineering). Homology searches were carried out using the SWISS-PROT protein sequence database or the DDBJ, EMBL and GenBank nucleotide sequence databases with BLAST programs (Altschul *et al.*, 1997). Results from the automated ORF prediction and functional assignment were manually controlled for the entire DNA contiguous sequence (70 kb). The nucleotide sequence was deposited at the DDBJ, EMBL and GenBank nucleotide sequence databases under the accession number AP008980.

RT-PCR analysis. Strain DBF63 was grown to mid-exponential phase in CFMM supplemented with FN, DF or protocatechuate solutions to a final substrate concentration of 0.1% (w/v). Both FN and DF were dissolved in dimethylformamide (DMF; 100 mg ml⁻¹), and protocatechuate was dissolved in ethanol (100 mg ml⁻¹). These solutions were sterilized by filtration with a 0.2 µm pore-size membrane filter, and 50 µl each solution was added to 5 ml CFMM. Two millilitres of each culture was then centrifuged, and total RNA was extracted from the harvested cells using a NucleoSpin RNA II purification kit (Macherey-Nagel), combined with RQ1 RNase-free DNase (Promega), according to the manufacturer's instructions. The gene-specific reverse primers used to synthesize cDNA were RT-Pr3R for

Fig. 1. The protocatechuate and catechol branches of the β -ketoadipate pathway in bacteria. The proposed pathway of FN degradation by *Terrabacter* sp. strain DBF63 is represented with bold arrows. Compound designations: I, fluorene; II, protocatechuate; III, β -carboxy-*cis,cis*-muconate; IV, γ -carboxymuconolactone; V, catechol; VI, *cis,cis*-muconate; VII, muconolactone; VIII, β -ketoadipate enol-lactone; IX, β -ketoadipate. Enzyme designations: PcaHG, protocatechuate 3,4-dioxygenase; PcaB, β -carboxy-*cis,cis*-muconate cycloisomerase; PcaC, γ -carboxymuconolactone decarboxylase; PcaD and CatD, β -ketoadipate enol-lactone hydrolase; PcaL, a fused γ -carboxymuconolactone decarboxylase/ β -ketoadipate enol-lactone hydrolase; CatA, catechol 1,2-dioxygenase; CatB, muconate cycloisomerase; CatC, muconolactone isomerase; PcaIJ and CatIJ, β -ketoadipate succinyl CoA transferase; PcaF and CatF, β -ketoadipyl CoA thiolase.

pcaHG, RT-Pr5R for *pcaDC*, RT-Pr7Rb for *pcaFJI* and RT-Pr1R for *pcaR* (Table 2). cDNA of each gene was synthesized from 10 ng total RNA using SuperScript III reverse transcriptase (Invitrogen), according to the manufacturer's instructions. The primers used are listed in Table 2. After the cDNA samples were properly diluted, PCR was performed as follows: (i) for *pcaHG* amplification, 95 °C for 10 min and 5 s, 60 °C for 10 s, and 72 °C for 30 s, for 40 cycles; (ii) for *pcaDC* amplification, 95 °C for 10 min and 30 s, 50 °C for 30 s, and 72 °C for 90 s, for 40 cycles; (iii) for *pcaFJI* amplification, 95 °C for

10 min and 5 s, 55 °C for 30 s, and 72 °C for 30 s, for 40 cycles; (iv) for *pcaR* amplification, 95 °C for 10 min and 30 s, 55 °C for 30 s, and 72 °C for 90 s, for 40 cycles.

Quantitative RT-PCR analysis. After a 4 day cultivation of strain DBF63 in 100 ml CFMM supplemented with DF, the cells were centrifuged and then washed twice with CFMM. The washed cells were resuspended in CFMM to an OD₆₀₀ of approximately 0.6, and 5 ml portions of the resulting cell suspension were transferred to four test

Table 1. Bacterial strains and plasmids

Strain or plasmid	Relevant characteristic	Reference or source
Strains		
<i>Terrabacter</i> sp. DBF63	FN utilization (FN ⁺)	Monna <i>et al.</i> (1993)
<i>Terrabacter</i> sp. DBF63W	FN ⁻ mutant of DBF63	Kasuga <i>et al.</i> (1997)
<i>E. coli</i> JM109	<i>recA1</i> , <i>endA1</i> , <i>gyrA96</i> , <i>thi-1</i> , <i>hsdR17</i> , <i>supE44</i> , <i>relA1</i> , $\Delta(lac-proAB)/F'[traD36, proAB^+, lacI^q, lacZ\Delta M15]$	Yanisch-Perron <i>et al.</i> (1985)
Cosmids		
pCC12	Ap ^r , Km ^r , SuperCos1 (Stratagene) with <i>Bam</i> HI insert of DBF63 DNA	Kasuga <i>et al.</i> (2001)
Plasmids		
pUC18/19	Ap ^r , <i>lacZ</i> ; pMB9 replicon	Yanisch-Perron <i>et al.</i> (1985)
pUC118/119	Ap ^r , <i>lacZ</i> ; pMB9 replicon	Yanisch-Perron <i>et al.</i> (1985)
pSTV28/29	Cm ^r , <i>lacZ</i> ; pMB9 replicon	Takara Shuzo
pT7Blue(R)	Ap ^r , T-vector	Novagen
pBluescript II KS(+/-)	Ap ^r , <i>lacZ</i> ; pMB9 replicon	Stratagene
pUCA811	Ap ^r , 3.2 kb <i>Eco</i> RI fragment from <i>Pseudomonas resinovorans</i> strain CA10 DNA inserted into <i>Eco</i> RI-digested pUC19; carries <i>catBC</i>	Nojiri <i>et al.</i> (2002b)
pUCA832	Ap ^r , 1.9 kb <i>Xho</i> I- <i>Eco</i> RI fragment from pUCA811 inserted into <i>Xho</i> I/ <i>Eco</i> RI-digested pBluescript II KS(-); carries <i>catBC</i> , <i>lac</i> orientation	This study
pARO523	Ap ^r ; carries <i>pcaHG</i> and <i>pcaB</i> from <i>Agrobacterium tumefaciens</i> A348 on mobilizable pUC19 derivative pARO190	Parke (1995)
pDFS603	Ap ^r , 0.4 kb PCR product (using pCC12 as a template) inserted into <i>Hind</i> III/ <i>Eco</i> RI-digested pSTV29; carries <i>pcaC</i> , <i>lac</i> orientation	This study
pDFS604	Ap ^r , 0.8 kb PCR product (using pCC12 as a template) inserted into <i>Hind</i> III/ <i>Eco</i> RI-digested pSTV29; carries <i>pcaD</i> , <i>lac</i> orientation	This study

tubes. The cells in each test tube were incubated at 30 °C with reciprocal shaking at 300 strokes min⁻¹ for 12 h as a starvation culture. The starved cells were washed twice as described above, and finally, resuspended in 5 ml CFMM supplemented with 0.1% (w/v) succinate. Following the addition of 50 µl ethanol (control), protocatechuate solution, DMF solution (control) or FN solution to each test tube, the cells were incubated at 30 °C for 1 h, harvested and then used for the extraction of total RNA. To isolate total RNA from harvested cells, a NucleoSpin RNA II purification kit (Macherey-Nagel) was used, combined with RQ1 RNase-free DNase (Promega), according to the manufacturer's instructions. The gene-specific reverse primers used to synthesize cDNA were qRT-pcaDRb2 for *pcaD*, and qRT-16SR for 16S rDNA (Table 2). Each cDNA was synthesized from 200 ng total RNA using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. Quantitative real-time PCR with SYBR Green PCR Master Mix (Applied Biosystems) was carried out using the synthesized cDNA as a template and the ABI7700 Sequence Detection System (Applied Biosystems). To calculate mRNA levels of the *pcaD* gene, the copy number of *pcaD* mRNA was determined by generating a standard curve using a series of samples of the targeted sequence with known concentrations. For normalization of the samples, 16S rRNA was used as an internal standard. For each sample, the mean value from triplicate real-time PCR experiments was adjusted and used to calculate the transcript abundance. The mean values of the *pcaD* mRNA levels in the control samples (supplemented with ethanol or DMF) were arbitrarily set as 1.0.

Construction of plasmids for *pcaD*, *pcaC* and *catBC* expression. The 820 bp and 415 bp DNA fragments containing *pcaD* and *pcaC*, respectively, were prepared by PCR using pCC12 (Kasuga *et al.*, 2001) as the template. The primer sets used are listed in

Table 2. PCR was carried out using LA *Taq* with GC buffer (Takara Shuzo) and PCR Thermal Cycler Dice (Takara Shuzo). The cycling conditions were as follows: 96 °C for 1 min, followed by 30 cycles of 96 °C for 30 s, 55 °C for 30 s, and 72 °C for 2 min, followed by 72 °C for 6 min. The PCR products were cloned using the pT7Blue (R) vector (Novagene), and the nucleotide sequences of the PCR products were confirmed by sequencing. The clones were digested with both *Hind*III and *Eco*RI (sites derived from the primer, Table 2), and then the fragments were cloned between the *Hind*III and *Eco*RI sites of pSTV29 to give pDFS604 (carrying *pcaD*) and pDFS603 (carrying *pcaC*), respectively. To construct the expression plasmid for *catBC* genes, the 1.9 kb *Xho*I-*Eco*RI fragment of pUCA811 (Nojiri *et al.*, 2002b) was cloned between the corresponding sites of pBluescript II KS(-) to give pUCA832 (carrying *catBC*).

Resting cell reactions and analysis of the products.

Transformation of *E. coli* cells with the constructed plasmids was performed as described by Hanahan (1983). An appropriate *E. coli* transformant was grown in 100 ml LB broth supplemented with ampicillin, chloramphenicol and IPTG, as appropriate, in a 500 ml flat-bottom flask at 30 °C with reciprocal shaking at 120 strokes min⁻¹. The culture broth was centrifuged (4000 g), and the pelleted cells were washed twice with 100 ml buffer (containing 2.2 g Na₂HPO₄, 0.8 g KH₂PO₄, and 3.0 g NH₄NO₃ per litre of distilled water; pH 7.0), and resuspended in the same buffer to an OD₆₀₀ of approximately 10. We added *cis,cis*-muconate (0.1% w/v) to 5 ml resting cell suspension in test tubes and incubated the tubes at 30 °C for 2 h. The reaction mixtures were acidified with 1 M HCl and extracted with 5 ml ethyl acetate. The extracts were analysed by GC-MS after derivatization with *N*-methyl-*N*-trimethylsilyltrifluoroacetamide. GC-MS analyses were performed with a JMS-Automass 150 GC-MS system (JEOL) fitted with a fused-silica chemically bonded capillary

Table 2. Primers used in RT-PCR, quantitative RT-PCR and plasmid construction

Target gene	Name of primer	Nucleotide sequence*
For RT-PCR		
<i>pcaR</i>	RT-Pr1R†	5'-CCGCTTGAGTGTGGGTAC-3'
	RT-Pr1F	5'-GCTGTCCGCCCTTTGACAC-3'
<i>pcaHG</i>	RT-Pr3R†	5'-ACGTCTGAAGCGCAGACTC-3'
	RT-Pr3F	5'-CAGCCCATGATCGACTTC-3'
<i>pcaDC</i>	RT-Pr5R†	5'-GTAGCGGGTGATCAGCTC-3'
	qRT-pcaDFb2	5'-TCGTCGGCCCATCTCTTG-3'
<i>pcaFIJ</i>	RT-Pr7R†	5'-AGAAGGAGATGCCGAAGAG-3'
	RT-Pr7Fb	5'-GAGCTGAATGAGGCCTTC-3'
Intergenic region between <i>pcaR</i> and <i>pcaH</i>	RT-Pr2HR†	5'-TGCCGCAGCAGTGAAGTG-3'
	RT-Pr2RR	5'-GTCGCTCAAGTTAGATG-3'
For quantitative RT-PCR		
<i>pcaD</i>	qRT-pcaDRb2†	5'-ACGACCTCGAAGTCATAG-3'
	qRT-pcaDFb2	5'-TCGTCGGCCCATCTCTTG-3'
16S rDNA	qRT-16SR†	5'-TGCCCGTACCCACTGCAAGTC-3'
	qRT-16SF	5'-TGACGGTACCTGCAGAAGAAG-3'
For plasmid construction		
<i>pcaD</i>	pcaD-F	5'-TTAAGCTTTAAGGAGGGCTGCGCATGGCTGAGCCGGACCTG-3'
	pcaD-R	5'-AAGAATTCTCATCGTTTCCCTCCGGTCGAAAGTCGTCC-3'
<i>pcaC</i>	pcaC-F	5'-TTAAGCTTTAAGGAGGGGAAACGATGAGCGACACCTACGACG-3'
	pcaC-R	5'-AAGAATTCTCATTTCATCAGGGGTCTCCAGGGAGTCGAGCG-3'

*The italic, underlined and double-underlined sequences are an efficient Shine–Dalgarno sequence (Shine & Dalgarno, 1975), a *Hind*III site and an *Eco*RI site, respectively.

†The gene-specific reverse primers were used for synthesizing cDNA.

column (DB-5; 0.25 mm inside diameter by 15 m, 0.25 µm film thickness; J&W Scientific). Each sample was injected into the column at 80 °C in the splitless mode. After 2 min at 80 °C, the column temperature was increased at 16 °C min⁻¹ to 280 °C. The head pressure of the helium carrier gas was 65 kPa.

Chemicals. β -Ketoadipic acid (3-oxoadipic acid) was purchased from Sigma-Aldrich. *cis,cis*-Muconic acid was synthesized by Mitsubishi Chemical. All other chemicals used in this study were of the highest purity commercially available (98–100%; Merck, Sigma-Aldrich, Kanto Chemical, Wako Pure Chemical, Nacalai Tesque).

RESULTS

Sequencing of the 70 kb DNA region of linear plasmid pDBF1

The shotgun sequencing of three cosmid clones produced a 70 631 bp contiguous sequence that included the 20 574 bp region previously analysed (Habe *et al.*, 2003, 2004). In addition to the already determined ORFs (ORF1–ORF20), 57 additional ORFs (ORF21–ORF77) and one incomplete ORF (ORF78) were annotated (Fig. 2; see the Supplementary Table with the online journal). The mean G+C content of the 70 631 bp region was 68.6%, close to that of the genome of strain DBF63 (70.5%; Kasuga *et al.*, 1997).

(i) Protocatechuate catabolic genes. Based on a homology search (see the Supplementary Table), nine genes, i.e. *pcaH* (ORF56), *pcaG* (ORF55), *pcaB* (ORF54), *pcaD* (ORF53), *pcaC* (ORF52), *pcaF* (ORF51), *pcaI* (ORF50) and *pcaJ* (ORF49), were identified that apparently coded for the protocatechuate 3,4-dioxygenase β subunit (PcaH), protocatechuate 3,4-dioxygenase α subunit (PcaG), β -carboxy-*cis,cis*-muconate cycloisomerase (PcaB), γ -carboxy-muconolactone decarboxylase (PcaC), β -ketoadipate enol-lactone hydrolase (PcaD), β -ketoadipate succinyl CoA transferase α subunit (PcaI), β -ketoadipate succinyl CoA transferase β subunit (PcaJ) and β -ketoadipyl CoA thio-lase (PcaF), respectively (Fig. 1). The start codons (ATG or GTG) of all *pca* genes overlapped with the stop codon (TGA) of the preceding *pca* gene by 1 or 4 bp (data not shown), indicating that these nine genes were transcribed together. Residues known to be important for catalytic function were well conserved in the DBF63 Pca protein sequences, e.g. the residues involved in coordinating the nonheme Fe³⁺ (Tyr408, Tyr447, His460 and His462), as well as active-site residues (e.g. Thr12, Pro15, Tyr324, Trp400, Trp449, Arg457 and Ile491) in PcaHG of *Acinetobacter* sp. strain ADP1 (Orville *et al.*, 1997; Vetting *et al.*, 2000), and the cysteine residue demonstrated to be critical for hydrolysis activity (Cys123) in PcaD from *Pseudomonas* sp. strain B13 (Pathak *et al.*, 1991). ORF57

Fig. 2. Diagram indicating the positions of ORFs present in the 70 kb DNA region of linear plasmid pDBF1. The blocks indicate the size, location and direction of transcription of the ORFs. The numbers underneath the blocks refer to ORF numbers, and those above the vertical lines refer to nucleotide sequence numbers. Hatched blocks, ORFs probably involved in plasmid replication and maintenance; grey blocks, ORFs showing significant similarity with the corresponding ORFs on the 350 kb linear plasmid SCP1 from *S. coelicolor* strain A3(2) (Yamasaki *et al.*, 2001) and the 210 kb linear plasmid pBD2 from *Rhodococcus erythropolis* strain BD2 (Stecker *et al.*, 2003); black blocks, transposase gene homologues; cross-hatched blocks, FN-degradative genes. For several ORFs, gene names or putative functions are indicated above the blocks. BP, binding protein; IS, insertion sequence; DNA-P, DNA polymerase.

is transcribed divergently from *pcaH*, and encodes a protein with the helix–turn–helix motif at its N-terminal region and shares identities with the putative transcriptional regulators belonging to the IclR family, so we designated it *pcaR*.

(ii) Insertion sequences (ISs). We found several ORFs homologous to transposase genes or IS-like elements around the *pca* and *dbf-fln* gene clusters (Fig. 2). The deduced amino acid sequence of ORF58 located downstream of *pcaR* exhibited 100% similarity to that of the transposase-encoding gene in ISTesp2, found just upstream of the *pht* genes in strain DBF63 (Habe *et al.*, 2003). We also determined that the incomplete ORF78 and its flanking region were homologous to ISTesp2 (Fig. 2).

The DNA sequence revealed that ORF68 and its flanking region (nucleotides 61 177–62 584; Fig. 2) were identical to ORF70 and its flanking region (nucleotides 64 306–65 713),

and that the full-length of these repetitive elements was 1408 bp. The deduced amino acid sequences of ORF42, ORF68 and ORF70 showed 82, 84 and 84% overall lengthwise similarity, respectively, to the transposase of IS1601-A from *Mycobacterium avium* (see the Supplementary Table), a predicted member of the IS256 family (Eckstein *et al.*, 2000). The similarity between the ORF42 product and ORF68 (ORF70) product was 85%. At both ends of ORF42 and ORF68 (ORF70), we found 20 and 23 bp imprecise inverted repeats (IRs), respectively. However, the formation of direct repeats, as duplicated target sequences at both ends, was not observed in the immediate vicinity of these IRs.

The polypeptide encoded by ORF67 and ORF66 showed 66–68% and 52–69% overall lengthwise similarity to the transposase subunits encoded by MC34 and MC35, respectively, from the plasmid pSD10 of *Micrococcus* sp. strain 28 (AY034092.1), and by PDB2.181 and PDB2.182, respectively, from plasmid pBD2 of *Rhodococcus erythropolis* strain

BD2 (Stecker *et al.*, 2003). This IS is potentially a member of the IS3 family. The 24 bp imperfect IR was found around the two ORFs, but a direct repeat, duplicated target site, was not observed in the immediate vicinity of the IR.

Transcriptional analysis of *pca* genes by RT-PCR

We used RT-PCR analysis to investigate whether *pca* genes were transcribed in strain DBF63 during its growth with FN. Total RNA was prepared from DBF63 cells grown with FN, DF or protocatechuate. As the *pcaHGBDCFIJ* gene cluster has been shown, by RT-PCR, to be transcribed as an operon (data not shown), we only present here the results with the three primers specific for amplifying the *pcaHG*, *pcaDC* and *pcaFIJ* genes within this cluster (Table 2; Supplementary Fig. 1a with the online journal). Clear amplification of the appropriate portion of the respective *pca* genes (1.2 kb for *pcaHG*, 0.85 kb for *pcaDC* and 1.4 kb for *pcaFIJ*) was detected when RNA prepared from both FN- and protocatechuate-grown cells were used, whereas PCR products were not detected when RNA from DF-grown cells was used. No amplified products were detected in the negative control lacking reverse transcriptase (data not shown). These results indicate that *pcaHGBDCFIJ* genes were transcribed in DBF63 cells grown with FN, as well as protocatechuate, suggesting that the genes are probably required for FN degradation in this strain. As strain DBF63 degrades DF via gentisate (Monna *et al.*, 1993), it is reasonable that PCR products are not detected from cells grown on DF. In contrast, the *pcaR* gene was not specifically expressed during the degradation of FN or protocatechuate, and the same levels of PCR products were detected from RNA prepared from DBF63 cells grown on all three substrates (see Supplementary Fig. 1a). This result suggests that the *pcaR* gene may be constitutively expressed in this strain. RT-PCR with primers specific for amplifying the intergenic region between *pcaR* and *pcaH* (Table 2) produced no PCR products of the expected size (310 bp; data not shown).

Quantitative RT-PCR

To analyse the expression pattern of *pca* genes, the *pcaD* mRNA levels in protocatechuate-induced or non-induced DBF63 cells, within 1 h of induction, were determined by quantitative RT-PCR with SYBR Green. Transcriptional levels were normalized with 16S rRNA as an internal control. As shown in Supplementary Fig. 1(b) (left panel), the *pcaD* mRNA in cells with protocatechuate was about 21-fold more abundant than in cells with ethanol (control). Similarly, the *pcaD* mRNA levels in FN-induced or non-induced DBF63 cells were examined (see Supplementary Fig. 1b, right panel), and expression levels 1.7-fold higher were observed in cells with FN. The differences in water solubility of these inducer substrates, and the solvent types, may have had some effect on induction levels. In addition, incubation times longer than 1 h would produce higher induction levels in FN-induced cells. Nevertheless, the patterns of sharp and gradual increase in the *pcaD* mRNA levels in

response to protocatechuate and FN, respectively, indicate that protocatechuate itself or its metabolite, e.g. β -ketoadipate, is the inducer of *Pca* enzymes.

β -Ketoadipate enol-lactone hydrolase (*PcaD*) activity

We constructed the plasmids pUCA832 (carrying *catBC*) and pDFS604 (carrying *pcaD*; Table 1), and performed a biotransformation experiment with *cis,cis*-muconate using *E. coli* cells harbouring both pUCA832 and pDFS604. *E. coli* JM109 cells carrying both pUCA832 and pSTV29 were used as a control. GC-MS analysis of products from *cis,cis*-muconate after treatment with *N*-methyl-*N*-trimethylsilyl-trifluoroacetamide, gave a mass spectrum that exhibited fragment ions at m/z 378 (M^+ , 6), 361 (27), 317 (6), 286 (16), 259 (6), 231 (21), 169 (59), 147 (39), 125 (8) and 73 (100), where the numbers in parentheses represent relative intensity, detected with a retention time of 8.4 min. This fragmentation pattern of the trimethylsilyl derivative of the metabolite was identical to that of authentic β -ketoadipate (data not shown). In contrast, β -ketoadipate was not detected in the control sample (data not shown). Instead, the trimethylsilyl derivative of β -ketoadipate enol-lactone was tentatively identified because the mass spectrum (m/z) 286 (M^+ , 14), 271 (5), 169 (100), 147 (19) and 73 (100) and the retention time (6.3 min) were consistent with those of the product when we performed biotransformation experiments with protocatechuate using *E. coli* cells harbouring both pARO523 (carrying *pcaHGB*; Parke, 1995) and pDFS603 (carrying *pcaC*; see Fig. 1). The above result indicates that the *pcaD* gene encodes β -ketoadipate enol-lactone hydrolase.

DISCUSSION

In this study, we examined the gene organization of a 70 631 bp DNA region derived from the FN-catabolic linear plasmid pDBF1 of *Terrabacter* sp. DBF63 and found it to include β -ketoadipate pathway genes (Fig. 2). The β -ketoadipate pathway is widely distributed among diverse soil microbes, and plays an important role in the metabolism of naturally occurring aromatic compounds and environmental pollutants. The pathway genes are usually located chromosomally in proteobacteria (Harwood & Parales, 1996), although there are some exceptions such as pWW174 from *Acinetobacter calcoaceticus* (Winstanley *et al.*, 1987) and pSymB from *Sinorhizobium meliloti* (Finan *et al.*, 2001). Similarly, in actinobacteria the *pca* genes are not located on plasmids. For example, Iwagami *et al.* (2000) reported that the *pca* genes from *Streptomyces coelicolor* strain A3(2) are located on the genetically stable region of its chromosome. In addition, the *pca* genes are not located on the linear megaplasmid p1CP of *Rhodococcus opacus* strain ICP (König *et al.*, 2004) nor on linear plasmids (e.g. pRHL1) of *Rhodococcus* sp. strain RHA1 (Patrauchan *et al.*, 2005). Hence, the underlying importance of the pathway for the metabolism of aromatic compounds that are produced abundantly by plants may make the pathway genes typically

chromosomally encoded. In the present study, however, we discovered that strain DBF63 possessed the *pca* genes on the linear plasmid pDBF1. As we did not obtain any PCR products for *pca* genes (*pcaH*, *pcaG*, *pcaD* or *pcaC*) when total DNA from strain DBF63W, which was obtained by the continuous culture of the wild-type strain on nutrient broth and lacks the FN-degrading linear plasmids (Kasuga *et al.*, 1997; Nojiri *et al.*, 2002a), was used as a template (data not shown), this *pca* gene cluster was not located (duplicated) on the chromosome. Moreover, strain DBF63W did not grow with protocatechuate as a carbon source (data not shown), suggesting that the protocatechuate catabolic gene cluster is only located on the plasmid. Recently, other DF-degrading bacteria, *Terrabacter* sp. strain DFA1 and *Janibacter* sp. strain DFA10, both of which possess *dbfA1A2* gene homologues, were isolated, and the gene organizations of the regions flanking the *dbfA1A2* gene homologues were analysed. Sequence comparisons of these DNA regions among strains DBF63, DFA1 and DFA10 showed that the organizations of both *pht* and *dbf-fln* gene clusters were completely conserved (Noumura *et al.*, 2004). These results strongly suggest that pDBF1 of strain DBF63 is a transmissible plasmid, and may play an important role in the distribution of FN-degradation genes. Together with the fact that many IS-like elements and transposase genes are flanked with the *pca* gene cluster, it suggests that the β -ketoadipate pathway genes may be acquired via transposition events and subsequently distributed among actinobacteria via the horizontal transfer of plasmids.

Eulberg *et al.* (1998) first found a *pcaL* gene that encoded a merged enzyme with β -ketoadipate enol-lactone hydrolase and γ -carboxymuconolactone decarboxylase in the rhodococcal *pca* gene cluster, although these two enzymes are encoded by separate genes, *pcaD* and *pcaC*, respectively, in most proteobacteria (Buchan *et al.*, 2004). In addition, Iwagami *et al.* (2000) identified the *pcaL* gene homologue in the streptomycete *pca* gene cluster (strain 2065). Recent genome sequence analyses also show that *S. coelicolor* strain A3(2) (SCO939128), *Streptomyces avermitilis* MA-4680 (AP005027.1) and *Corynebacterium glutamicum* strain ATCC13032 (BX927155) possess the *pcaL* gene. This fusion of the two proteins into one is predicted to be beneficial to the stabilization of the enzyme itself, and to the efficient delivery of substrates and products by minimizing the distance between the active sites of each protein (Eulberg *et al.*, 1998). Furthermore, it is hypothesized that the separate *pcaD* and *pcaC* gene arrangement of proteobacteria may be more ancient, and that the presence of a fused *pcaL* gene is a Gram-positive trait (Eulberg *et al.*, 1998; Iwagami *et al.*, 2000). However, in this study we found that pDBF1 carried the separate *pcaD* and *pcaC* genes that had not been found in the *pca* gene cluster of actinobacteria (Fig. 3). The *pcaD* gene was expressed in both FN- and protocatechuate-induced DBF63 cells (Supplementary Fig. 1), and the PcaD enzyme exhibited its activity as a single protein, indicating that PcaD functions during FN or protocatechuate degradation by strain DBF63. Therefore, the presence of the *pcaL* gene is not quite a Gram-positive trait. Considering that the

Terrabacter sp. DBF63

Streptomyces coelicolor A3(2)

Rhodococcus opacus 1CP

Acinetobacter sp. ADP1

Fig. 3. Gene organization of *pca* gene clusters from actinobacteria [*Terrabacter* sp. strain DBF63 (accession number AP008980), *S. coelicolor* strain A3(2) (AL079355) and *Rhodococcus opacus* strain 1CP (AF003947)]. As a comparison, a *pca* gene organization from proteobacteria [*Acinetobacter* sp. ADP1 (L05770)] is shown under the dotted line. Putative regulatory genes (*pcaR*, *pcaU* and SCO6704) are indicated by grey shading. Gene designations are given in Supplementary Table 1. *pcaK* is a proposed aromatic-acid-transport gene. The organization of *pca* gene clusters from *S. coelicolor* strain A3(2) and *Streptomyces* sp. strain 2065 were identical (Iwagami *et al.*, 2000). Also, the *pca* gene organizations of *Rhodococcus opacus* strain 1CP and *Rhodococcus* sp. strain RHA1 were identical (Patrauchan *et al.*, 2005).

merged *pcaL* gene also exists in *Ralstonia metallidurans* (a β -proteobacteria) and *Caulobacter crescentus* (an α -proteobacteria) within the *pca* gene cluster (Jimenez *et al.*, 2002), whether the Pca enzyme arrangement of actinobacteria is more evolved than that of proteobacteria may be irrelevant.

The structure of the *pca* gene cluster from strain DBF63 was compared with that from *S. coelicolor* strain A3(2) and *Rhodococcus opacus* strain 1CP (Fig. 3). The gene organizations from *pcaH* to *pcaL* (*pcaD* and *pcaC* in strain DBF63) were highly similar among these strains. However, the locations of *pcaF*, *pcaIJ* and *pcaR* genes were markedly different. In strain DBF63, *pcaFIJ* was located at the end of the cluster, whereas in strain A3(2), *pcaIJF* was located upstream of *pcaH*. In these two species, both structural gene clusters *pcaHGBDCFIJ* and *pcaIJFHGBL* seemed to be located in a single operon, and the transcriptional regulatory gene, *pcaR*, was present upstream of the respective operons (Fig. 3). In contrast, *pcaIJ* of strain 1CP was transcribed divergently from other *pca* genes, and interestingly, *pcaR* seemed to be cotranscribed with the preceding *pcaHGBL* genes as a single operon because *pcaR* overlapped *pcaL* by 4 bp (Eulberg *et al.*, 1998; Fig. 3). These differences in gene arrangement affect the regulation of *pca* genes. In strain DBF63, the expression of the *pcaR* gene is probably not elicited by protocatechuate (Supplementary Fig. 1a), and is likely to be constitutive. However, the expression of the *pcaR* gene in strain 1CP, which may be cotranscribed with the structural genes, is likely to be induced. We have not yet established a procedure for gene knockouts in *Terrabacter* sp. strain DBF63, but if this can be achieved, the detailed transcriptional mode of not only *pca* genes, but also both *dbf-fln* and *pht* genes in concert with the putative corresponding regulators (PcaR, FlnR and PhtR, respectively), can be performed. It would be interesting to compare the regulatory systems among *pca* gene clusters in strains DBF63, A3(2) and 1CP.

In conclusion, this report is believed to be the first to (i) elucidate all catabolic pathway genes for the conversion of FN to TCA cycle intermediates, (ii) find the genes encoding the protocatechuate branch of the β -ketoadipate pathway on a linear plasmid, and (iii) find separate *pcaD* and *pcaC* genes, but not the merged *pcaL* gene, within the *pca* gene cluster of an actinobacteria species. The study of further examples of actinobacterial *pca* gene clusters, and comparisons of both the organization and regulatory systems of these *pca* operons may illuminate an evolutionary trait in actinobacteria that could be a consequence of the distinctive selection pressures faced by organisms maintaining the β -ketoadipate pathway.

ACKNOWLEDGEMENTS

We thank Dr Donna Parke of Yale University for providing pARO523. This work was partially supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences (PROBRAIN) in Japan.

REFERENCES

- Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3420.
- Buchan, A., Neidle, E. L. & Moran, M. A. (2004). Diverse organization of genes of the β -ketoadipate pathway in members of the marine *Rosebacter* lineage. *Appl Environ Microbiol* **70**, 1658–1668.
- Cerniglia, C. E. (1992). Biodegradation of polycyclic aromatic hydrocarbons. *Biodegradation* **3**, 351–368.
- Dagley, S., Evans, W. C. & Ribbons, D. W. (1960). New pathways in the oxidative metabolism of aromatic compounds by microorganisms. *Nature* **188**, 560–566.
- Eaton, R. W. (2001). Plasmid-encoded phthalate catabolic pathway in *Arthrobacter keyseri* 12B. *J Bacteriol* **183**, 3689–3703.
- Eckstein, T. M., Inamine, J. M., Lambert, M. L. & Belisle, J. T. (2000). A genetic mechanism for deletion of the *ser2* gene cluster and formation of rough morphological variants of *Mycobacterium avium*. *J Bacteriol* **182**, 6177–6182.
- Eulberg, D., Lakner, S., Golovleva, L. A. & Schlömann, M. (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.
- Finan, T. M., Weidner, S., Wong, K. & 9 other authors (2001). The complete sequence of the 1,683-kb pSymB megaplasmid from the N_2 -fixing endosymbiont *Sinorhizobium meliloti*. *Proc Natl Acad Sci U S A* **98**, 9889–9894.
- Grifoll, M., Selifonov, S. A. & Chapman, P. J. (1995a). Transformation of substituted fluorenes and fluorene analogs by *Pseudomonas* sp. strain F274. *Appl Environ Microbiol* **61**, 3490–3493.
- Grifoll, M., Selifonov, S. A., Gatlin, C. V. & Chapman, P. J. (1995b). Actions of a versatile fluorene-degrading bacterial isolate on polycyclic aromatic compounds. *Appl Environ Microbiol* **61**, 3711–3723.
- Habe, H. & Omori, T. (2003). Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria. *Biosci Biotechnol Biochem* **67**, 225–243.
- Habe, H., Miyakoshi, M., Chung, J.-S., Kasuga, K., Yoshida, T., Nojiri, H. & Omori, T. (2003). Phthalate catabolic gene cluster is linked to the angular dioxygenase gene in *Terrabacter* sp. strain DBF63. *Appl Microbiol Biotechnol* **61**, 44–54.
- Habe, H., Chung, J.-S., Kato, H., Ayabe, Y., Kasuga, K., Yoshida, T., Nojiri, H., Yamane, H. & Omori, T. (2004). Characterization of the upper pathway genes for fluorene metabolism in *Terrabacter* sp. strain DBF63. *J Bacteriol* **186**, 5938–5944.
- Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557–580.
- Harwood, C. S. & Parales, R. E. (1996). The β -ketoadipate pathway and the biology of self identity. *Annu Rev Microbiol* **50**, 553–590.
- Iwagami, S. G., Yang, K. & Davies, J. (2000). Characterization of the protocatechuic acid catabolic gene cluster from *Streptomyces* sp. strain 2065. *Appl Environ Microbiol* **66**, 1499–1508.
- Jensen, T. J. & Hites, R. A. (1983). Aromatic diesel emissions as a function of engine conditions. *Anal Chem* **55**, 594–599.
- Jimenez, J. I., Minambres, B., Garcia, J. L. & Diaz, E. (2002). Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. *Environ Microbiol* **4**, 824–841.
- Kanally, R. A. & Harayama, S. (2000). Biodegradation of high-molecular-weight polycyclic aromatic hydrocarbons by bacteria. *J Bacteriol* **182**, 2059–2067.

- Kasuga, K., Nojiri, H., Yamane, H., Kodama, T. & Omori, T. (1997). Cloning and characterization of the genes involved in the degradation of dibenzofuran by *Terrabacter* sp. strain DBF63. *J Ferment Bioeng* **84**, 387–399.
- Kasuga, K., Habe, H., Chung, J.-S., Yoshida, T., Nojiri, H., Yamane, H. & Omori, T. (2001). Isolation and characterization of the genes encoding a novel oxygenase component of angular dioxygenase from the gram-positive dibenzofuran-degrader *Terrabacter* sp. strain DBF63. *Biochem Biophys Res Commun* **283**, 195–204.
- Keith, L. H. & Telliard, W. A. (1979). Priority pollutants. I. A perspective view. *Environ Sci Technol* **13**, 416–423.
- König, C., Eulberg, D., Gröning, J., Lakner, S., Seibert, V., Kaschabek, S. R. & Schlömann, M. (2004). A linear megaplasmid, p1CP, carrying the genes for chlorocatechol catabolism of *Rhodococcus opacus* 1CP. *Microbiology* **150**, 3075–3087.
- Monna, L., Omori, T. & Kodama, T. (1993). Microbial degradation of dibenzofuran, fluorene, and dibenzo-*p*-dioxin by *Staphylococcus auriculans* DBF63. *Appl Environ Microbiol* **59**, 285–289.
- Nojiri, H., Kamakura, M., Urata, M., Tanaka, T., Chung, J.-S., Takemura, T., Yoshida, T., Habe, H. & Omori, T. (2002a). Dioxin catabolic genes are dispersed on the *Terrabacter* sp. DBF63 genome. *Biochem Biophys Res Commun* **296**, 233–240.
- Nojiri, H., Maeda, K., Sekiguchi, H., Urata, M., Shintani, M., Yoshida, T., Habe, H. & Omori, T. (2002b). Organization and transcriptional characterization of catechol degradation genes involved in carbazole degradation by *Pseudomonas resinovorans* strain CA10. *Biosci Biotechnol Biochem* **66**, 897–901.
- Noumura, T., Habe, H., Widada, J., Chung, J.-S., Yoshida, T., Nojiri, H. & Omori, T. (2004). Genetic characterization of the dibenzofuran-degrading Actinobacteria carrying the *dbfA1A2* gene homologues isolated from activated sludge. *FEMS Microbiol Lett* **239**, 147–155.
- Orville, A. M., Lipscomb, J. D. & Ohlendorf, D. H. (1997). Crystal structures of substrate and substrate analog complexes of proto-catechuate 3,4-dioxygenase: endogenous Fe^{3+} ligand displacement in response to substrate binding. *Biochemistry* **36**, 10052–10066.
- Parke, D. (1995). Supraoperonic clustering of *pca* genes for catabolism of the phenolic compound protocatechuate in *Agrobacterium tumefaciens*. *J Bacteriol* **177**, 3808–3817.
- Parke, D. (1997). Acquisition, reorganization, and merger of genes: novel management of the β -ketoadipate pathway in *Agrobacterium tumefaciens*. *FEMS Microbiol Lett* **146**, 3–12.
- Parke, D., D'Argenio, D. A. & Ornston, L. N. (2000). Bacteria are not what they eat: that is why they are so diverse. *J Bacteriol* **182**, 257–263.
- Pathak, D., Ashley, G. & Ollis, D. (1991). Thiol protease-like active site found in the enzyme dienelactone hydrolase: localization using biochemical, genetic and structural tools. *Proteins* **9**, 267–279.
- Patrauchan, M. A., Florizone, C., Dosanjh, M., Mohn, W. W., Davies, J. & Eltis, L. D. (2005). Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. *J Bacteriol* **187**, 4050–4063.
- Pinyakong, O., Habe, H. & Omori, T. (2003). The unique aromatic catabolic genes in sphingomonads degrading polycyclic aromatic hydrocarbons. *J Gen Appl Microbiol* **49**, 1–19.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Shine, J. & Dalgarno, L. (1975). Determination of cistron specificity in bacterial ribosomes. *Nature* **254**, 34–38.
- Shuttleworth, K. L. & Cerniglia, C. E. (1995). Environmental aspects of PAH biodegradation. *Appl Biochem Biotechnol* **54**, 291–302.
- Stecker, C., Johann, A., Herzberg, C., Averhoff, B. & Gottschalk, G. (2003). Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. *J Bacteriol* **185**, 5269–5274.
- Vetting, M. W., D'Argenio, D. A., Ornston, L. N. & Ohlendorf, D. H. (2000). Structures of *Acinetobacter* strain ADP1 protocatechuate 3,4-dioxygenase at 2.2 Å resolution: implications for the mechanism of an intradiol dioxygenase. *Biochemistry* **39**, 7943–7955.
- Winstanley, C., Taylor, S. C. & Williams, P. A. (1987). pWW174: a large plasmid from *Acinetobacter calcoaceticus* encoding benzene catabolism by the β -ketoadipate pathway. *Mol Microbiol* **1**, 219–227.
- Yamasaki, M., Redenbach, M. & Kinashi, H. (2001). Integrated structures of the linear plasmid SCP1 in two bidirectional donor strains of *Streptomyces coelicolor* A3(2). *Mol Gen Genet* **264**, 634–642.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.