Characterization of the *pcbE* Gene Encoding 2-Hydroxypenta-2,4-Dienoate Hydratase in *Pseudomonas* sp. DJ-12

Jong-Chul Lim[1,2], Jeongrai Lee[1], Jeong Duk Jang[2], Jai-Yun Lim[2], Kyung Rak Min[1], Chi-Kyung Kim[2], and Youngsoo Kim[1]

[1]*College of Pharmacy, Chungbuk National University, Cheongju 361-763, Korea and* [2]*School of Life Sciences, Chungbuk National University, Cheongju 361-763, Korea*

(Received Jaunary 5, 2000)

Nucleotide sequence extending 2,3-dihydroxybiphenyl 1,2-dioxygenase gene (*pcbC*) and 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase gene (*pcbD*) of *Pseudomonas* sp. DJ-12 was previously analyzed and the two genes were present in the order of *pcbD-pcbC* preceded by a promoter from *Pseudomonas* sp. DJ-12. In this study, a 3.8-kb nucleotide sequence located downstream of the *pcbC* gene was analyzed to have three open reading frames (ORFs) that are designated as *orf1, pcbE* and *orf2* genes. All of the ORFs were preceded by each ribosome-binding sequence of 5-GGAXA-3 (X=G or A). However, no promoter-like sequence and transcription terminator sequence were found in the analyzed region, downstream of *pcbC* gene. Therefore, the gene cluster appeared to be present in the order of *pcbD-pcbC-orf1-pcbE-orf2* as an operon, which is unique organization characterized so far in biphenyl- and PCB-degrading bacteria. The *orf1* gene was composed of 1,224 base pairs which can encode a polypeptide of molecular weight 44,950 containing 405 amino acid residues. A deduced amino acid sequence of the *orf1* gene product exhibited 21-33% identity with those of indole dioxygenase and phenol hydroxylase components. The *pcbE* gene was composed of 783 base pairs encoding 2-hydroxypenta-2,4-dienoate hydratase involved in the 4-chlorobiphenyl catabolism. The *orf2* gene was composed of 1,017 base pairs encoding a polypeptide of molecular weight 37,378 containing 338 amino acid residues. A deduced amino acid sequence of the *orf2* gene product exhibited 31% identity with that of a nitrilotriacetate monooxygenase component.

Key words: *Pseudomonas* sp. DJ-12, 4-Chorobiphenyl Catabolism, *pcbE* Gene, Unique Gene Cluster, Nucleotide Sequence

INTRODUCTION

Polychlorinated biphenyls (PCB) are recalcitrant aromatics that are widely distributed throughout the environment by their use in past as dielectric fluids, plasticizers, and hydraulic fluids (Hutzinger *et al.*, 1974). Since the first report of PCB as the environmental pollutants appeared 3 decades ago, there is much interest in the biodegradation of this class of aromatic compounds to understand the metabolic fate of PCB in the environment and to develop the remediation process of PCB-contaminated areas.

Several bacterial strains that can degrade PCB have been isolated from the environment (Ahmed and Focht, 1973; Bedard *et al.*, 1997; Bopp, 1986; Furukawa and Chakraberty, 1982; Furukawa *et al.*, 1978; Kim, *et al.*, 1987). Most of the PCB-degrading bacteria are *Pseudomonas, Alcaligenes, Achromobacter,* and *Rhodococcus* strains. Microbial degradation of PCB is mediated by the same enzymes that are involved in the catabolic pathway of biphenyl. Catabolic pathway of biphenyl is initiated by insertion of two atoms of oxygen at carbon positions 2 and 3 of one of the aromatic rings, which is catalyzed by biphenyl dioxygenase (Fig. 1) (Abramowicz, 1990). The resulting biphenyl dihydrodiol is converted to 2,3-dihydroxybiphenyl by a dihydrodiol dehydrogenase, and then the dihydroxylated ring is cleaved at *meta* position to form 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate by 2,3-dihydroxybiphenyl 1,2-

Correspondence to: Youngsoo Kim, College of Pharmacy, Chungbuk National University, Cheongju 361-763, Korea
E-mail: youngsoo@cbucc.chungbuk.ac.kr

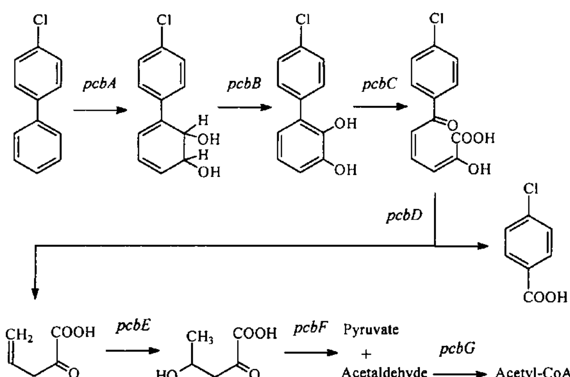

Fig. 1. A catabolic pathway of 4-chlorobiphenyl in *Pseudomonas* sp. DJ-12. The enzymes are biphenyl 1,2-dioxygenase (*pcbA*), a dihydrodiol dehydrogenase (*pcbB*), 2,3-dihydroxybiphenyl 1,2-dioxygenase (*pcbC*), 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase (*pcbD*), 2-hydroxypenta-2,4-dienoate hydratase (*pcbE*), 2-oxo-4-hydroxyvalerate aldolase (*pcbF*), and acetaldehyde dehydrogenase (*pcbG*).

dioxygenase. The *meta* cleavage compound is metabolized to benzoate and 2-hydroxypenta-2,4-dienoate by a hydrolase. It has been reported that some of PCB-degrading strains accumulate chlorinated benzoates as dead-end products. These kinds of bacterial strains utilize 2-hydroxypenta-2,4-dienoate as carbon and energy sources for growth, which is produced from the non-chlorinated aromatic ring (Kikuchi *et al.*, 1994; Kim *et al.*, 1996a). The 2-hydroxypenta-2,4-dienoate is converted to 2-oxo-4-hydroxyvalerate by a hydratase, and then to acetyl-CoA by sequential activities of aldolase and dehydrogenase.

The gene cluster encoding enzymes involved in the upper catabolic pathway of PCB to form corresponding chlorobenzoates and 2-hydroxypenta-2,4-dienoate has been cloned from several strains, and designated as *bphABCD, cbpABCD* or *pcbABCD* (Furukawa *et al.*, 1989; Hayase *et al.*, 1990; Hofer *et al.*, 1993; Khan and Walia, 1991; Kim *et al.*, 1996b; Kimbara *et al.*, 1989; Masai *et al.*, 1995; Mondello, 1989; Seeger *et al.*, 1995; Taira *et al.*, 1992) However, genes responsible for catabolism of chlorobenzoates and 2-hydroxypenta-2,4-dienoate in PCB-degrading bacteria have been yet studied extensively.

A 6.4-kb *Eco*RI fragment including 2,3-dihydroxybiphenyl 1,2-dioxygenase gene (*pcbC*) and 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase gene (*pcbD*) was previously cloned from chromosomal DNA of *Pseudomonas* sp. DJ-12, a soil bacterium that can grow in biphenyl or 4-chlorobiphenyl as the sole carbon and energy source (Kim *et al.*, 1994). The two genes were sequenced and appeared to be present in the order of *pcbD-pcbC* with a putative promoter (Kim, *et al.*, 1996b).

In this study, a 3.8-kb nucleotide sequence located downstream of the *pcbC* gene from *Pseudomonas* sp. DJ-12 has been determined, where three open reading frames (ORFs) were identified and designated as *orf1*, *pcbE*, and *orf2* genes. However, no promoter-like sequence and transcription terminator sequence were found in the region analyzed. Therefore the three ORFs seemed to be translatable in the same direction with *pcbDC* genes as an operon, where *pcbE* gene specifies 2-hydroxypenta-2,4-dienoate hydratase.

MATERIALS AND METHODS

Bacterial strains and plasmids

Strains and plasmids used and prepared in this study are described in Table I. *E. coli* HB101 was used as the recipient strain of recombinant plasmids, and grown in LB medium (10 g tryptone, 5 g yeast extract and 10 g NaCl per liter). For antibiotic selection, ampicillin was supplemented to the medium with 50 μg/ml as the final concentration.

DNA restriction, ligation and transformation

Plasmid was isolated by the alkali lysis method (Sambrook *et al.*, 1989). DNA fragmentation with restriction endonuclease and ligation of DNA fragments were accomplished according to the conditions recomended by the supplier (Boehringer Mannheim). DNA transformation was accomplished by the calcium chloride method (Sambrook *et al.*, 1989).

Gel electrophoresis and nucleotide sequencing

DNA was resolved on 0.7% or 1% agarose gel with TAE buffer (40 mM Tris-acetate and 2 mM EDTA) by electrophoresis, stained with 0.5 μg/ml ethidium bromide, and visualized by UV irradiation (Sambrook *et al.*, 1989). Nucleotide sequencing was carried out by using an Applied Biosystems DNA Sequencer (Pharmacia Biotech) with a dideoxy-chain termination kit from the same company. Nucleotide sequences obtained were analyzed by using softwares of DNASIS, PROSIS, and Clustal V.

RESULTS

pCU1 and its subclones for nucleotide sequencing

pCU1 isolated previously from the genomic library of *Pseudomonas* sp. DJ-12 is a recombinant plasmid of a 6.4-kb *Eco*RI fragment of DJ-12 inserted into the same site of pBluescript SK(+) (Kim *et al.*, 1994). Nucleotide sequence extending 2,3-dihydroxybiphenyl 1,2-dioxygenase gene (*pcbC*) and 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase gene (*pcbD*) of *Pseudomonas* sp.

Table I. Bacterial strains and plasmids used and prepared in this study

Strain or plasmid	Description	Source
Strains		
E. coli HB101	*supE44 hsdS58*($r_B^- m_B^-$) *recA13 ara-14 proA2 lacY1 galK2 respL20 xyl-5 mtl-1*	Stratagene Ltd
Pseudomonas sp. DJ-12	A soil isolate which can grow in 4-chlorobiphenyl or biphenyl as the sole carbon and energy source	Kim *et al.*, 1987
Plasmids		
pBluescript SK(+)	A cloning vector carrying *lacZ* promoter	Stratagene Ltd
pCU1	A plasmid clone selected from the genomic library of DJ-12, a 6.4-kb *EcoRI* fragment of the DJ-12 inserted into the same site of SK(+)	Kim *et al.*, 1994
pCNU1165	A 5.4-kb *Not*I fragment, deletion derivative of pCU1 lacking a *Not*1 fragment	This study
pCNU1166	A 4.3-kb *Apa*I fragment, deletion derivative of pCNU1165 lacking a *Apa*I fragment	This study
pCNU1167	A 3.7-kb *Hin*dIII fragment, deletion derivative of pCNU1165 lacking 3 *Hin*dIII fragments	This study
pCNU1168	A 6.0-kb *Sac*I fragment, deletion derivative of pCU1 lacking 2 *Sac*I fragments	This study
pCNU1169	A 3.3-kb *Pst*I fragment, deletion derivative of pCU1 lacking 5 *Pst*I fragments	This study
pCNU1170	A 4.8-kb *Hin*dIII fragment, deletion derivative of pCU1 lacking 5 *Hin*dIII fragments	This study
pCNU1171	A 3.5-kb *Hin*dIII fragment, deletion derivative of pCNU1168 lacking 4 *Hin*dIII fragments	This study
pCNU1172	A 3.3-kb *Sal*I fragment, deletion derivative of pCNU1165 lacking a *Sal*I fragment	This study
pCNU1173	A 2.3-kb *Sal*I-*Eco*RI fragment of pCU1 inserted into the same sites of SK(+)	This study
pCNU1174	An 0.6-kb *Not*I-*Sac*I fragment of pCNU1168 inserted into the same sites of SK(+)	This study
pCNU1175	An 1.6-kb *Pst*I-*Sac*I fragment of pCNU1168 inserted into the same sites of SK(+)	This study
pCNU1176	A 3.8-kb *Pst*I fragment, deletion derivative of pCNU1173 lacking a *Pst*I fragment	This study
pCNU1177	A 8.2-kb *Apa*I fragment, deletion derivative of pCU1 lacking a *Apa*I fragment	This study
pCNU1178	A 6.0-kb *Sal*I fragment, deletion derivative of pCU1 lacking 2 *Sal*I fragments	This study
pCNU1179	A 5.8-kb *Kpn*I fragment, deletion derivative of pCU1 lacking a *Kpn*I fragment	This study
pCNU1180	A 4.1-kb ApaI fragment, deletion derivative of pCNU1173 lacking a *Apa*I fragment	This study
pCNU1181	An 0.7-kb *Hin*dIII-*Eco*RI fragment of pCNU1173 inserted into the same sites of SK(+)	This study
pCNU1182	A 3.3-kb *Sac*I fragment, deletion derivative of pCNU1178 lacking a *Sac*I fragment	This study
pCNU1183	A 3.3-kb *Pst*I fragment, deletion derivative of pCNU1179 lacking 2 *Pst*I fragments	This study
pCNU1184	A 3.3-kb *Pst*I fragment, deletion derivative of pCNU1180 lacking a *Pst*I fragment	This study
pCNU1185	An 1.9-kb *Not*I-*Sal*I fragment of pCNU1177 inserted into the same sites of SK(+)	This study
pCNU1186	A 3.4-kb *Sac*I fragment, deletion derivative of pCNU1185 lacking a *Sac*I fragment	This study
pCNU1187	An 0.6-kb *Pst*I-*Cla*I fragment of pCNU1179 inserted into the same sites of SK(+)	This study

DJ-12 was previously analyzed and the two genes were present in the order of *pcbD-pcbC* as an operon (Kim *et al.*, 1996b). To analyze the remaining nucleotide sequence located downstream of the *pcbC* gene of *Pseudomonas* sp. DJ-12, a variety of subclones containing defined restriction fragments within the region have been constructed as described in Table I.

Nucleotide sequence of *orf1-pcbE-orf2* gene cluster

A 3.8-kb nucleotide sequence located downstream of the *pcbC* gene of *Pseudomonas* sp. DJ-12 is represented in Fig. 2. Three ORFs were identified, and designated as *orf1*, *pcbE*, and *orf2* genes which are linked tandemly. However, no promoter-like sequence and transcription terminator sequence were found in the sequenced region.

The *orf1* gene was started with ATG codon at position 1 and terminated with TGA codon at position 1224. A purine-rich region of 5-GGAGA-3, complementary to the sequence near the 3 end of 16S rRNA of *P. aeruginosa*, regarding as the ribosome-binding sequence (RBS) for *orf1* gene was identified at positions -11 to -7. ORFs for *orf1* and *pcbC* genes were overlapped at positions 1 to 13. The G + C content of *orf1* gene was 57%.

The *pcbE* gene was composed of 783 base pairs extending from ATG initiation codon at position 1310 to TAG termination codon at position 2092, and its G + C content was 57%. An intergenic sequence with 82 base pairs was identified between *orf1* and *pcbE* genes, where a RBS of 5-GGAAA-3 at positions 1302 to 1306 is present.

The *orf2* gene, located immediately downstream

Position	Sequence
-90	CAG TGA CTG GGG CCA TGC ACG TCC TGG CCA AGC CGT GCG TAA GCA
-45	CTT CGA CCT CTC TGT ATC TGT CCC TAT CTA TCA ***AGGAGA*** ATC CAA
	RBS
1	ATG TCT ACC GTG ACC GAC AAG CCT TTC GCA GCG ACT GCG ATC CCT
	Met Ser Thr Val Thr Asp Lys Pro Phe Ala Ala Thr Ala Ile Pro
46	TCC GTT GAA GAA CTC GTG CAA CGC GCT AGA GAT CTG TTG CCC ATG
	Ser Val Glu Glu Leu Val Gln Arg Ala Arg Asp Leu Leu Pro Met
91	CTG AAG GAA AAG GCT GCT TCC GTG GAA GCG AAC CGC ATG GTG TCC
	Leu Lys Glu Lys Ala Ala Ser Val Glu Ala Asn Arg Met Val Ser
136	AAG GAG ACC ATT CAG GCT TTC GTG GAT GCT GGC TTC TTC AAG ATC
	Lys Glu Thr Ile Gln Ala Phe Val Asp Ala Gly Phe Phe Lys Ile
181	CTG CAG GAA AAA CGT TGG GGT GGC TGG GAG CTA GAC CCG TGG TTT
	Leu Gln Glu Lys Arg Trp Gly Gly Trp Glu Leu Asp Pro Trp Phe
226	TGG CGC GTG CTC ATG GAG CTC GGC CGT GGC TGC CCG GCC AGC TGC
	Trp Arg Val Leu Met Glu Leu Gly Arg Gly Cys Pro Ala Ser Cys
271	TGG AAC ATG ATG ATT CTC GGC GTG CAT CAG TGG GAA TTT GGT CAC
	Trp Asn Met Met Ile Leu Gly Val His Gln Trp Glu Phe Gly His
316	ATG GAC CCC CGC GCT GCT GAA GAT GTT TGG GGC TCG GAT AAC ACC
	Met Asp Pro Arg Ala Ala Glu Asp Val Trp Gly Ser Asp Asn Thr
361	ACC ATC ATT GCT TCT TCG TAC GGT CCG ACA GGT CGA TGC GAA AAG
	Thr Ile Ile Ala Ser Ser Tyr Gly Pro Thr Gly Arg Cys Glu Lys
406	GTG GAC GGC GGC TAC AAG CTC AGT GGC CAA TGG AAG TGC TCT AGC
	Val Asp Gly Gly Tyr Lys Leu Ser Gly Gln Trp Lys Cys Ser Ser
451	GGT ACC GAT CAT GGT CAA TGG GCC TTC GTC GGC GGT CTG GTG CAT
	Gly Thr Asp His Gly Gln Trp Ala Phe Val Gly Gly Leu Val His
496	GAC GAG AAC GGT GTT GCC GTC GAC CGT ATC GCG CTT CTA GTG CCA
	Asp Glu Asn Gly Val Ala Val Asp Arg Ile Ala Leu Leu Val Pro
541	CGC ACT CAT TAC ACA ATT ATC GAT GAC TGG TAT GTG TTC GGT CTC
	Arg Thr His Tyr Thr Ile Ile Asp Asp Trp Tyr Val Phe Gly Leu
586	TGC GGT ACG GGC AGC AAG AGC TTG GAG ATG AAG GAC GTT TTC GTT
	Cys Gly Thr Gly Ser Lys Ser Leu Glu Met Lys Asp Val Phe Val
631	CCC GAA TAT CGC ACT CAC AGC CTG TCG GCC TAC CAG TTG AAC GAT
	Pro Glu Tyr Arg Thr His Ser Leu Ser Ala Tyr Gln Leu Asn Asp
676	CGC CCA TCG TCG TAT CTT CAC CCG TTC GGC ACC ATC TTC TGC GGC
	Arg Pro Ser Ser Tyr Leu His Pro Phe Gly Thr Ile Phe Cys Gly
721	TCT GTG TCG GCT GCA ATC GCC GGT TTT GCT CAA GGT GCG ATC GAC
	Ser Val Ser Ala Ala Ile Ala Gly Phe Ala Gln Gly Ala Ile Asp
766	ACC TAT ATC GAG CAG ATG AGC GGT CGT ACA CTC AGT GCG AAT GCG
	Thr Tyr Ile Glu Gln Met Ser Gly Arg Thr Leu Ser Ala Asn Ala
811	AGC ATG AAG GCC GCC AAC GTC AAC CCG GCG GTG CGT GAT CGC TTA
	Ser Met Lys Ala Ala Asn Val Asn Pro Ala Val Arg Asp Arg Leu
856	GGC CAT GCT GTG GCA AAG GTG CGC AGC TGC CGA GCT CGC CTG CTA
	Gly His Ala Val Ala Lys Val Arg Ser Cys Arg Ala Arg Leu Leu
901	TAC CTG ATG CAA GAG TCG GCT GAG TAT GTG TAC CGC CGC GAG TTG
	Tyr Leu Met Gln Glu Ser Ala Glu Tyr Val Tyr Arg Arg Glu Leu
946	GTG CCC GAA GAC ATG CGT GTG GCC ATT CCT CCG ACA TCT CCC GGT
	Val Pro Glu Asp Met Arg Val Ala ile Pro Pro Thr Ser Pro Gly
991	GTG GGT TAT GAA TGC TCG GAA GCC GTA ATG AAG CTG TTC CCC GCA
	Val Gly Tyr Glu Cys Ser Glu Ala Val Met Lys Leu Phe Pro Ala
1036	CTA GGA GCA CGC GGT CTG TAT TCC GAC AAT CCG ATG CAG CGC TTT
	Leu Gly Ala Arg Gly Leu Tyr Ser Asp Asn Pro Met Gln Arg Phe
1081	CTG CGC GAC ATT CTG GCA GGC TCC AAC CAT ATG ACC CAG AAC ATC
	Leu Arg Asp Ile Leu Ala Gly Ser Asn His Met Thr Gln Asn Ile

Fig. 2. Nucleotide sequence of downstream region of *pcbC* gene in *Pseudomonas* sp. DJ-12. Nucleotide sequence of the 3-end of *pcbC* gene and putative ribosome-binding sequences (RBS) for each open reading frame (ORF) are underlined. The termination codon of each ORF is indicated by***.

1126 GAC GAC TCC GCA GGC TTT ACA GGT GCA CGT CTG CTC GGC ACC GAT
Asp Asp Ser Ala Gly Phe Thr Gly Ala Arg Leu Leu Gly Thr Asp

1171 CTG CCG CCA ACG GCT TAC GGT ATC AAG CGC ACG GTG CCT GCT GAA
Leu Pro Pro Thr Ala Tyr Gly Ile Lys Arg Lys Val Pro Ala Glu

1216 AAG CGT TGA GTT ATT TTT TCT CGT CCC GAG ACA CCT GCA TCC AAG
Lys Arg ***

1261 CGG TGT TAT CGG GCG AGA AGG ATG AGC CAG CTTCGAGCCAA***GGAAA***
RBS

1307 AAC ATG CAG AGT GAA AAA ATT GTG AAA GCG GCG GAG ACT CTC AGG
Met Gln Ser Glu Lys Ile Val Lys Ala Ala Glu Thr Leu Arg

1352 CAG TCT CGT GCA ACG CTT ACT GCC GTA GAC AAG GTG TCT GTA ACC
Gln Ser Arg Ala Thr Leu Thr Ala Val Asp Lys Val Ser Val Thr

1397 CAT GGC ATT GGT TCG GTT GAT GAA GCT TAC ATG GTC TCG GAC CTC
His Gly Ile Gly Ser Val Asp Glu Ala Tyr Met Val Ser Asp Leu

1442 AAC CGC CAA TTC AAG GAG GCA CAG GGC GGC CGC GTT GTG GGC AAG
Asn Arg Gln Phe Lys Glu Ala Gln Gly Gly Arg Val Val Gly Lys

1487 AAG ATC GGC CTA ACA TCC AAG GCG GTT CAG GCC CAG TTG GGC GTA
Lys Ile Gly Leu Thr Ser Lys Ala Val Gln Ala Gln Leu Gly Val

1532 GAT CAA CCT GAC TTC GGA TTC CTG TTT GAC GAC ATG GAG TTT TTG
Asp Gln Pro Asp Phe Gly Phe Leu Phe Asp Asp Met Glu Phe Leu

1577 AAT GGT CAG GTC GTT CCC GCA AAT CGT CTG ATT CAA CCC AAG GCT
Asn Gly Gln Val Val Pro Ala Asn Arg Leu Ile Gln Pro Lys Ala

1622 GAA GCA GAA ATT GCC TTT GTG GTG CAG CAG AGC ATG GGT GGA GAG
Glu Ala Glu Ile Ala Phe Val Val Gln Gln Ser Met Gly Gly Glu

1667 GTG CCC ACC TAC AGC GAA TTT TTG GCT GCA TCG GTT ATG CGG TGC
Val Pro Thr Tyr Ser Glu Phe Leu Ala Ala Ser Val Met Arg Cys

1712 CCC GCC ATC GAG ATT GTC GAC AGC GCG ATC GCC GGT TGG AAG ATC
Pro Ala Ile Glu Ile Val Asp Ser Ala Ile Ala Gly Trp Lys Ile

1757 ACC ATC ATC GAC ACG GTC GCC GAC AAC GCT TCT TCG GGT CTC TAT
Thr Ile Ile Asp Thr Val Ala Asp Asn Ala Ser Ser Gly Leu Tyr

1802 GTC CTC GGT GAT CAG CCA GTA CCT GTC GCG CAG TTG GCC CTG AGT
Val Leu Gly Asp Gln Pro Val Pro Val Ala Gln Leu Ala Leu Ser

1847 GAA CTG GGA ATG ACT CTC GAA AAG AAC GGC TCC ATG GTT TCC TCA
Glu Leu Gly Met Thr Leu Glu Lys Asn Gly Ser Met Val Ser Ser

1892 GGT GTC GGC TCA GCG TGC CTG GGG CAC CCG CTG CGC GCC GCG TAC
Gly Val Gly Ser Ala Cys Leu Gly His Pro Leu Arg Ala Ala Tyr

1937 TGG CTG GCT TGC GAA ATG ATT CGC CGT GGT CAT GGC CTA GCC GCC
Trp Leu Ala Cys Glu Met Ile Arg Arg Gly His Gly Leu Ala Ala

1982 GGT GAG GTG ATT CTC TCG GGC GCA CTA GGA CCA ATG GTT CCG ATT
Gly Glu Val Ile Leu Ser Gly Ala Leu Gly Pro Met Val Pro Ile

2027 CAG GCC GGT GAC AGG GTC GAG GCA CGC ATA CAG GGC CTG GGC TCG
Gln Ala Gly Asp Arg Val Glu Ala Arg Ile Gln Gly Leu Gly Ser

2072 GTG CAT TTC AGT TTG GGC TAG TCC GCT CCG TTT TTT AAG CTT CAG
Val His Phe Ser Leu Gly ***

2117 ACA ATT AAC AGA ACA TCG CAG AGC CAG TC***G GAGA***CA TAC ATG CAA
RBS Met Gln

2162 GAT CAG TCG ATC CAG ACA TTT GAC CCC AAA GCC TTC CGC GCA GCA
Asp Gln Ser Ile Gln Thr Phe Asp Pro Lys Ala Phe Arg Ala Ala

2207 CTC GGA AGC TTT GCC ACC GGG GTG ACT GTC ATC ACT ACC GTC GGA
Leu Gly Ser Phe Ala Thr Gly Val Thr Val Ile Thr Thr Va Gly

2252 CAA GAC GGA AAG CCT GTT GGT CTG ACG GCC AAC AGC TTC AAC TCT
Gln Asp Gly Lys Pro Val Gly Leu Thr Ala Asn Ser Phe Asn Ser

2297 GTG TCG CTG GAC CCG CCC TTG GTT CTT TGG AGT CTC GCG AAA AAG
Val Ser Leu Asp Pro Pro Leu Val Leu Trp Ser Leu Ala Lys Lys

2342 GCG TTC AGT TTG CAG GAT TTT GTC GCA GCT AAA CAT TGG GCA GTG
Ala Phe Ser Leu Gln Asp Phe Val Ala Ala Lys His Trp Ala Val

2387 CAT ATT CTG AGC GCT GAC CAG GAG CAT ATC TCC AAT CAG TTT GCG
His Ile Leu Ser Ala Asp Gln Glu His Ile Ser Asn Gln Phe Ala

Continued Fig. 2.

```
2432  CGA GGT GGT GCA GAC AAA TTC GCA GGT GTC GAG CCC GTA ACC TCC
      Arg Gly Gly Ala Asp Lys Phe Ala Gly Val Glu Pro Val Thr Ser
2477  GAG CAT GGG GTG CCA CTG CTG CAG AAC TGC GCA GCA CGG TTT GAA
      Glu His Gly Val Pro Leu Leu Gln Asn Cys Ala Ala Arg Phe Glu
2522  TGC ACC AGC ACC TTT CAG TAC GAA GGT GGT GAT CAC ATC ATT TTT
      Cys Thr Ser Thr Phe Gln Tyr Glu Gly Gly Asp His Ile Ile Phe
2567  GTT GGC GAA GTC AAG AAA TTT GAA CGC AAT GAT CAC TCG CCG CTG
      Val Gly Glu Val Lys Lys Phe Glu Arg Asn Asp His Ser Pro Leu
2612  GTC TTT CAT TCG GGT AGA TAT GCG CTG GCA TCC CTC AAA GAC AAT
      Val Phe His Ser Gly Arg Tyr Ala Leu Ala Ser Leu Lys Asp Asn
2657  GGC TTT AGT CCT CCC AAA CCT GAT TGC ACC GAG CCG CTC ACG TTC
      Gly Phe Ser Pro Pro Lys Pro Asp Cys Thr Glu Pro Leu Thr Phe
2702  CGC GCA GAC ATG CTG GGC TAT ATG GTG GGC CGG GCC CGC AAG AAT
      Arg Ala Asp Met Leu Gly Tyr Met Val Gly Arg Ala Arg Lys Asn
2747  TTC CTC GAC AGC ATG CGC ATC CAT CTG GAG AGC AAT GGA CTC AAC
      Phe Leu Asp Ser Met Arg Ile His Leu Glu Ser Asn Gly Leu Asn
2792  GAC TAC GAA TGG CGT CTG CTG ACC ATC ATC CTG ACG AAG AAA AGT
      Asp Tyr Glu Trp Arg Leu Leu Thr Ile Ile Leu Thr Lys Lys Ser
2837  CTG ACT GCA TCC ATG TTC GAG CGA TTC AAC AAA GAC GCT GCG CTG
      Leu Thr Ala Ser Met Phe Glu Arg Phe Asn Lys Asp Ala Ala Leu
2882  GAA ACC ATT CTG GCT ACG CTG GAC TCC TTG CAG CAA AAG GGC TGG
      Glu Thr Ile Leu Ala Thr Leu Asp Ser Leu Gln Gln Lys Gly Trp
2927  ATT GCC GTG TCT TCG GAT GGC GAG CAC GGA GAG GCG ACC TAC GGT
      Ile Ala Val Ser Ser Asp Gly Glu His Gly Glu Ala Thr Tyr Gly
2972  CTC ACG CAA AAG GGT GTC TAC GAC TCG GTT CGC CTG CTG GCC ATT
      Leu Thr Gln Lys Gly Val Tyr Asp Ser Val Arg Leu Leu Ala Ile
3017  GCA AAG TCC CAT GAG GAG CAT TTG ATG GAA AAA ATG GGC TAT GCG
      Ala Lys Ser His Glu Glu His Leu Met Glu Lys Met Gly Tyr Ala
3062  GAT GGA ATG CTG CTT AAA AGC TTG CTC AAT AAA TTT ATT CAG AAA
      Asp Gly Met Leu Leu Lys Ser Leu Leu Asn Lys Phe Ile Gln Lys
3107  ACG GAG GGA TCT GCC CCC AAC CTC TGG AGC GAT GAA TTT GTG TTG
      Thr Glu Gly Ser Ala Pro Asn Leu Trp Ser Asp Glu Phe Val Leu
3152  AAA GAA AAG GTG GTT CTG TAA GAG TGA GCA ACT AAT AAA AGA CGA
      Lys Glu Lys Val Val Leu ***
3197  GGC TAA GCA CCT CAA TAA ATA GAA GTA TTT CCT CAG GAG ACA AAA
3242  TGT CGA AAG TAT TCA CAT CGA AGG TAT TGA TTC CGG CAC TGT TGG
3287  GTT TGA GTG CGA GCC TGG CAC ATG CTC AAA ATC CCA ACC CAG TTC
3332  GAA TCG TCA TTG GTT TTG CGG CAG GAG GTA ATG GCG ACA TCA TTG
3377  CAC GCA TGC TGG CCA ACG AAT TGC GTC CCA TCA TGG GCC GCA ATG
3422  TGA TCG TGG AAA ACA AGC CAG GAG CTG GCG GGC GCC TTG CCG CGC
3467  TGC AGT TGA AGT CTT CTC CTG CTG ACG GAA GCT ACT ATC TAC TGG
3512  CTC CGG ATA GCT GGT CGA TTT TTC CGA CCA TCC TGC AGA CAG AAA
3557  GTC AGC TGC GCT ACA ACC TCC AAA AGG ACT TCG CGC CCG TCG CGC
3602  GTA TCG TAT CTT ATC CCC TTG GCC TGT TCG TGT CTG AAG GCG CTG
3647  GGG TCC ATA GTC TCA AGG AAT ACA TCG AGA AGG CCA AGG TGA ACC
3692  CCG CCA TTG CCT CTT ACG GTT CGT CAG GTG CAG GCA GCA TTA CTG
3737  AGT TTT TGG GCA TTG TGA TGT CCA AGG AAT TC
                                          EcoRI
```

Continued Fig. 2.

of *pcbE* gene, was started at position 2156 and terminated at position 3172. The G + C content of *orf2* gene was 53%. An intergenic sequence between *pcbE* and *orf2* genes was composed of 63 base pairs, where a RBS of 5-GGAGA-3 is identified at positions 2146 to 2150.

Comparison of the deduced amino acid sequences

The *orf1* gene can encode a polypeptide of predicted molecular weight 44,950 containing 405 amino acid residues (Fig. 2). The amino acid sequence of *orf1* gene product exhibited 21-33% identity with those of a hydroxylase (Orf11) from *Rhodococcus erythropolis* TA421, a pigment polypeptide (Rv309c) from *Mycobacterium tuberculosis* H37Rv, an indole dioxygenase component (BpfA) from *R. opacus* NCIB12038, and a phenol hydroxylase component (PheA) from *Bacillus stearothermophilus* BR219 (Table II).

The *pcbE* gene encodes a 2-hydroxypenta-2,4-dieno-

Table II. The homology of amino acid sequences of polypeptides encoded in *orf1, pcbE,* and *orf2* genes from *Pseudomonas* sp. DJ-12 with those of corresponding polypeptides from other bacterial strains.

ORF	Identity(%)		Strain (gene)	GenBank accession number
	AA	NT		
orf-1	33	53	*Rhodococcus erythropolis* TA421 (*orf11*)	D88013
	26	48	*Mycobacterium tuberculosis* H37Rv (*Rv3094c*)	Z95150
	25	50	*Rhodococcus opacus* NCIB12038 (*bpfA*)	AJ005688
	21	48	*Bacillus starothermophilus* BR219 (*pheA*)	U17960
pcbE	50	56	*Pseudomonas* sp. KKS102 (*bphE*)	D16407
	49	55	*Sphingomonas aromaticivorans* F199 (*xylJ*)	AF079317
	48	54	*P. putida* F1 (*todG*)	U09250
	47	58	*P. putida* CF600 (*dmpE*)	X60835
	47	53	*Pseudomonas* sp. DJ77 (*phnH*)	U97697
	46	56	*P. putida* mt-2 (*xylJ*)	M64747
	46	54	*E. coli* CS520 (*mhpD*)	Y0955
	37	52	*M. tuberculosis* H37Rv (*Rv3536c*)	Z82089
	31	49	*E. coli* ATCC11105 (*hpaH*)	Z37980
orf-2	31	49	*Chelatobacter heintzii* ATCC29600 (*ntaB*)	U39411

Polypeptides are a hydroxylase (*orf11*), pigment polypeptide (*Rv3094c*), indole dioxygenase component (*bpfA*), phenol hydroxylase component (*pheA*), 2-hydroxypenta-2,4-dienoic acid hydratase (*bphE, xylJ, todG, dmpE, phnH, mhpD*), 2-hydroxyhepta-2,4-diene-1,7-dioic acid hydratase (*Rv3536c, hpaH*), and nitrilotriacetate monooxygenase component (*ntaB*). Abbreviations are amino acid sequence (AA) and nucleotide sequence (NT).

ate hydratase with molecular weight 27,397 containing 260 amino acid residues. The deduced amino acid sequence of *pcbE* gene product exhibited 46-50% identity with those of 2-hydroxypenta-2,4-dienoate hydratases of BphE from *Pseudomonas* sp. KKS102, XylJ from *Sphingomonas aromaticivorans* F199, TodG from *P. putida* F1, DmpE from *P. putida* CF600, PhnH from *Pseudomonas* sp. DJ77, XylJ from *P. putida* mt-2, and MhpD from *E. coli* CS520, and did 31-37% identity with those of 2-hydroxyhepta-2,4-diene-1,7-dioate hydratases of Rv3536c from *M. tuberculosis* H37Rv and HpaH from *E. coli* ATCC11105 (Table II).

The *orf2* gene can encode a polypeptide of molecular weight 37,378 containing 338 amino acid residues. A deduced amino acid sequence of *orf2* gene product exhibited 31% identity with that of nitrilotriacetate monooxygenase component (NtaB) from *Chelatobacter heintzii* ATCC29600 (Table II).

Organization of the *pcb* gene cluster in *Pseudomonas* sp DJ-12

A promoter-like sequence (-35 and -10) was previously identified upstream of *pcbD* gene (Kim *et al.*, 1996b). The *pcbD* gene was located immediately upstream of the *pcbC* gene, where both of the genes can be translatable in the same direction as an operon (Kim et *al.*, 1996b). The 3.8-kb nucleotide sequence located downstream of the *pcbC* gene has revealed three ORFs

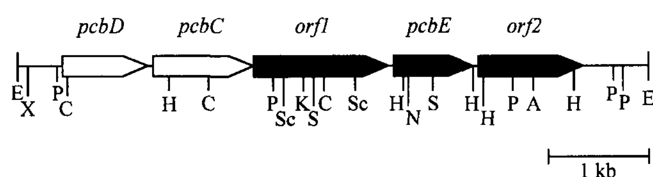

Fig. 3. Physical and genetic maps of recombinant plasmid pCU1. Open reading frames (ORFs) are 2-hydroxy-6-oxo-6phenylhexa-2,4-dienoate hydrolase (*pcbD*), 2,3-dihydroxybiphenyl 1,2-dioxygenase (*pcbC*), 2-hydroxypenta-2,4-dienoate hydratase (*pcbE*), and additional two ORFs (*orf1* and *orf2*). Restriction enzymes are *ApaI* (A), *ClaI* (C), *EcoRI* (E), *HindIII* (H), *KpnI* (K), *NotI* (N), *PstI* (P), *SacI* (Sc), *SalI* (S), and *XhoI* (X).

of *orf1-pcbE-orf2* genes (Fig. 3). No promoter-like sequence and transcription terminator sequence were found downstream region of the *pcbC* gene. Therefore, the gene cluster appeared to be present in the order of *pcbD-pcbC-orf1-pcbE-orf2* as an operon.

DISCUSSION

Microbial catabolism of PCB is initially proceeded to form corresponding chlorobenzoates and 2-hydroxypenta-2,4-dienoate by sequential activities of four enzymes encoded in *bphABCD, cbpABCD* or *pcbABCD* (Fig. 1) (Abramowicz, 1990; Furukawa et al., 1989; Hayase et *al.*, 1990; Hofer et *al.*, 1993; Khan and Walia,

1991; Kim *et al.*, 1996b; Kimbara *et al.*, 1989; Masai *et al.*, 1995; Mondello, 1989; Seeger *et al.*, 1995; Taira *et al.*, 1992). Some of biphenyl-degrading bacteria cannot degrade the chlorobenzoates any further, and thus the aromatic intermediates are accumulated during PCB catabolism.

Pseudomonas sp. DJ-12 is a soil isolate that can grow in biphenyl or 4-chlorobiphenyl as carbon and energy sources for growth (Kim *et al.*, 1987). The bacterial strain was known to degrade both 4-chlorobenzoate and 2-hydroxypenta-2,4-dienoate resulted from 4-chlorobiphenyl catabolism. Previously, we reported the nucleotide sequence extending a promoter followed by *pcbD* and *pcbC* genes of *Pseudomonas* sp. DJ-12 (Kim *et al.*, 1996b). In this study, a 3.8-kb nucleotide sequence located downstream of the *pcbC* gene in *Pseudomonas* sp. DJ-12 was determined, where three ORFs were identified, and designated as *orf1*, *pcbE* and *orf2* genes (Fig. 2). All of the ORFs were preceded by each RBS of 5-GGAXA-3 (X = G or A) which is complementary to the sequence near 3 end of 16S rRNA of *P. aeruginosa*. The G + C contents of the ORFs were *orf1* (57%) and *pcbE* (57%) and *orf2* (53%), which are slightly lower than those of other *Pseudomonas* strains (Normore, 1976).

A deduced amino acid sequence of *pcbE* gene product from *Pseudomonas* sp. DJ-12 exhibited 46-50% identity with those of 2-hydroxypenta-2,4-dienoate hydratases from *Pseudomonas*, *Sphingomonas*, and *Escherichia* strains, and did 31-37% identity with those of 2-hydroxyhepta-2,4-diene-1,7-dioate hydratases from *Mycobacterium* and *Escherichia* strains (Table II). Recently, a crystal structure of enoyl-CoA hydratase from rat liver mitochondria was resolved, where Glu 164 acts as catalytic acid donating a proton to the substrate (Engel *et al.*, 1996). An alignment of amino acid sequences of 2-hydroxypenta-2,4-dienoate hydratases from *Pseudomonas* sp. DJ-12 and enoyl-CoA hydratase from rat liver mitochondria has revealed that the catalytic Glu residue was, corresponding to Glu 107 in *Pseudomonas* sp. DJ-12, well conserved. The *orf1* gene was composed of 1,224 base pairs encoding a polypeptide of molecular weight 44,950, whose deduced amino acid sequence exhibited 21-33% identity with those of pigment polypeptide, indole dioxygenase and phenol hydroxylase components from other bacterial strains (Table II). The *orf2* gene, encoding a polypeptide of molecular weight 37,378 containing 338 amino acid residues, exhibited 31% identity at amino acid sequence with nitrilotriacetate monooxygenase component from *Chelatobacter heintzii* ATCC29600 (Table II). At this moment, importance of the polypeptides encoded in *orf1* and *orf2* genes in the 4-chlorobiphenyl or other aromatics catabolism of *Pseudomonas* sp. DJ-12 is not clarified, which will be elucidated in a future study.

No promoter-like sequence and transcription terminator sequence were found in the region downstream of *pcbC* gene of *Pseudomonas* sp. DJ-12. Therefore, the gene cluster appeared to be present in the order of *pcbD-pcbC-orf1-pcbE-orf2* as an operon (Fig. 3). Nucleotide sequence upstream of the *pcbD* gene did not exhibit significant homology with that of biphenyl dioxygenase gene (*pcbA*) or a dihydrodiol deydrogenase gene (*pcbB*) (Kim *et al.*, 1996b). This means that *pcbA* and *pcbB* genes would be separated in a distance from *pcbDCE* gene cluster, and transcribed by using different promoters in *Pseudomonas* sp. DJ-12. The *cbpDCB* gene cluster in *P. putida* OU83 is separated by 3.0 kb from *cbpA* gene (Khan and Walia, 1991). The *bphABC* and *bphDEF* gene clusters are located on different plasmids in *Rhodococcus* sp. RHA1 (Masai *et al.*, 1995). However, *bphABCD* genes in *P. putida* KF715, *bphABCEGFD* genes in *P. pseudoalcaligenes* KF707 and *Pseudomonas* sp. LB400, and *bphEGFABCD* genes in *Pseudomonas* sp. KKS102 are known to present as the clusters (Furukawa *et al.*, 1989; Hayase *et al.*, 1990; Hofer *et al.*, 1993; Kikuchi *et al.*, 1994; Taira *et al.*, 1992). Thus, organization of the genes responsible for 4-chlorobiphenyl catabolism in *Pseudomonas* sp. DJ-12 is unique.

ACKNOWLEDGEMENTS

This work was financially supported by a grant (KOSEF 1999-2-202-004-3) from Korea of Science and Engineering Foundation.

REFERENCES

Abramowicz, D. A., Aerobic and anaerobic biodegradation of PCBs: a review. *Biotechnology*, 10, 241-251 (1990).

Ahmed, M., and Focht, D. D., Degradation of polychlorinated biphenyls by two species of *Achromobacter*. *Can. J. Microbiol.*, 19, 47-52 (1973).

Bedard, D. L., Wagner, R. E., Brennan, M. J., Haberl, M. L., and Brown, Jr. J. F., Extensive degradation of Aroclors and environmentally transformed polychlorinated biphenyls by *Alcaligenes eutrophus* H850. *Appl. Environ. Microbiol.*, 53, 1094-1102 (1997).

Bopp, L. H., Degradation of highly chlorinated PCBs by *Pseudomonas* strain LB400. *J. Ind. Microbiol.*, 1, 23-29 (1986).

Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K., and Wierenga, R. W., Crystal structure of enoyl-coenzyme A (CoA) hydratase at 2.5 A resolution: a spiral fold defines the CoA-binding pocket. *EMBO J.*, 15, 5135-5145 (1996).

Furukawa, K., and Chakraberty, A. M., Involvement of plasmids in total degradation of chlorinated biphenyls.

Appl. Environ. Microbiol., 44, 543-548 (1982).

Furukawa, K., Matsumura, F., and Tonomura, K., *Alcaligenes* and *Acinectobacter* strains capable of degrading polchlorinated biphenyls. *Agric. Biol. Chem.,* 42, 543-548 (1978).

Furukawa, K., Hayase, N., Taira, K., and Tomizuka, N., Molecular relationship of chromosomal genes encoding biphenyl/polychlorinated biphenyl catabolism: some soil bacteria possess a highly conserved *bph* operon. *J. Bacteriol.,* 171, 5467-5472 (1989).

Hayase, N., Taira, K., and Furukawa, K., *Pseudomonas putida* KF715 *bphABCD* operon encoding biphenyl and polychlorinated biphenyl degradation: cloning, analysis, and expression in soil bacteria. *J. Bacteriol.,* 172, 1160-1164 (1990).

Hofer, B., Eltis, L. D., Dowling, D. N., and Timmis, K. N., Genetic analysis of a *Pseudomonas* locus encoding a pathway of biphenyl/polychlorinated biphenyl degradation. *Gene,* 130, 47-55 (1993).

Hutzinger, O., Safe, S., and Zitko, V., The Chemistry of PCB's. CRC Press, Inc., Cleveland (1974).

Khan, A. A., and Walia, S. K., Expression, localization, and functional analysis of polychlorinated biphenyl degradation genes *cbpABCD* of *Pseudomonas putida. Appl. Environ. Microbiol.,* 57, 1325-1332 (1991).

Kikuchi, Y., Yasukochi, Y., Nagata, Y., Fukuda, M., and Takagi, M., Nucleotide sequence and functional analysis of the *meta*-cleavage pathway involved in biphenyl and polychlorinated biphenyl degradation in *Pseudomonas* sp. KKS102. *J. Bacteriol.,* 176, 4269-4276 (1994).

Kim, C.-K., Chae, J. C., and Kim, Y.-C., Isolation and characterization of bacteria degrading chlorinated aromatic hydrocarbons. *Kor. J. Microbiol.,* 25, 122-128 (1987).

Kim, C.-K., Kim, E., Chae, J. C., and Kim, Y., Structure of the *pcbC* encoding 2,3-dihydroxybiphenyl dioxygenase of *Pseudomonas* sp. P20. *Biochem. Biophys. Res. Commun.,* 226, 15-20 (1996a).

Kim, C.-K., Sung, T.-K., Nam, J. H., Kim, Y.-C., and Lee, J. K., Cloning and expression of *pcbCD* genes in *Escherichia coli* from *Pseudomonas* sp. DJ-12. *Kor. J. Microbiol.,* 32, 40-46 (1994).

Kim, E., Kim, Y., and Kim, C.-K., Genetic structure of the genes encoding 2,3-dihydroxybiphenyl 1,2-dioxygenase and 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoic acid hydrolase from biphenyl- and 4-chlorobiphenyl-degrading *Pseudomonas* sp. strain DJ-12. *Appl. Environ. Microbiol.,* 62, 262-265 (1996b).

Kimbara, K., Hashimoto, T., Fukuda, M., Koana, T., Takagi, M., Oishi, M., and Yano, K., Cloning and sequencing of two tandem genes involved in degradation of 2,3-dihydroxybiphenyl to benzoic acid in the polychlorinated biphenyl-degrading soil bacterium *Pseudomonas* sp. strain KKS102. *J. Bacteriol.,* 171, 2740-2747 (1989).

Masai, E., Yamada, A., Healy, J. M., Hatta, T., Kimbara, K., Fukuda, M., and Yano, K., Characterization of biphenyl catabolic genes of Gram-positive polychlorinated biphenyl degradaer *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol.,* 61, 2079-2085 (1995).

Mondello, F. J., Cloning and expression in *Escherichia coli* of *Pseudomonas* strain LB400 genes encoding polychlorinated biphenyl degradation. *J. Bacteriol.,* 171, 1725-1732 (1989).

Normore, W. M., *Handbook of Biochemistry and Molecular Biology,* Fasman, G. D. (ed.), Vol. 3, p. 65-235, The Chemical Rubber Co., Cleveland (1976).

Sambrook, J., Fritsch, E. F., and Manniatis, T., *Molecular Cloning: A Laboratory Manual,* Cold Spring Harbor Laboratory, Cold Spring Harbor, New York (1989).

Seeger, M., Timmis, K. N., and Hofer, B., Conversion of chlorobiphenyl into phenylhexadienoates and benzoates by the enzymes of the upper pathway for polychlorobiphenyl degradation encoded by *bph* locus of *Pseudomonas* sp. strain LB400. *Appl. Environ. Microbiol.,* 61, 2654-2658 (1995).

Taira, K., Hirose, J., Hayashida, S., and Furukawa, K., Analysis of *bph* operon from the polchlorinated biphenyl-degrading strain of *Pseudomonas pseudoalcaligenes* KF707. *J. Biol. Chem.,* 267, 4844-4853 (1992).

Yagi, O., and Sudo, R., Degradation of polychlorinated biphenyls by microorganisms. *J. Water Pollut. Control. Fed.,* 52, 1035-1043 (1980).