## Identification and Characterization of Genes Encoding Polycyclic Aromatic Hydrocarbon Dioxygenase and Polycyclic Aromatic Hydrocarbon Dihydrodiol Dehydrogenase in *Pseudomonas putida* OUS82

NOBORU TAKIZAWA,<sup>1</sup>\* NAOFUMI KAIDA,<sup>1</sup> SHIN TORIGOE,<sup>1</sup> TSUYOSHI MORITANI,<sup>1</sup> TAKASHI SAWADA,<sup>1</sup> SACHIKO SATOH,<sup>2</sup> and HOHZOH KIYOHARA<sup>1,2</sup>

Biotechnology Laboratory, Department of Applied Chemistry, Faculty of Engineering,<sup>1</sup> and Central Institute for Research,<sup>2</sup> Okayama University of Science, Okayama 700, Japan

Received 12 October 1993/Accepted 25 January 1994

Naphthalene and phenanthrene are transformed by enzymes encoded by the *pah* gene cluster of *Pseudomonas putida* OUS82. The *pahA* and *pahB* genes, which encode the first and second enzymes, dioxygenase and *cis*-dihydrodiol dehydrogenase, respectively, were identified and sequenced. The DNA sequences showed that *pahA* and *pahB* were clustered and that *pahA* consisted of four cistrons, *pahA<sub>a</sub>*, *pahA<sub>b</sub>*, *pahA<sub>d</sub>*, and *pahA<sub>d</sub>*, which encode ferredoxin reductase, ferredoxin, and two subunits of the iron-sulfur protein, respectively.

*Pseudomonas putida* OUS82 can assimilate naphthalene and phenanthrene as its sole carbon sources. The strain converts naphthalene and phenanthrene to salicylate and 1-hydroxy-2naphthoate, respectively, by a shared catabolic pathway (the upper pathway; Fig. 1). Salicylate and 1-hydroxy-2-naphthoate are further degraded by other catabolic enzymes. The enzymes in the upper pathway have broad substrate specificities, and various polycyclic aromatic hydrocarbons other than naphthalene and phenanthrene are oxidized by a high-density suspension of OUS82 cells (9).

Previously, we cloned the gene cluster encoding the enzymes of the upper pathway and named it *pah* (polycyclic aromatic hydrocarbon; 9). The *pah* region strongly hybridized to a corresponding region of plasmid NAH7 of *P. putida* G7, which degrades naphthalene (4). All recombinant plasmids carrying *pahA* have 6.5- and 3.0-kb SalI fragments. The two fragments were seen to be necessary for the dioxygenase phenotype (PahA). A restriction endonuclease map of a region in the fragments resembles that of the *nahA* region of NAH7 and pDTG1 in *P. putida* G7 and NCIB 9816-4, which degrade naphthalene (2, 4, 23). The *pahA* gene was expected to be in that region.

Here, we describe the identification and characterization of the *pahA* and *pahB* genes, which encode dioxygenase PahA, which is the first enzyme of the pathway and converts polycyclic aromatic hydrocarbon (PAH) to the corresponding *cis*-dihydrodiol, and dehydrogenase PahB, the second enzyme of the pathway, which converts the product of PahA to the corresponding diol.

*P. putida* OUS8211 (*trp-82 \Delta pah-821*), a derivative of strain OUS82 that is defective in naphthalene and phenanthrene utilization, and plasmid pDI1, which carries the *pahAB* gene cluster, were described previously (9). Plasmid NAH7 was described elsewhere (4, 5). *Escherichia coli* JM109 and plasmid pUC119 were described by Yanisch-Perron et al. (21) and Vieira and Messing (20), respectively. Broad-host-range vector

pTS1210 (Apr Kmr mob<sup>+</sup>), a derivative of pSa, was a gift from A. Nakazawa, (Yamaguchi University) (11). Helper plasmid pRK2013 (Km<sup>r</sup> tra<sup>+</sup>) was a gift from M. Fukuda (Nagaoka University of Technology) (7). tac promoter expression vector pKK223-3 (1) was purchased from Pharmacia LKB Biotechnology, Uppsala, Sweden. The rich medium (LB) and the minimal medium used in this study were described previously (9). SMT medium was the minimal medium plus 0.3% disodium succinate and 30 µg of tryptophan per ml. Restriction enzymes and T4 DNA ligase were obtained from Toyobo Co., Ltd., Osaka, Japan, and Nippon Gene Co., Ltd., Toyama, Japan. An in vitro packaging kit was obtained from Amersham International plc, Buckinghamshire, United Kingdom. DNA and amino acid sequence similarities were analyzed with DNASIS-Mac software (version 2.0; Hitachi Software Engineering Co., Ltd., Yokohama, Japan).

Nucleotide sequencing and characterization of pahA and pahB. Previously, the pahA gene was suggested to be in a 4.1-kb EcoRI-HpaI fragment in pDI1 (Fig. 2). The pahB gene was thought to flank the pahA gene because genes that encode catabolic enzymes are often clustered (13, 19, 22). We sequenced the nucleotides of a 6-kb region between the EcoRV and SacI sites (Fig. 2). DNA was sequenced by the dideoxychain termination procedure (15) with alkali-denatured plasmid DNA and biotinylated oligonucleotides (New England BioLabs, Beverly, Mass.) as the primer and a Sequenase DNA sequencing kit (United States Biochemical Corp., Cleveland, Ohio) or a Bca-Best DNA sequencing kit (Takara Shuzo Co., Ltd., Kyoto, Japan). Electrophoresis was done with  $0.5 \times TBE$ buffer (1 M Tris base, 83 mM boric acid, 1 mM disodium EDTA) in the upper chamber (anode) and a mixture of  $1 \times$ TBE and 0.5 volume of 3 M sodium acetate in the lower chamber (cathode) (17). The result of electrophoresis was electroblotted onto a Biodyne A (Pall Biosupport Div., East Hills, N.Y.) nylon membrane with an electroblotting apparatus (NB-1600; Nihon Eido Co., Ltd., Tokyo, Japan) at 100 mA (constant current) for 15 min with  $0.2 \times$  TBE buffer and detected with a Uniplex chemiluminescence detection subkit (Millipore Corp., Bedford, Mass.). Five complete open reading frames (ORF1 to ORF5) and the 5'-terminal sequence of ORF6, each preceded by an E. coli consensus ribosome-

<sup>\*</sup> Corresponding author. Mailing address: Department of Applied Chemistry, Faculty of Engineering, Okayama University of Science, 1-1 Ridai-cho, Okayama 700, Japan. Phone: +81-86-252-3161, Ext. 4635. Fax: +81-86-252-6891.



FIG. 1. Possible upper pathway for degradation of PAHs in *P. putida* OUS82.

binding site, were found in that region (Fig. 3). ORF1 to ORF5 spanned 987, 315, 1,350, 585, and 780 nucleotides, respectively, and encoded polypeptides with deduced molecular masses of 35.6, 11.5, 49.3, 22.9, and 27.5 kDa, respectively. In this region, there was a pair of putative -10 and -35 promoter sequences similar to the *E. coli* consensus sequence upstream from the initiation codons of ORF1. The nucleotides and deduced amino acid residues of ORF1 to ORF4 were very similar to those of the naphthalene dioxygenase (NDO) genes of *P. putida* G7, NCIB 9816-4 (18), and NCIB 9816 (Table 1) (10). Rieske-type iron-sulfur centers (that is, [2Fe-2S]-binding sites)



FIG. 2. Cloned DNA regions (thick lines) of *P. putida* OUS8211 in pDI1, pKTS1, and pNA1. The thin lines represent vector plasmids. The open boxes indicate the gene locations determined by nucleotide sequencing.  $P_{tac}$  and the arrow indicate the *tac* promoter and the direction of transcription, respectively. MCS, multiple cloning site in pUC119.

were found in the deduced amino acid sequences of ORF2 (Cys-45 to His-47 and Cys-64 to His-67) and ORF3 (Cys-81 to His-83 and Cys-101 to His-104). A sequence similar to an NahR-binding sequence (-70 sequence) (16, 24) was found upstream from ORF1.

These results indicate that the *pahA* gene encodes an NDO-type enzyme with four components: ferredoxin reductase, ferredoxin, and the iron-sulfur protein large and small subunits (6). We defined ORF1 to ORF4 as  $pahA_a$ ,  $pahA_b$ ,  $pahA_c$ , and  $pahA_d$ , respectively.

The protein predicted by ORF5 (27.4 kDa) was similar in molecular mass to the naphthalene dihydrodiol (NDD) dehydrogenase NahB (25.5 kDa) and the toluene dihydrodiol dehydrogenase TodD (27.0 kDa) (12, 14), and its primary structure was similar to the structures of the toluene dihydrodiol, biphenyl dihydrodiol, and benzene dihydrodiol dehydrogenases (TodD, [40.0% similarity], BphB [39.0% similarity], and protein 5 [40.9% similarity], respectively) listed in the GenBank DNA data base. The 0.5-kb *Eco*RI-*StuI* region in ORF5 was strongly hybridized to a 4.7-kb *SalI* fragment of NAH7 containing the *nahB* gene (5; data not shown). These results suggest that ORF5 encodes dihydrodiol dehydrogenase, and we defined ORF5 as *pahB*.

Conversion of naphthalene by P. putida OUS8211 carrying pahA. A 7.4-kb EcoRI fragment of pDI1 harboring pahA was cut out and inserted into pTS1210 to obtain pKTS1 (Fig. 2). A high-density suspension of OUS8211 cells carrying pKTS1 was exposed to naphthalene, and metabolites were extracted from the supernatant and analyzed by high-pressure liquid chromatography (D-6100 three-dimensional chromatography system; Hitachi Ltd., Tokyo, Japan; column, Intersil ODS-2 [4.6 by 250 mm]; G-L Science, Tokyo, Japan; mobile phase, 60% acetonitrile containing 0.05% trifluoroacetic acid; pressure, 150 kg/ cm<sup>2</sup>; flow rate, 0.5 ml/min). One main metabolite was purified and crystallized. The metabolite was converted to 1-naphthol and 2-naphthol by boiling under acidic conditions. Because NDD is converted to naphthol under acidic conditions (8), the metabolite was probably NDD. The metabolite was further analyzed by <sup>13</sup>C and <sup>1</sup>H nuclear magnetic resonance analyses. Peaks representing two hydroxylated carbons in positions 1 and 2 were detected in the <sup>13</sup>C nuclear magnetic resonance spectrum (8 68.03 and 70.85), and the <sup>1</sup>H nuclear magnetic resonance spectra of the metabolite ( $\delta$  4.68 [d, J = 5.2 Hz, 1H, H-1],  $\delta$  4.38 [dd, J = 5.2 and 4.0 Hz, 1H, H-2],  $\delta$  6.05 [dd, J = 4.0 and 10 Hz, 1H, H-3],  $\delta$  6.53 [d, J = 10 Hz 1H, H-4],  $\delta$  2.04 and 2.38 [each s, 1H, OH],  $\delta$  7.10 to 7.55 [m]) agreed with those

1	AGGCAGTGATCTCTGCATTCTCCGATGGTGAGACATCACCTAGCAGCGTACTGCTGAGAGCGAGACTAGAGACTAGAGTCGGAAGTACCGGCCGTACGGCTCATAACCAAAACCTCAGTG Dah
1 121	n a h - d o x GTCGACCGTCGTCTTCCTGGANTACGCACCTCACCGTCGAGCCAGGCTGAATGGATGGACGGCCAGGCTGAGC TACGATTTTTTCAGTCTCCTCTATTCGACCTGCGTTGCGAACGTTCGATCGA
69 241	GTCCCAGCTATTTGGCCAGGAGTGGCATATGATCGATCTAGTCAAACCCGTTTTTGGTGAAAAGCTGCACCTGGTCACCATCGATCCGTGCTAACAGACGATACCAGTCGTACTTGATCT GACAACACCAGAGCGTTATCCTGCGGGTTCAGGCAGAGCCCAACCACATGACTACCTCCGGCAAAATGAGGGTCGTTACTAATTTTGAACGTCTTAAGCGATGGGGTTTTAGGTCGAA
189 361	CGTATCGCCATGCTCCTGCCGACGACAGTGTCAGTAGAGTAACCAGTTGTGGCGCGATCCAACACGGGTTTGGTGCCTTGGGCGGGTTTCACTATCGCCACCTCGCTATCAGGAACATTCC AGCTGCCCAGGCCTGCGGAACCTACCACATCGTAACCCGAGCATATTCGGCCGCGGCGGCTAAAAACACAGAAATGAGCGGGGTGACCCGATCGCCTTTGATCGATTCTCCGC
309 481	TTC AGGGTGGGTGGGTGCGTAATTTCTGTGAAAGGGGAGCCAGGTTATGAGTATTCACATTGGTGATAAACAACCTCACTTATGGCGTAATTGACATATAAACGTCGTATTCACGATTATTTACCAAAATAAAGCACCATCACGATCATTCACGATGTTTCACGATGATAAATTCAACTATGCCTTTATTGACAAATAAAAGCAACCATCACCATCACGACAAATAAAAGCAACCATCACGATGTTTCACGATGTGATAAATTCAACTATGCACTGACAATAAAAAGCAACCACCACCAACCA
429 601 1	$ \begin{array}{c} $ n = h a \\ matrix $ n = h \\ n = h \\ matrix $ n = h \\ n $
549 721 31	TTTCCTACAGTTGCTGGCGGTATGCGGGAACCTGTCGCTGCCGGGTTATGATGGCAGTGTCATTGATTCTGGGCGGAAAATGGCAATCAAACCTCACCGACAAGCAGTATGTGC TTTCCTACAGTTGTATGTCTGGGCGTTGCGGAACCTGCCGGGGTTACAGATGGTAGTATGGTAGTATTGATTCGGGGAAGCGGGTTACCACACCTCGTGGACAAGCAGTATGTGC I S Y S C M S G R C G T C R C R V T D G S V I D S G C L F E L V D E X V
669 841 71	TCGCCTGTCAGTCAGTACTACTGGCAATTGCGCTATCGAAGTCCCAGAAGCCGACGAAATTGTCACTCAC
789 961 111	ACGATATCCGTCGCTTACGCGTACGCCTCTCAAGCCCTTCGAGTTCTCACCCGGACAGTACGCGACACTGCAGTTCAGCCCTGAGCATGCGCGTCCGTATTCAATGGCAGGTTTGCCAG ACGATATCCGTCGCCTACGCGTACGCCTCGCCAAGCCCTTCGAGTTCTCACCCGGACAGTACGCGACATTGCAGTTCAGTCCTGAGCATGCGCGTCCGTATTCAATGGCAGGTCTGCCAG H D I R R L R V R L & K P F E F S P G Q Y A T L Q F S P E H A R P Y S M A G L P
909 1081 151	ATGACCAAGAATGGAGTTCCACATACGCAAGGTGCCGGGGTGGCGCGCTACGGAGTATGTTTTCGAACACGTCCGCGAAGGTACAAGCATCAAGTAGAGGGGCCCTTGGTACGGCT ATGACCAAGAAATGGAGTTCCACATACGCAAGGTGCCGGGGGGGG
1029 1201 191	ATCTACGTCAGAAGCACACCGGACCGATGCTGTGTGTGGGGGGGG
1149 1321 231	ATTICGGGGTGCGCAGTCAGCAGACCTCTACGACGCAGAGCGATTGCACAAACTCGCCGCTGACCACCCTCAACTGACCGACGCAGGTGATTGCAACGGGCCCGATTAATGAGGTC ATTICGGAGTGCGCAGTCAGCAAGACCTCTACGACGCAGGCGAGTGATCACCTCGCGGCTGATCACCTCAACTGACCGTAACACGGGTAATGCAAGGGCCCGATTAATGAGAGTC Y F G V R S Q Q D L Y D A B R L H S L A A D H P Q L T V H T V I A S G P I N E S
1269 1441 271	AGCGAGCCGGCCTAATTACCGATGTGATCGAAAAAGACATCCTTTCGCTGGCTG
1389 1561 311	TTGGAATATCACCCGAACATATTTATGCCGATGCCTTCTATCCCGGTGGGATCTGAATAGTTCCCGGCCATGCACCTTCGTCCATCGACAATTCAACAGGAAGACATTCAAATGAACGTAA TTGGAATATCACCGGAACATATTTATGCCGATGCCTTCTATCCCGGTGAAATGGTCCCCTTCCCCCACCTCGGTCCATTGAGGACTCATCAAGGAGGAGAATCTCAAATGACGCAA L G I S P E I I X A D A F Y P G E I $*$
1510 1681 1	
1630 1801 15	$\begin{array}{c} CTGAAGGTGACGTCCTCGGCGTGACTGTCGAGGGCAAGGAGCTGGCGCTGTATGAAGGTGAAGGCGAAATCTACGCTACCGACAACCTGTGCACGCATGGTTCCGCCCCCATGAGTGATCCCCCCCC$
1750 1921 55	$ \begin{array}{c} GGTTATCTCGAGGGGTAGAGAAATCGAATGCCCCTTGCATCAAGGTCGGTTTGACGTTTGCACAGGCAAAGCCCTGTGGCCACGGCCAGAACATCAAAACATATCCAGTCAAGATGGGTTGACGTTTGACGTTGGAGGGGAGAGCCCCTGGCCCCCGGGACAGAACATCAAAACATATCCAGTCAAGATGCAAGATGCAAGATGCAGAGAGGCCGCCCCGGGACAGAACAACAAAAACATATCCAGTCAAGATGCAAGATGCAAGATGCAGAGAGGCCCCTGGCCCCCGGGACAGAACAACAAAAACATATACCAGTCAAGATGCAAGATGCAAGATGCAGAGAGGCCCCTGGCCCCCGGGACAGAACAACAATATGCAGTCAAGATGCAAGATGCAAGATGCAAGACATCAAAACATATACCAGTCAAGATGCAAGACGTTGACGTTTGACGTTTGACGAGAGCCCCTGGCCCCCGGGACAGAACAACAATATACAATATGCAGTCAAGATGCAAGACATCAAAGCATAAAACATATACCAGTCAAGATGCAAGACGTTGACGTTGGACGCCCCGGGCCCCGGGACAGAACAACAAAAAAAA$
1870 2041 95	GAGAACCTGCGCGTAATGATTGAGTTGAGCTAAGAATTTT-AACAGGAGGCACCCCGGGCCCTAGAGCGTAATCACCCCCATTCCATCTTTTT-AGGTGAAAACATGAATTAAAAAC GAGAACCTGCGCGTAATGATTGATTGAATTTAAGCTGAGAATTTTTAATAGGCGGCGCCCCGGACCATAGAGCGTGATTATCCCCATTCCATCTTTTTTTAGGTGAAAACATGAATTACAAAAAC B N L R V M I D L S * RB M V K N ORF3 ( $pahA_{c}$ ;
1988 2161 6	AAAATCTTGGTAAGTGAATCTGGTCTGGACCAAAAGCACCTGATTCATGGCGATGAAGAACTTTTCCAACATGAAAACCATTTTTGCGCGGAACTGGCTTTTCTCACTCA
2108 2281 46	$\begin{array}{llllllllllllllllllllllllllllllllllll$
2228 2401 86	A A GACGCTGGTGGAGCGTGGAAGCCGGCAATGCCAAAGGTTTTGTTTG
2348 2521 126	GAGTCGCTCAATAAAAAATGTCTGGGGTTGAAAGAAGTCGCTCGC
2468 2641 166	$ \begin{array}{llllllllllllllllllllllllllllllllllll$
2588 2761 206	$ \begin{array}{cccc} GCATACCACGTGGGTTGGACGCACGCGTCTTCGCTCGGCGGGAGTCTATCTTCTTCTGTGGTGCTGGCAATGCGCCGCTACCACGTGGAAGGCGCAGGCTTGCAAATGACGTCCAAAGGCGCAGGCTTGCAAATGACGTCCGAGGCTGCAAGGCGCTGCAAGGCGGCTTGCAAATGACGTCCAAAGGCGAAGGCGCGGGCTGCGAGGGCTGCAAGGCGGCTGCAAATGACGTCCAAAGGCGCAGGCGGCTGCAAGGCGGCGGGCTGCAAAGGCGCGGGCTGCAAAGGCGCGGGCTGGCAAGGCGGCGGGCTGGCAAGGCGGCGGGCTGGCAAGGCGGCGGGGGGGG$
2648 2881 246	TACGGCAGCGGCATGGGTGTGTGTGTGTGGGACGGATATTCAGGTGCGATAGCGCAGACTTGGTTCCGGAATTGATGGCATTCGGAGGCGCAAAGCAGGAAAGGCTGAACAAAGAAATTGGC TACGGCAGCGGTATGGGTGTGTGTGTGGGACGGATATTCAGGCGTGCATAGCGCAGACTTGGTTCCGGAATTGATGGCATTCGGCGCGCGAAAGAAGGCTGAACAAAGAAATTGGC Y G S G M G V L W D G Y S G V H S A D L V P E L M A F G G A K Q E R L N K E I G
2768 3001 286	GATGTTCGCGCTCGGATTTATCGCAGCCACCTCAACTGCACCGTTTTCCCGAACAACAGCATGGACCTGCTGGCCGGGTGTTTTCAAAGTATGGAACCCGATCGACGCAAACACCACCGAG GATGTTCGCGCCCGGATTTATCGCAGCCACCTCAACTGCACCGTTTTCCCGAACAACAGCGTGGTGGCTGACCTGCTCGGGGTGTTTTCAAAGTATGGAACCCGATCGACGCAAACACCACCGAG D V R A R I Y R S H L N C T V F P N N S V L T C S G V F K V W N P I D A N T T E
2888 3121 326	GTCTGGACCTACGCCATTGTCGAAAAAGACATGCCTGAGGATCTCAAGCGCCGCTTGGCCGACCTGTTCAGGGAACGTTCGGGCCTGCTGGGAAAGCGACGACGACAATGACAAT GTCTGGACCTACGCCATTGTCGAAAAAGACATGCCCGAGGATCTCAAGCGCCGCGTTGGCCGACGGGTTCAGCGAACGTTCGGGCCTGCTGGCTATGGCGACGACGACGAC V W T Y A I V E K D M P E D L K R R L A D & V Q R T F G P A G F W E S D D N D N
3008	ATGGAAACAGCATCGCAAAACGGCAAAAAATATCAATCAA

3128 3361 406	TCGGCGATCGGCGAGACCAGTTATCGTGGTTTCTACCGGGCTTACCAGGCACACGTCAGCAGCTCCAACTGGGCTGAGTTCGAGCATGCCTCTAGTACTTAGGATACTGAACTACGAAG TCGGCGATCGGCGAGACCAGTTATCGTGGTTTCTACCGGGCTTACCAGGGCACACGTCAGCAGCTCCAACTGGGCTGAGTTCGAGGATGCCTCTAGTACTTGGCATACCGAACTGACGAAC S A I G E T S Y R G F Y R A Y Q A H V S S S N W A E F E D A S S T W H T E L T K
3248 3481 446	$\begin{array}{cccc} &   \ naha_{d}(ndoC) \\ \text{ACTACTGATCGATACAGACGAGTCGATCATATCAATATCAAGAAGACAAGCTGGTTTCCGCCCACGACGCCCGAAGAGATTCTTCGTTTCTTCAATTGCCACGACTCTGCTTTGCACTACTGATCGCTAACAGACGAGTCGACCATGATGATCATCAAGAAGACAAGCTGGTATCCGCCCATGACGCCCGAAGAGATTTCTTCGTTTCTTCAATTGCCACGACTCGCCTTTGCT T D R * RBS M M I N T Q E D K L V S A H D A E E F L R F F N C H D S A L ORF4 (pahA_{d}; ISP small subunit) \\\end{array}$
3368 3601 31	$ \begin{array}{cccc} AACAAGAAGCACCACTACGCTGACCCAGGAAGCGCATTTGTTGGACATTCAGGCTACCGTGGTTAGAGGACCACGGGGGGGTCAGAGGTCATATCAGGTCATTTCACGCGACCACGGACGTGGGGCGCGGGGGGGCAGAGGGTCAATATCAGGTCATTCACGCGACGACGACGGGCGGG$
3488 3721 71	TEGEGGEGAGCTTCCAGAGCGTCGTTATAAGCTCAATGAAGCCATGAACGTTAACGGAAAATTTTCCAGCAACTGAAAGTTCGAGTCGAGCAACATGGATCGACCGCCAACAAACTGGAGCCAACAAACTGGAGCGACGACAACATGGAGCGACGACAACATGGAGCGACGACAACATGGAGCGACGAACATCGACGAGCGACGAACATCGACGAGCGACGAACATCGACGAGCGACAACATGGAGCGACGAACATGGAGCGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAACGGAACGGAACATGGAGCGAACATGGAGCGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGACGAACATGGAGCGAACATGGAGCGAACATGGAGCGAACATGGAGCGAACATGGAGGAACATGGAGGAACATGGAACGGAACGAAC
3608 3841 111	$ \begin{array}{c} GCCGGAAGCTGGGCTTTACTGGCTTTATCACCAACGTCCAGGCCGCAATGGACGTAAATGACGAAAGAGCTACTTCACATCGGCTCCAAAGGTCATTTCGCACCGGGCACGACGTGGGCAATGGACGAAGACGAACGTCGGCCTCCAACGTCCAACGTCGACGTCGGCCAATGGACGAAGGCGAATGGACGAAGACGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGAAGGCGCAATGGACGAAGGCGAATGGACGAACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGAAGGCGAATGGACGAAGGCGAATGGACGAAGGCGAATGGACGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGAAGGCGCAATGGACGAAGGCGAATGGACGAAGGCGAATGGACGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGACGTCGAAGGCGAATGGACGAAGGCGAATGGACGACGTGGCGAATGGACGACGTGGCCAATGGACGACGTCGACGTCGACGTGGCGAATGGACGACGTCGACGTGGACGTGGACGACGACGTCGACGTGGACGACGACGTCGACGTGGACGACGTGGCCGAATGGACGACGTGGACGACGTGGACGACGACGTGGACGACGTGGACGACGTGGACGACGTGGACGACGTGGACGACGTGGACGACGACGTGGACGACGACGACGACGACGACGACGACGACGACGACGAC$
3728 3961 151	$ \begin{array}{llllllllllllllllllllllllllllllllllll$
3848 4081 191	$ \begin{array}{c} do x B \\ \text{TOGTCTTTCTGTGATTCAGGATATCAAAATGGTCACTGCTACCGCCGGTCACCATTAATC-AAAGGGAATGTACGGGTATCGGCAATCAACAAGTCGTTTCGATAACCGGTGCACCATTAATCAAAAGGGGAATGTACGGGTATGGGCAATCAACAAGTCGTTTCAATAACCGGTGCCG \\ \text{TOGTCTTTCTGTGATCCGGTGACCACTTTTACAAAATGGTGACTGCTACCGCGGTCACCATTAATCAAAAGGGGAATGTACGGTGATGGGCAATCAACAAGTCGTTTCAATAACCGGTGCCCG \\ \text{M V F L } \\ \text{RBB } \\ \text{M G N Q V V S I T G A } \\ \text{ORF5 (pahB; dehydrogenase)} \end{array} $
3967 4201 13	GCTCAGGAATCGGTCTCGAACTGGTTCGGTCCTTTAAGTCGGCCGGTTATTACGTATCCGCTCTCGTACGAAACGAGGAGCAAGAGGCGCTTCTTTGCAAAGAGTTCAAGGACGCCACTCG GCTCAGGAATCGGTCTCGAACTGGTTCGATCCTTCAAGTCGGCCGGTTATTCCGTATCCGCTCTCGTACGAAACGAGGAGCGCTTCTTTGCAAAGAGTTCAAGGACGCACTCG G S G I G L E L V R S F K S A G Y C V S A L V R N E E Q E A L L C S E F K D A L
4087 4321 53	AGATTGTAGTGGGGGATGTCCGGGACCACGCAAATAAATGAGAAGCTGATAAAGCAAACAATCGATAGATTCGGTCATCTTGATTGTTTTATTGCAAATGCCGGTATCTGGGATTACATGC AGATCGTTGTGGGGGATGTCCGAGATCACGCAATAAATGAGAAGCTGATCAAGCAGACAATCGCTAGATTGGGTCATCTGGATTGTTTCATCGCAAATGCCGGTATTTGGGATTACATGC E I V V G D V R D H A I N E K L I K Q T I & R F G H L D C F I A N A G I W D Y M
4207 4441 93	TGAGCATCGAAGAGCCTTGGGAGAAAATATATCGAGCAGTTTTGACGAAATATTCGACATTAATGTCAAGAGCTATTTCAGTGGCGATCAGTGCCGCCCTGCCGGAACTGAAAAAGACTAACG TGAGCATCGAAGAGCCTTGGGAGAAAATCTCCCAGCAGTTTTGACGAAATATTCGACATCAATGTAAAGAGCTATTTCAGTGGCATCAGTGCCAGCTCTGCCGGAACTGAAAAAGACGAACG L S I E E P W E K I S S S F D E I F D I N V K S Y F S G I S A A L P E L K K T N
4327 4561 133	$ \begin{array}{llllllllllllllllllllllllllllllllllll$
4447 4681 173	$\label{eq:ccg} CCGAAGTTCGCGTGAACGCTGTTTCGCCGGGGGCCACCGTGACGTCTCTGTGCGGTCCCCGCGACCGCGGTTTCGACAAAATGCACATGAAAAGACATGCCCGGCATCGACGATATGATCACAAAATGCACATGAAAAATGCCCGGCATCGACGATATGATCACAAAATGCACATGAAAAATGCACATGACGATGACGATGACGATATGATCACAAAATGCACATGAAAAATGCACATGACGATGACGATGACGATATGATCACAAAATGCACATGAAAAATGCACATGACGATGACGATGACGATATGATCACAAAATGCACATGAAAATGCACATGACGATGACGATGACGATATGATCACAAAATGCACATGAAAAATGCACATGAACGATGCACGGATGACGATGATGATGATGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGACGATGATGATGATGACGATGACGATGATGATGATGATGACGATGATGATGATGATGACGATGATGATGATGATGATGATGATGATGATGATGATGATG$
4567 4801 213	AAGGTCTCACGCCCTTTGGGTTTGCAGCCAAGACGTGGTGGGCACCCTATTTGTTGCTGGCTTCGCGAAAGCAAGGAAAATTCATCACCGGGCACCGTGATTAGCATTGATGGCG AAGGCCTCACTCCTCTTGGGTTTGCAGCCAAGACGTGGTGGAGCCCTATTGTTGCTGGCTTCGCGAAAGCAGGGAAAATTCATCACCGGGCACCGTGATTAGCATTGATGGCG K G L T P L G F A A K P E D V V E P Y L L A S R K Q G K F I T G T V I S I D G
4687 4921 253	GTATGGCGCTCGGTCGCAAGTGAGCTTGTAGCCGATCAGAAGTTATAGACACATTT-CAGGTGACGCCCCATGAAGACAAAACTGTTTATCAATAACGCCTGGATCGATTCTAGTGACCA   GTATGGCGCTCGGTCGCAAGTGAGCCTGCAGCCGATCAAAGGTTATAGACACTTTTTAGGTGACGCCCCATGAAGACAAAACTGTTTATCAACAACGCCTGGATCGATTCCACAGTGACCA   GTATGGCGCTCGGTCGCAAGTGAGCCGATCAAAGGTTATAGAAACATTTTTAGGTGACGCCCCATGAAGACAAAAACTGTTTATCAACAACGCCTGGATCGATTCCACAGTGACCA   G M A L G R K *   ORF6 (pah F?)
4806 5041 18	$ \begin{array}{c} GCAGACCTTCGAGCGCATACACCCCGTCAGCAGCGATGTGGTGACTGAGAGCGCAAACGCCACAGTGACGGACG$
4926 5161 58	$\begin{array}{llllllllllllllllllllllllllllllllllll$
5046 5281 98	$\begin{array}{c} \label{constraint} CGGATTCAACGTCCATCGCTATGTCTCCGAGAGGCTGCCTCGCTGCCTGC$
5166 5401 138	GCCGGTCGGACCGATCCTAAGCATCGTTCCATGGAACGGCACGCAGTGCTTGCGGCACGAGCCATCGCTTGTCCGCTGGTCGTGGCAACACTGTGGTGTTCAAAGGCTCGAATTTAG GCCGGTCGGCCCGATCCTGAGCATCGTCCCATGGAACGGCACCGCAGTGCTGGCGGCACGAGCCATCGCGTATCCGCTGGTCGTGGCCAACACGGTGGTGTTCAAAGGCTCTGAATTTAG P V G P I L S I V P W N G T A V L A A R A I A Y P L V C G N T V V F K G S E F S
5286 5521 178	TCCCGCGACGCATGCCCTGATCACCCAGTGCGGGGGGGGG
5406 5641 218	TGCCAAGGAGATCCGCCGCATCAACTTCACGGGTTCCACCCGCGTGGGCAGCAGCAGCAGAAAGCCGCGCAACACCTCCAAGCGCTGCTGCCTGGAGCTCGGCGGCAAGTCCCCCGCT CGCGAAGGAGATCCGCCGCATCAACTTCACGGGTTCAACCCGCGTGGGCAGCAGAAGGCCGCGCAAAGCCCGCGCAACACCTCAAGCGCTGCCTACTGGAGCTCGGCGGCAAGTCCCCCGCT A K E I R R I N F T G S T R V G S I I A Q K A A Q H L K R C L L E L G G K S P L
5526 5761 258	TATTGTTCTGGATGATGCAGACATCGATGCGGCGGTCAAGGCAGCGGTGTTCGGTAGCTTCCTGTTCCAAGGTCAGATCT TATTGTTCTGGATGACGCAAACATTGACGCGGGGGTCAAGGCAGCGGTGTTCGGTAGCTTCCGGTAGCTCCAAGGTCAGATCC I V L D D A M I D A A V K A A V F G S F L F Q G Q I

FIG. 3. Comparison of the nucleotide sequences of the *pahA* and *pahB* genes and the flanking regions identified in this study (*pah*; middle line) and the corresponding regions of *nah*, *ndo*, and *dox* (*nah-dox*; upper line), i.e., the *nahA* operon of *P*. *putida* NCIB 9816-4 and 9816 (positions 1 to 3911) and the *dox* operon of *Pseudomonas* sp. strain C18 (positions 1477 to 5605), respectively. The sequences of *nah* and *dox* between positions 1477 and 3911 are identical. The symbol  $\lceil$  and three dots above *nah-dox* indicate initiation and stop codons of ORFs predicted by Simon et al. (18), Kurkela et al. (10) and Denome et al. (3). The deduced amino acid sequences for the *pah* genes (bottom line) are shown in the one-letter code. Amino acid residues not conserved between the *pah* products and the *nahA<sub>a</sub>* and the *doxEF* products are outlined. The sterisks indicate stop codons. A putative promoter (-35 and -10) and probable ribosome-binding sequences (RBS) are underlined. Estimated [2Fe-2S]-binding sites are doubly underlined. The sequence similar to the NahR binding sequence (-70) is also underlined.

of NDD reported by Jerina et al. (8). These results indicate that the *pahA* gene encodes dioxygenase and that naphthalene was converted to NDD by PahA.

Phenanthrene was also converted to cis-3,4-phenanthrene dihydrodiol by a high-density suspension of OUS8211 cells carrying pKTS1 (data not shown). The efficiency of conversion of phenanthrene was 1/10 of that of naphthalene.

**Conversion of NDD by the** *pahB* gene product. Because we could not find a typical product of naphthalene catabolism in the high-density suspension of OUS8211 cells carrying the *pahAB* cluster in a preliminary test, the *pahB* gene was joined to the *tac* promoter and expressed in *E. coli*. A 0.9-kb SacII-HincII fragment containing the *pahB* gene was inserted into the SmaI site of tac promoter expression vector pKK223-3,

		Ţ	ABLE 1. Nuk	cleotide	and ami	no acid	l sequence sim	ilarities b	etween ]	ahA a	and other NC	00-type	enzyme	a				
	P. putida C	DUS82		P. putida	67		P. putida NC	IB 9816-4,	NCIB 98	16	P.	putida F	L		P. pseudo	alcaligen	s KF707	
NDO-type enzyme	Gene	No. of amino	Gene	No. of amino	% Sim	ilarity <sup>b</sup> f:	Gene	No. of amino	% Simila of:	rity	Gene	No. of amino	% Simil of:	arity <sup>b</sup>	Gene	No. of amino	% Simil of:	arity <sup>b</sup>
	(% G+C)	acid residues	(% G+C)	acid residues	Nucleo- tides	Amino acids	(% G+C)	acid 1 residues	Nucleo- A tides	mino	1 (% G+C)	acid ] residues	Nucleo- tides	Amino acids	(% G+C)	acid esidues	Nucleo- , tides	Amino acids
Ferredoxin reductase	: pahA. (54.4)	) 328	nahA <sub>a</sub> (55.1)	328	76	96	nahA <sub>a</sub> (55.7)	328	96	90	odA (63.7)	410	53	28	bphA4 (69.0)	408	52	27
Ferredoxin	pahA, (51.2)	) 104	nahA <sub>h</sub> (52.0)	107	<b>9</b> 4	93	nahA <sup>b</sup> <sup>c</sup> (50.3)	104	91	90 1	odB (57.0)	107	58	38	bphA3 (56.3)	109	58	39
ISP <sup>d</sup> large subunit	pahA, (53.7)	(449	nahA, (53.6)	449	95	96	nahA <sup>e</sup> (52.6)	449	93	94 t	odC1 (60.4)	450	56	39	bphA1 (62.2)	458	56	37
ISP small subunit	pahA <sub>d</sub> (51.5)	) 194	$nahA_{d}(52.1)$	194	95	91	$nahA_b^{f}(50.0)$	194	93	92 t	odC2 (58.7)	187	55	32	bphA2 (58.2)	213	55	50
						-												

The nucleotide and amino acid sequences of the dioxygenase are from the GenBank DNA data base. Compared with P. putida OUS82 gene and enzyme

nahA<sub>b</sub> ndoA.

We thank K. Yano and M. Fukuda, Nagaoka University of Technology, for helpful discussion and H. Hamada, Department Fundamental Science, Okayama University of Science, for nuclear magnetic resonance analysis. <sup>1</sup> ISP, iron-sulfur protein. <sup>2</sup> nahA<sub>b</sub> ndoB. nahA<sub>b</sub> ndoB.

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, and Science of Japan to H.K. Part of this study was supported by a grant to N.T. from the Okayama Foundation for Science and Technology.

## REFERENCES

1. Brosius, J., and A. Holy. 1984. Regulation of ribosomal RNA promoters with a synthetic lac operator. Proc. Natl. Acad. Sci. USA 81:6929-6933.

and pNA1 was obtained (Fig. 2). E. coli JM109 carrying pNA1 was cultured with induction by isopropyl-B-D-thiogalactopyranoside (IPTG), and an extract was prepared from the cells. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of the extract gave an IPTG-induced protein band (data not shown). The molecular mass of the protein was estimated to be 27 kDa.

Conversion of NDD by the extract was analyzed by highpressure liquid chromatography. NDD, which was prepared by conversion of naphthalene by P. putida OUS8211 carrying pKTS1; NAD+; and the cell extract prepared from E. coli carrying pNA1 were mixed and incubated at 30°C. Naphthalene diol was transiently detected after 1 min of incubation, and accumulation of  $\beta$ -naphthoquinone, to which naphthalene diol was converted by oxidation in air, was detected after 15 min of incubation. The amino-terminal amino acid sequence of the 27-kDa protein was MGNQQVVSITG, which agreed with that of the deduced protein product of the pahB gene. These results indicate that the pahB gene encodes PAH dihydrodiol dehydrogenase.

In this study, we identified the PAH dioxygenase gene, pahA, and the PAH dihydrodiol dehydrogenase gene, pahB, and sequenced them. PahA is a multicomponent enzyme, like NDO (6). The nucleotide and amino acid sequences of PahA are similar to those of NDOs (more than 90% similarity), but the similarities between PahA and toluene dioxygenase or biphenyl dioxygenase are less at 52 to 58% for nucleotides and 27 to 39% for amino acids. The pahB gene is expected to be similar to the nahB gene, the nucleotide sequence of which has not been reported, because pahB and nahB hybridized strongly (data not shown). The pah and nah clusters were probably derived from the same ancestor.

A search of a DNA sequence data base showed that the region downstream from the pahB gene containing the 5'terminal region of ORF6 is similar to the sequences that encode E. coli aldehyde dehydrogenase and Alcaligenes eutrophus acetaldehyde dehydrogenase. In plasmid NAH7, the nahF gene encoding aldehyde dehydrogenase is located between *nahB* and *nahC* (5). ORF6 is probably the *pahF* gene.

After we submitted this report for publication, the nucleotide sequence of the dox operon of Pseudomonas strain sp. C18 was reported by Denome et al. (3). The sequence between doxA and doxF is very similar to the sequence between  $pahA_{b}$ and pahB reported here, but the dox sequence does not contain the region corresponding to the promoter and  $pahA_a$  (Fig. 3). Some of the ORFs in the dox operon predicted by them do not initiate at an ATG or GTG codon and do seem strange. We have also determined the sequence of all regions of the pah cluster. We will report that nucleotide sequence, together with biochemical evidence, elsewhere.

Nucleotide sequence accession number. The nucleotide sequence in Fig. 3 will appear in the DDBJ, EMBL, and GenBank data bases under accession no. D16629.

- Davies, J. I., and W. C. Evans. 1964. Oxidative metabolism of naphthalene by soil pseudomonads: the ring fission mechanism. Biochem. J. 91:251–261.
- 3. Denome, S. A., D. C. Stanley, E. S. Olson, and K. D. Young. 1993. Metabolism of dibenzothiophene and naphthalene in *Pseudomonas* strains: complete DNA sequence of an upper naphthalene catabolic pathway. J. Bacteriol. **175**:6890–6901.
- Dunn, N. W., and I. C. Gunsalus. 1973. Transmissible plasmid coding early enzymes of naphthalene oxidation in *Pseudomonas putida*. J. Bacteriol. 114:974–979.
- Eaton, R. W., and P. J. Chapman. 1992. Bacterial metabolism of naphthalene: construction and use of recombinant bacteria to study ring cleavage of 1,2-dihydroxynaphthalene and subsequent reactions. J. Bacteriol. 174:7542–7554.
- Ensley, B. D., D. T. Gibson, and A. L. Laborde. 1982. Oxidation of naphthalene by a multicomponent enzyme system from *Pseudo*monas sp. strain NCIB 9816. J. Bacteriol. 149:948–954.
- 7. Figurski, D. H., and D. R. Helinski. 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA **76**:1648–1652.
- Jerina, D. M., J. W. Daly, A. M. Jeffrey, and D. T. Gibson. 1971. cis-1,2-Dihydroxy-1,2-dihydronaphthalene: a bacterial metabolite from naphthalene. Arch. Biochem. Biophys. 142:394–396.
- Kiyohara, H., S. Torigoe, N. Kaida, T. Asaki, T. Iida, H. Hayashi, and N. Takizawa. 1994. Cloning and characterization of a chromosomal gene cluster, *pah*, that encodes the upper pathway for phenanthrene and naphthalene utilization by *Pseudomonas putida* OUS82. J. Bacteriol. 176:2439–2443.
- Kurkela, S., H. Lehväslaiho, E. T. Palva, and T. H. Teeri. 1988. Cloning, nucleotide sequence and characterization of genes encoding naphthalene dioxygenase of *Pseudomonas putida* strain NCIB 9816. Gene 73:355–362.
- 11. Nakazawa, A. (Yamaguchi University). 1990. Personal communication.
- Patel, T. R., and D. T. Gibson. 1974. Purification and properties of (+)-cis-naphthalene dihydrodiol dehydrogenase of *Pseudomonas* putida. J. Bacteriol. 119:879–888.

- Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin. 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. 172:2351–2359.
- Rogers, J. E., and D. T. Gibson. 1977. Purification and properties of *cis*-toluene dihydrodiol dehydrogenase from *Pseudomonas putida*. J. Bacteriol. 130:1117–1124.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463-5467.
- Schell, M. A., and E. F. Poser. 1989. Demonstration, characterization, and mutational analysis of NahR protein binding to *nah* and *sal* promoters. J. Bacteriol. 171:837–846.
- Sheen, J.-Y., and B. Seed. 1988. Electrolyte gradient gels for DNA sequencing. BioTechniques 6:942–944.
- Simon, M. J., T. D. Osslund, R. Saunders, B. D. Ensley, S. Suggs, A. Harcourt, W.-C. Suen, D. L. Cruden, D. T. Gibson, and G. J. Zylstra. 1993. Sequences of genes encoding naphthalene dioxygenase in *Pseudomonas putida* strains G7 and NCIB 9816-4. Gene 127:31–37.
- Taira, K., N. Hayase, N. Arimura, S. Yamashita, T. Miyazaki, and K. Furukawa. 1988. Cloning and nucleotide sequence of the 2,3-dihydroxybiphenyl dioxygenase gene from the PCB-degrading strain of *Pseudomonas paucimobilis* Q1. Biochemistry 27:3990– 3996.
- Vieira, J., and J. Messing. 1987. Production of single-stranded plasmid DNA. Methods Enzymol. 153:3–11.
- Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119.
- Yen, K.-M., and I. C. Gunsalus. 1982. Plasmid gene organization: naphthalene/salicylate oxidation. Proc. Natl. Acad. Sci. USA 79: 874–878.
- Yen, K.-M., and C. M. Serdar. 1988. Genetics of naphthalene catabolism in pseudomonads. Crit. Rev. Microbiol. 15:247–268.
- You, I., D. Ghosal, and I. C. Gunsalus. 1988. Nucleotide sequence of plasmid NAH7 gene *nahR* and DNA binding of the *nahR* product. J. Bacteriol. 170:5409-5415.