The Journal of Microbiology, December 2001, p.334-337
Vol. 39, No. 4

NOTE

Cloning and Sequence Analysis of the *hpaD* Gene Responsible for Homoprotocatechuate 2,3-Dioxygenase from *Pseudomonas* sp. DJ-12

Sang-Mahn Lee, Jong-Chan Chae[1], Youngsoo Kim[2], and Chi-Kyung Kim[1]*

Department of Life Science, Chongju University, Cheongju 360-764, Korea
[1]Department of Microbiology and Biotechnology, and [2]Department of Pharmacy, Chungbuk National University, Cheongju 361-763, Korea

(Received October 10, 2001 / Accepted December 13, 2001)

The degradative pathway of homoprotocatechuate (HPC) is the bacterial route whereby 3,4-dihydroxyphenylacetic acid is catabolized to pyruvate and succinate by a series of sequential reactions. The HPC is catalized by homoprotocatechuate 2,3-dioxygenase (HPC-2,3O) to form 5-carboxymethyl-2-hydroxymuco semialdehyde. In this study, the *hpaD* gene encoding HPC-2,3O was cloned from the chromosomal DNA of *Pseudomonas* sp. DJ-12 and its nucleotide sequence was analyzed. The open reding frame of *hpaD* gene was found to be composed of 864 nucleotide pairs and to encode a polypeptide with 287 amino acid residues. The deduced amino acid sequence of the HPC-2,3O from *Pseudomonas* sp. DJ-12 exhibited 60~64% homology with those of the corresponding enzymes from *E. coli*, *Salmonella enterica*, and *Klebsiella pneumoniae*.

Key words: *hpaD*, homoprotocatechuate 2,3-dioxygenase, nucleotide sequence, homology

Homoprotocatechuate (HPC) is the intermediate product from the bacterial catabolism of phenylalanine and tyrosine, and also a compound released during the degradation of lignin. It is an important reaction in the environment in terms of degradation of such natural organic compounds. The *meta*-cleavage pathway catalyzed by homoprotocatechuate-2,3-dioxygenese (HPC-2,3O) is the central process in homoprotocatechuate metabolism. The catecholic substrate can be cleaved by a dioxygenase which breaks the bond between (intradiol, or *ortho* cleavage) or adjacent to (extradiol, or *meta* cleavage) the two carbons with a hydroxyl group (19). As shown in Fig. 1(I), HPC is converted initially by *meta*-cleavage with two oxygens to produce 5-carboxyl-2-hydroxylmuconate semialdehyde (CHMS) which is oxidized again by an NAD-dependent reaction to give 5-carboxylmethyl-2-hyroxymuconate (CHM). The final products of the pathway are the molecules that enter the tricarboxylic acid cycle (1, 5, 7, 10). From various bacteria, various enzymes and genes involved in the *meta*-fission pathway for HPC were investigated, (6, 9, 12, 13, 14, 15, 17). HPC-2,3O was purified from *Brebibacterium fuscum* and characterized (8), and the gene encoding HPC-2,3O was studied in *E. coli* (11, 16).

Pseudomonas sp. DJ-12 is a bacterial strain that was previously isolated from contaminated waste (20). The strain has been reported to utilize 4-chlorobiphenyl, biphenyl, and 4-chlorobenzoate as sole sources of carbon and energy through the *meta*-cleavage pathway (3, 20). The *fcb* genes for dechlorination of 4-chlorobenzoate was also analyzed by Chae *et al* (3, 4). In this study, the *hpaD* gene encoding HPC-2,3O was cloned from the locus adjacent to the *fcbB* gene of *Pseudomonas* sp. DJ-12. The complete nucleotide sequence of the *hpaD* gene was analyzed and its amino acid sequence was compared with those from other strains.

Cloning and nucleotide sequencing of hpaD gene

The bacterial cells of *Pseudomonas* sp. DJ-12 and host *E. coli* were cultivated in Luria-Bertani medium (1% tryptone, 0.5% yeast extract, 1% NaCl) supplemented with ampicillin (50 μl/ml) as a selective agent. *Pseudomonas* sp. DJ-12 was also grown at 30°C in MM2 minimal medium [1M $FeSO_4 \cdot 7H_2O$, 100 mM $CaCl_2 7H_2O$, 1 μM $MgSO_4 \cdot 7H_2O$, 8.5 mM NaCl, 18 mM $(NH_4)_2SO_4$, 10 mM potassium phosphate buffer (pH 7.0), 1.5% agar] containing 0.5 mM 4CBA. Transformation and DNA

* To whom correspondence should be addressed.
(Tel) 82-43-261-2300; (Fax) 82-43-264-9600
(E-mail) environ@trut.chungbuk.ac.kr

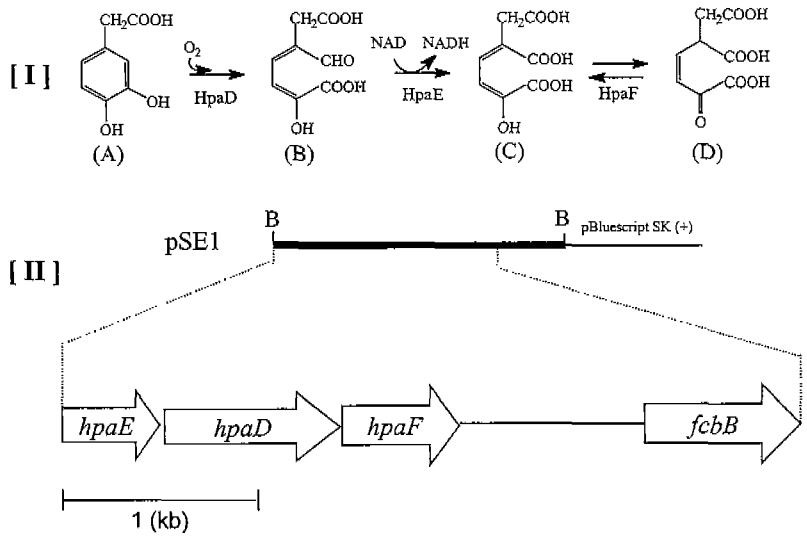

Fig. 1. Degradative pathway of homoprotocatechuate via *meta*-cleavage (I) and arrangement of the *hpa* gene cluster of *Pseudomonas* sp. DJ-12 (II). A, homoprotocatechuate (HPC); B, 5-carboxy-methyl-2-hydroxymuconate semialdehyde (CHMS); C, 5-carboxymethyl-2-hydroxymuconate (CHM); D, 5-carboxymethyl-2-oxo-hexa-3-ene-1,6-dioate (COHED); HpaD, HPC dioxygenase; HpaE, CHMS dehydrogenase; HpaF, CHM isomerase; B, *Bam*HI; X, *Xho*I; C, *Cla*I.

manipulation for cloning of the *hpaD* genes were carried out in general as described by Sambrook *et al.* (18). The cosmid pWE15 and pBluescript SK(+) vectors (Stratagene, USA) were used for cloning and other molecular work. The plasmids were isolated with a Wizard DNA purification system (Promoga, USA). Nucleotide sequences of both strands were determined with an automated-sequencing apparatus (ALFwin™ DNA sequencer, Pharmacia, USA). Sequencing reactions were performed with an ALFexpress™AutoRead™sequencing kit.

The pKC1 carrying *hpaD* gene was previously cloned from the chromosomal DNA of *Pseudomonas* sp. DJ-12 (3, 4). To obtain the *hpaD* gene, the pKC1 plasmid were futher subcloned to construct pKC157 (4). Finally the pSE1 harboring the *hpaD* gene was constructed from pKC157 plasmid. It contained the 3.4 kb *Bam*HI fragment derived from pKC157. The three ORFs involved in the catabolizing the homoprotocatechuate metabolism is located upstream with *fcbB* gene encoding the 4CBA-CoA dehalogenase which is related to the 4CBA degradation (Fig. 1(II)). The nucleotide sequences of 3.4 kb containing the *hpaD* gene for ring fission of HPC compound was determined. The open reading frame corresponding to the *hpaD* gene consists of 864 base pairs, and ATG codon is located 10 bp downstream of the putative ribosome binding sequence GGGGA (Fig. 2). The *hpaD* gene of strain DJ-12 is located between the *hpaE* and *hpaF* which code for 5-carboxymethyl-2-hydroxy-muconic semialdehyde dehydrogenase and 5-carboxymethyl-2-hydroxy-muconic acid isomerase, respectively. The HPC-2,3O encoded by *hpaD* was putatively determined to be a polypeptide chain with 287 amino acid residues. The HPC-2,3 dioxygenase enzymes of *E.coli* C (HpcB) (16),

```
5'TGCCTGGGTTCGCACCACATTCCGCACTGGGGGACCTAAagacATGGGCAAGCTCGCGCT   60
 C  L  G  S  H  H  I  P  H  W  G  T  *     M  G  K  L  A  L
 hpaE ──►                RBS              hpaD ──►
GGCGGCCAAGATCACGCACGTGCCGTCGATGTACCTCTCGGAGTTCCCGGGACCCAACCA   120
 A  A  K  I  T  H  V  P  S  M  Y  L  S  E  F  P  G  P  N  Q
GGGATGCCGGGAGGCGGCGATCAACGGGCACAAGGAGATTGACCGCCGCTGCCGGGAGCT   180
 G  C  R  E  A  A  I  N  G  H  K  E  I  D  R  R  C  R  E  L
GGGCGTGGACACCATCGTGGTGTTCGACGTGCACTGGCAGGTGAACAGCGAGTACCACAT   240
 G  V  D  T  I  V  V  F  D  V  H  W  Q  V  N  S  E  Y  H  I
CAACTGCGGGCCGAGGTTCGAAGGCGTGTACACCAGCAACGAGCTGCCGCACTTCATCAA   300
 N  C  G  P  R  F  E  G  V  Y  T  S  N  E  L  P  H  F  I  K
GAACTTGCCTTTCTCGTACCCGGGCAATCCCGGGCTCGGGCACCTGATCGCCGACGTCGC   360
 N  L  P  F  S  Y  P  G  N  P  G  L  G  H  L  I  A  D  V  A
CAACGAAATGGGAGTCAAGAGCCGCGCCCACTCCGACACCACGCTGGAGCTGGAGTACGG   420
 N  E  M  G  V  K  S  R  A  H  S  D  T  T  L  E  L  E  Y  G
CACGCTGGTGCCCATGCGTTACATGAACGGCGACCAGCACTACAAGGTGGTCAGCGTCAG   480
 T  L  V  P  M  R  Y  M  N  G  D  Q  H  Y  K  V  V  S  V  S
CGGCTGGTGCGACTGGCACGACCTGCACGAATCCGGCCGCTTCGGGCTGGCGGTGCGCAA   540
 G  W  C  D  W  H  D  L  H  E  S  G  R  F  G  L  A  V  R  K
GGCCATCGAGGAGCGCCACCACGGCACGGTGGCCGTGTTCGCCAGCGGTTCGCTGTCGCA   600
 A  I  E  E  R  H  H  G  T  V  A  V  F  A  S  G  S  L  S  H
CCACTTCGCCGATAACGGCCGCTCGCCCGATTTCATGCACAAGGTGTACGACCCGTTCCT   660
 H  F  A  D  N  G  R  S  P  D  F  M  H  K  V  Y  D  P  F  L
GGAGCAGGTGGACCAGCGCGTGGTGGACCTGTGGAAGTCCGGCGACTGGAAGACGTTCGT   720
 E  Q  V  D  Q  R  V  V  D  L  W  K  S  G  D  W  K  T  F  V
GGGCATGCTGCCGATGTACGCCGACAAGTGCTGGGGCGAAGGCGGCATGCACGACACCGC   780
 G  M  L  P  M  Y  A  D  K  C  W  G  E  G  G  M  H  D  T  A
CATGCTGCTGGGCCTGCTGGGCTGGGACCGCTACCGCGCGCCGGTGGAGATCGTCACGCC   840
 M  L  L  G  L  L  G  W  D  R  Y  R  A  P  V  E  I  V  T  P
CTACTTCGGCAGCTCGGGCACCGGCCAGATCAACGCAATCTTCCCGGTCACGCCGCTGCC   900
 Y  F  G  S  S  G  T  G  Q  I  N  A  I  F  P  V  T  P  L  P
GGCCTGAggtgcaccccATGCCGCACCTAGTGATCCTCTACACGCCGAACGTCGAGCTCG   960
 A  *  RBS     M  P  H  L  V  I  L  Y  T  P  N  V  E  L  D
                hpaF ──►
```

Fig.2. Nucleotide and deduced amino acid sequence of the *Pseudomonas* sp. DJ-12 HPC 2,3-dioxygenase gene (*hpaD*). The sequence is written in the 5'-3' direction of the coding strand with the deduced amino acid sequence below. The RBS sequence is underlined, the start codon is indicated by the symbol ATG. Stop codon is indicated by asterisk.

E.coli W (HpaD) (12), *Salmonella enterica* (HpaD), and *Klebsiella pneumoniae* (HpaD) (5) have been reported to

```
DJ-12          MGKLALAAKITHVPSMYLSEFPGPNQGCREAAINGHKEIDRRCRELGVDTIVVFDVHWQV
E.coli_C       MGKLALAAKITHVPSMYLSELPGKNHGCRQGAIDGHKEISKRCREMGVDTIIVFDTHWLV
E.coli_W       MGKLALAAKITHVPSMYLSELPGKNHGCRQGAIDGHKEISKRCREMGVDTIIVFDTHWLV
Salmonella     MGKLALAAKITHVPSMYLSELPGKNHGCRQAAIDGHIEIGKRCREMGVDTIIVFDTHWLV
Klebsiella     MGKLALAAKTTHVPSMYLSELPGKNHGCRQGAIDGHKEIGKRCRELGVDTFIVFDTHWLV
               ********** ********** ** * ***  ** ** **  **** ****  *** ** *

DJ-12          NSEYHINCGPRFEGVYTSNELPHFIKNLPFSYPGNPGLGHLIADVANEMGVKSRAHSDTT
E.coli_C       NSAYHINCADHFEGVYTSNELPHFIRDMTYNYEGNPELGQLIADEALKLGVRAKAHNIPS
E.coli_W       NSAYHINCADHFEGVYTSNELPHFIRDMTYNYEGNPELGQLIADEALKLGVRAKAHNIPS
Salmonella     NSAYHINCADHFQGVYTSNELPHFIRDMTYDYDGNPELGHLIADEAVKLGVRAKAHNIPS
Klebsiella     NSAYHINCADHFQGVYTSNELPHFIRDMTYDYDGNPELGHLIADKTVKLGVRAKAHNIPS
               ** *****  * ************      * *** ** ****     **  **

DJ-12          LELEYGTLVPMRYMNGDQHYKVVSVSGWCDWHDLHESGRFGLAVRKAIEERHHGTVAVFA
E.coli_C       LKLEYGTLVPMRYMNEDKHFKVVSISAFCTVHDFADSRKLGEAILKAIEQ-YDGTVAVLA
E.coli_W       LKLEYGTLVPMRYMNEDKHFKVVSISAFCTVHDFADSRKLGEAILKAIEQ-YDGTVAVLA
Salmonella     LKLEYGTLVPMRYMNSDKHFKVVSISAFCTVHDFADSRRLGEAILKAIEK-YDGTVAVFA
Klebsiella     LKLDYGTLVPMRYMNADKHFKVVSISAFCTVHDFADSRKLGEAIRKAIEK-YDGTVAVLA
               * * ********** * * **** * *  **  *  * * ****    ***** *

DJ-12          SGSLSHHFADNGRSPDFMHKVYDPFLEQVDQRVVDLWKSGDWKTFVGMLPMYADKCWGEG
E.coli_C       SGSLSHRFIDDQRAEEGMNSYTREFDRQMDERVVKLWREGQFKEFCNMLPEYADYCYGEG
E.coli_W       SGSLSHRFIDDQRAEEGMNSYTREFDRQMDERVVKLWREGQFKEFCNMLPEYADYCYGEG
Salmonella     SGSLSHRFIDDQRAEEGMNSYTREFDHQMDERVVKLWREGKFKEFCTMLPEYADYCYGEG
Klebsiella     SGSLSHRFIEDQRAEEGMNSYTREFDHQMDERVVKLWREGKFKEFCTMLPENAEYCYGEG
               ****** *   *   *   * * * *** ** * * * *  ***  *  * ***
```

Fig. 3. Comparison of the predicted amino acid sequence of *hpaD* gene from *Pseudomonas* sp. DJ-12 with those of other HPC dioxygenase. The amino acid sequences from the homoprotocatechuate 2,3-dioxygenase are as follow: DJ-12 in this study, *Salmonella enterica* (accession number AAD53499), *E. coli* W (accession number CAA86042), *E. coli* C (accession number Q05353), and *Klebsiella pneumoniae* (accession number CAB65144).

consist of 276, 283, 283 and 276 amino acid residues, respectively.

Homology of homoprotocatechuate 2,3-dioxygenase

The DNA sequences of the *hpaD* gene were deposited in the GenBank database under accession number AF051771. Nucleotide and deduced amino acid sequences were analyzed using DNASIS and PROSIS software (Hitachi version 7.0, Japan). Searches for the similarities of nucleotide and amino acid sequences were done using the FASTA and BLAST programs (2), and EMBL and GenBank databases.

When a computer-assisted search of various databases, four significant matching sequences for HPC-2,3 dioxygenase were found from *E. coli*, *Salmonella* and *Klebsiella*. The deduced amino acid sequence of *hpaD* of *Pseudomonas* sp. DJ-12 exhibits 64, 63, 60, 62% identities to those of the corresponding enzymes from *Salmonella enterica* (accession number AAD53499), *E. coli* W (accession number CAA86042), *E .coli* C (accession number Q05353), and *Klebsiella pneumoniae* (accession number CAB65144), respectively. The sequence alignment shown in Fig. 3 indicates that the HPC 2,3-dioxygenase has conserved residues among different bacterial strains. The relatedness of the homology in the amino acid sequence

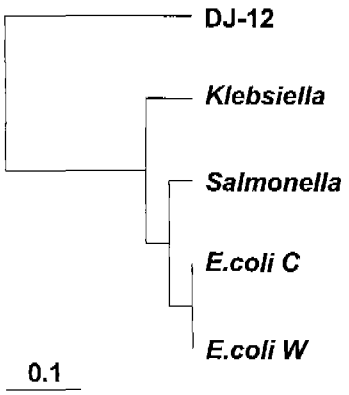

Fig. 4. Phylogenetic tree produced by comparison of amino acid sequences from different homoprotocatechuate 2,3-dioxygenase. The tree analysis was performed with the FASTA and BLASTA program. The scale bar indicates 0.1 substitution per site.

of the dioxygenase among different strains is shown in Fig. 4. It suggests that the *hpaD* gene of *Pseudomonas* sp. DJ-12 is quite different from the genes of other Gram negative strains in evolutionary relationships.

Acknowledgment

This work was supported by a grant (1999-2-202-004-3) from the Korea Science and Engineering Foundation. We gratefully acknowledge Marco Bazzicalupo for helpful discussion which was provided by a visiting program (01H020000500) of the International Cooperative Research Project between Korea and Italy.

References

1. Arunachalam, U., V. Massey, and S.M. Miller. 1994. Mechanism of p-hydroxyphenylacetate-3-hydroxylase. *J. Biol. Chem.* 269, 150-155.
2. Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402.
3. Chae, J.C., and C.K. Kim. 1997. Dechlorination of 4-chlorobenzoate by *Pseudomonas* sp. DJ-12. *J. Microbiol.* 35, 290-294.
4. Chae, J.C., Y.S . Kim, Y.C. Kim, G.J. Zylstra, and C.K. Kim. 2000. Genetic structure and functional implication of the *fcb* gene cluster for hydrolytic dechlorination of 4-chlorobenzoate from *Pseudomonas* sp. DJ-12. *Gene* 258, 109-116.
5. Gibello, A., E. Ferrer, J. Sanz, and M. Martin. 1995. Polymer production by *Klebsiella pnumoniae* 4-hydroxyphenylacetic acid hydroxylase gense cloned in *Escherichia coli*. *Appl. Environ. Microbiol.* 61, 4167-4171.
6. Jenkins, J.R., and R.A. Cooper. 1988. Molecular cloning, expression, and analysis of the genes of the homoprotocatechuate catabolic pathway of *Escherichia coli* C. *J. Bacterol.* 170, 5317-5324.
7. Mohamed, M.E., B. Seyfried, A. Tschech, and G. Fuchs. 1993.

Anaerobic oxidation of phenylacetate and 4-hydroxyphenylacetate, to benzoyl-coenzyme A and CO_2 in denitrifying *Pseudomonas* sp. *Microbiology* 20, 563-573.

8. Miller, M.A., and J.D. Lipscomb. 1996. Homoprotocatechuate 2, 3-dioxygenase from *Brevibacterium fuscum*. *J. Biol. Chem.* 271, 5524-5535.
9. Masai, E., K. Sugiyama. N. Iwashita. S. Shimizu. J.E. Hauschild, T. Hatta, K. Kimbara, K. Yano, and M. Fukuda. 1997. The *bpcDEF meta-* cleavage pathway genes involved in biphenyl/polychlorinated biphenyl degradation are located on a linear plasmid and separated from the initial *bph ACB* genes in *Rhodococcus* sp. strain RHA1. *Gene* 187, 141-149.
10. Prieto, M.A., A.P. Aranoa, and J.L. Garcia. 1993. Characterization of an *Escherichicl coli* aromatic hydroxylase with a broad substrate range. *J. Bacteriol.* 175, 2162-2167.
11. Prieto, M.A., E. Diaz, and J.L. Garcia. 1996. Molecular characterization of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli* W: Engineering a mobile aromatic degradative cluster. *J. Bacteriol.* 178, 111-120.
12. Prieto, M.A., and J.L. Garcia. 1997. Identification of the 4-hydroxyphenylacetate transport gene of *Escherichia coli* W: consturction of a highly sensitive cellular biosensor. *FEBS Letters* 414, 293-297.
13. Prieto, M.A., and J.L. Garcia. 1997. Identification of a novel positive regulator of the 4-hydroxyphenylacetate catabolic pathway of *Escherichia coli*. *Biochem. Biophys. Res. Comm.* 232, 759-765.
14. Roper, D.I., J.M. Stringfellow, and R.A. Cooper. 1995. Sequence of the *hpcC* and *hpc*G genes of the *meta*-fission homoprotocatechuic acid pathway of *Escherichia coli* C: nearly 40% amino-acid identity with the analogous of the catechol pathway. *Gene* 156, 47-51.
15. Roper, D.I., and R.A. Cooper. 1990. Purification, some properties and nucleotide sequence of 5-carboxymethyl-2-hydroxymuconate isomerase of *Escherichia coli* C. *FEBS Letters* 266, 63-66.
16. Roper, D.I., and R.A. Cooper. 1990. Subcloning and nucleotide sequence of the 3,4-dihydroxyphenylacetate (homoprotocatechuate) 2,3-dioxygenase gene from *Escherichia coli* C. *FEBS Letters* 275, 53-57.
17. Stringfellow, J.M., B. Turpin, and R.A. Cooper. 1995. Sequence of the *Escherichia coli* C homoprotocatechuic acid degradative operon completed with that of the 2,4-dihydroxyhept-2-ene-1, 7-dioic acid aldolase-encoding gene (*hpcH*). *Gene* 166, 73-76.
18. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning; A Laboratory Manual. 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.
19. Vallaeys, T., and G. Soulas. 1997. Pesticides: Microbial degradation and effects on microorganisms. p 547-575, *In* J.D.V. Elsas (ed), Modern soil microbiology. Marcel-Dekker Inc., New York, USA.
20. Youn, D.J., J.J. Han, C.K. Kim, and Y. Kim. 1992. Divergence of the *cbp* genes in 4-chlorobiphenyl catabolizing bacteria. *Kor. J. Microbiol.* 30. 53-59.