BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **238,** 56–60 (1997)
ARTICLE NO. RC976959

Localization and Sequence Analysis of the *phnH* Gene Encoding 2-Hydroxypent-2,4-dienoate Hydratase in *Pseudomonas* sp. Strain DJ77

Sungje Kim,* Oh-Kyu Kweon,* Youngsoo Kim,† Chi-Kyung Kim,*
Ki-Sung Lee,‡ and Young-Chang Kim*,[1]

**School of Life Sciences and †Department of Pharmacy, Chungbuk National University, Cheongju 361-763, Korea; and ‡Department of Biology, Pai-Chai University, Taejon 302-735, Korea*

Received June 11, 1997

The *phnDEFG* genes of *Pseudomonas* sp. DJ77, which are responsible for the degradation of polyaromatic hydrocarbons and chlorinated aromatics, were located previously on the 6.8 kb *Xho*I fragment of chromosomal DNA. Here, we sequenced a downstream region hitherto unknown and identified the *phnH* gene encoding a 2-hydroxypent-2,4-dienoate hydratase, which is required for the conversion of 2-hydroxypent-2,4-dienoate to 4-hydroxy-2-oxovalerate in the *meta*-cleavage pathway of catechols. The relative position of the hydratase gene in the *phn* operon is unique compared to the other *meta*-cleavage operons which have a dehydrogenative branch of the pathway. The PhnH hydratase contains 264 amino acids with a *M*r of 28043. The deduced amino acid sequence of the PhnH enzyme is 60.9–31.6% identical to those of homologous enzymes encoded by the *todG, bphE, cmtF, bphH, bphX1, xylJ, dmpE, cumE,* MTCY03C7.20 and *etbE* genes. © 1997 Academic Press

In the *meta*-cleavage pathway for catechol or substituted catechols, 2-hydroxypent-2,4-dienoate (HPD) formed by either a hydrolytic (H) or dehydrogenative (D) route [1], is metabolized to 4-hydroxy-2-oxovalerate by the action of the HPD hydratase (HPDH) (Fig. 1) [2]. Up to now, several HPDHs were discovered from various strains including *Pseudomonas putida* CF600 [3], *P. putida* mt-2 [4], *P. putida* NCIB9816 [5], *P. putida* F1 [6,7], *Pseudomonas* sp. KKS102 [8], *Pseudomonas* sp. LB400 [9], *P. pseudoalcaligenes* KF707, *P. fluorescens* [10] and *Rhodococcus* sp. RHA1 [11], and the genes encoding these isofunctional enzymes were designated as *dmpE, xylJ, nahL, todG* and *cmtF* (both are from *P. putida* F1), *bphE, bphH, bphX1, cumE* and *etbE*, respectively.

Prior to this study, we have undertaken a detailed study on phenanthrene-utilizing *Pseudomonas* sp. strain DJ77 [12] which is able to use both H and D routes of the *meta*-cleavage pathway. The genes responsible for the *meta*-cleavage pathway were cloned from the chromosomal DNA and the recombinant plasmid was designated as pHENX7 [13]. On the cloned *Xho*I fragment of about 6.8 kb, the *phnDEFG* genes encoding a 2-hydroxymuconic semialdehyde hydrolase (HMSH), a catechol-2,3-dioxygenase (C23O), an unknown polypeptide and a 2-hydroxymuconic semialdehyde dehydrogenase (HMSD), respectively, were mapped (Fig. 2). The nucleotide sequences of the *phnD* and *phnE* genes were reported [14,15]. This 6.8 kb *Xho*I fragment contains a region of about 1.5 kb of hitherto unknown function, downstream of the *phnG* gene. In this report, we identified the *phnH* gene encoding a HPDH in this region, determined the nucleotide sequence and analyzed evolutionary relationships among hydratases which are involved in the *meta*-cleavage pathway.

MATERIALS AND METHODS

Bacterial strains, plasmids and DNA manipulations. *Pseudomonas* sp. strain DJ77 [12] is a natural isolate capable of degrading phenanthrene, biphenyl and 4-chlorobiphenyl. The recombinant plasmid pHENX7 [13] contains the *phnDEFG* genes, which are involved in the phenanthrene degradation pathway from *Pseudomonas* sp. strain DJ77. Isolation of plasmid DNA, transformation, restriction endonuclease digestion, ligation, agarose gel electrophoresis and other standard recombinant DNA techniques were performed as described by Sambrook et al. [16].

Nucleotide sequencing. Nucleotide sequences were determined directly from plasmids by using either the Sequenase version 2.0 kit from USB Co. or an Applied Biosystems automated DNA sequencer.

[1] To whom correspondence should be addressed (Fax: 82-431-68-2538; e-mail: youngkim@cbucc.chungbuk.ac.kr).

0006-291X/97 $25.00

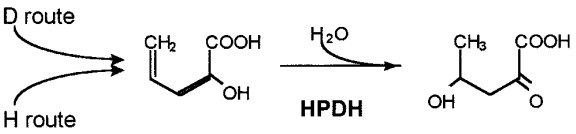

FIG. 1. Reaction catalyzed by 2-hydroxypent-2,4-dienoate hydratase (HPDH). The substrate formed by either a hydrolytic (H) or dehydrogenative (D) route is converted to 4-hydroxy-2-oxovalerate by the enzyme.

Suitable subcloning into the polycloning site of the pBluescript SK (+) sequencing vector allowed sequencing of both strands. Plasmid DNAs were purified by standard procedures using Qiagen Plasmid Kit (Qiagen Co., Cat. No. 12125).

Sequence analyses. The sequence of the *phnH* gene was deposited in the GenBank database under the accession number U97697. The nucleotide sequence and the deduced amino acid sequence were analyzed by using the DNASIS/PROSIS (Hitachi v. 7.0). Multiple alignments were carried out on a computer using the Clustal W algorithm [17] with all parameters set at their default values and fine-tuned manually. The EMBL and GenBank database accession numbers for the sequences used as reference sequences for analyses were as follows: *P. putida* CF600 (*dmpE*), X60835; *P. putida* F1 (*todG*), U09250; *P. putida* mt-2 TOL plasmid pWW0 (*xylJ*), M64747; *P. pseudoalcaligenes* KF707 (*bphX1*), D85853; *Mycobacterium tuberculosis* H37Rv (MTCY03C7.20), Z82098; *Pseudomonas* sp. LB400 (*bphH*), X76500; *P. putida* F1 (*cmtF*), U24215; *Pseudomonas* sp. KKS102 (*bphE*), D16407; *P. fluorescens* (*cumE*), D63377; *Rhodococcus* sp. RHA1 (*etbE*), D78322; *E. coli* MG1655 (*orf5*), U73857; *E. coli* CS520 (*mhpD*), Y09555; *E. coli* W3110 (*mhpD*), D86239.

Phylogenetic analysis. Phylogenetic analysis was performed using the maximum parsimony approach of PAUP (version 3.1.1; 37) [18]. The PAUP analysis was carried out using fairly standard default settings (heuristic search: 12 searches in which the input order of the sequences was randomly varied; the tree bisection-reconnection branch-swapping routine, treating gaps as individual insertion events) on Power Macintosh computer. For bootstrap analysis, the bootstrap program of the PAUP was used to generate 100 data sets, which were analyzed as described above. At least one gap in each run of gaps was treated as a new character state while the other gaps were treated as missing data to minimize the number of evolutionary events represented by a given run of gaps.

RESULTS AND DISCUSSIONS

Nucleotide sequence and location of the HPDH gene. In order to determine the sequence of the region hitherto unknown, subclones were constructed by cloning 1.8-kb *Hin*dIII-*Xho*I (pHX18) and 1.1-kb *Hin*dIII-*Hin*dIII (pHH11) fragments from pHENX7 into SK(+) vector. The entire nucleotide sequence and the deduced amino acid sequence are shown in Fig. 3. The sequence revealed one open reading frame (*phnH*) encoding a polypeptide. A database search by Blast E-mail server [19] detected a high degree of similarity between this polypeptide and the variety of hydratases. This suggests that the *phnH* encodes a HPDH. The calculated

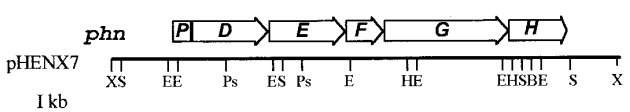

FIG. 2. Genetic map of the *phn* locus of pHENX7. Genes encode metabolic enzymes as follows: *phnD*, HMSH; *phnE*, C23O; *phnF*, a polypeptide with unknown function; *phnG*, HMSD; *phnH*, HPDH. The promoter is shown as P. Only relevant restriction sites are given. B, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; Ps, *Pst*I; S, *Sal*I; X, *Xho*I.

```
5'CCTCGAATTCTACACCGAGATCACCAACATCTGCGTAAAGCTTTAAGACGATGACTATCG  60
phnG:L  E  F  Y  T  E  I  T  N  I  C  V  K  L  *  phnH:M  T  I
   ACCCTAAAACCATCGAACAAGCGGCTCTGGTGCTGCGCGGCGCTGCCGAAAGCGGCACAC 120
   D  P  K  T  I  E  Q  A  A  L  V  L  R  G  A  A  E  S  G  T
   CGGTCAGTCCGATCCGAGACCTGATTGCCGCCGGTGGCGTTGAGGCCGCCTATGCTGTGC 180
   P  V  S  P  I  R  D  L  I  A  A  G  G  V  E  A  A  Y  A  V
   AGGAATCCAACACCCGGCACTACCTGGCTAGCGGGCGGCGGCTCGTAGGGCGCAAGATCG 240
   Q  E  S  N  T  R  H  Y  L  A  S  G  R  R  L  V  G  R  K  I
   GCCTGACGTCGCTTGCAGTCCAGCGCCAGTTGGGGGTGGACCATACCGATTACGGGATGC 300
   G  L  T  S  L  A  V  Q  R  Q  L  G  V  D  H  T  D  Y  G  M
   TGTTCGCAGACATGGACGTACCGGAAGGTATCCCGGTGTCGCTTGATCAGGTCATCCAAC 360
   L  F  A  D  M  D  V  P  E  G  I  P  V  S  L  D  Q  V  I  Q
   CCAAGATCGAGGCCGAGATTGCGATCGTCGTTGGGCGGGATTTGCCCCACCCCGACATGA 420
   P  K  I  E  A  E  I  A  I  V  V  G  R  D  L  P  H  P  D  M
   CCACCGCCGAGATGATCCGCGCGGTCGAATATGTCGTTCCGACAATCGAGATCGTCGACA 480
   T  T  A  E  M  I  R  A  V  E  Y  V  V  P  T  I  E  I  V  D
   GCCGCGTCACCAACTGGGACATCAAGATCTGGGACACGATCGCCGACAACGCGTCGAGCG 540
   S  R  V  T  N  W  D  I  K  I  W  D  T  I  A  D  N  A  S  S
   GACTGTTCGTGCTCGTCGCGGTGCCGCGCAAGCTAGATAGGCTGGATTTGCGCACGTGCG 600
   G  L  F  V  L  V  A  V  P  R  K  L  D  R  L  D  L  R  T  C
   GCATGGTTATGGAGGTAAAAGGCGAACCGATTTCGGTCGGCGCCGGGATCGCCTGTCTCG 660
   G  M  V  M  E  V  K  G  E  P  I  S  V  G  A  G  I  A  C  L
   GTAGACCAATCACCTCTTCCCTGTGGCTGGCGCGGGTCATGGCGAATGCCGGGCGGCCCC 720
   G  R  P  I  T  S  S  L  W  L  A  R  V  M  A  N  A  G  R  P
   TGCTGGAAGGCGACGTGATCCTTTCGGGCGCGCTCGGCCCGATGGCCGGGGTTTCTCGCG 780
   L  L  E  G  D  V  I  L  S  G  A  L  G  P  M  A  G  V  S  R
   GAGATGTCGTTGAAGCGCGGATCAATGGGCTTGGCACTGTCCGAGCTACGTTTGCTGCTG 840
   G  D  V  V  E  A  R  I  N  G  L  G  T  V  R  A  T  F  A  A
   ACTGAATTGGAAGTGGAGGTTGGTCAAATG 3'                             870
   D  *
```

FIG. 3. Nucleotide sequence of the *phnH* gene (GenBank accession number U97697). The amino acid sequences are also shown in one letter codes beneath the corresponding codons and the stop codons are marked with an asterisk. A putative ribosome binding site is underlined.

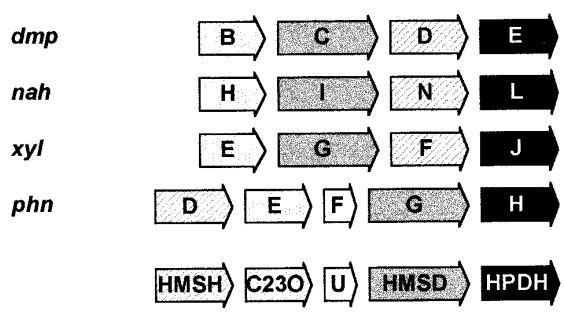

FIG. 4. Differences in the genetic organization of the *phn* operon from the other three operons, which can utilize both hydrolytic (H) and dehydrogenative (D) routes of the *meta*-cleavage pathway, such as *dmp*, *nah* and *xyl*. Only the genes analyzed in this study are shown. Arrows indicate the direction of transcription. Abbreviations are the same as the legend of Fig. 2 and also described in the introduction.

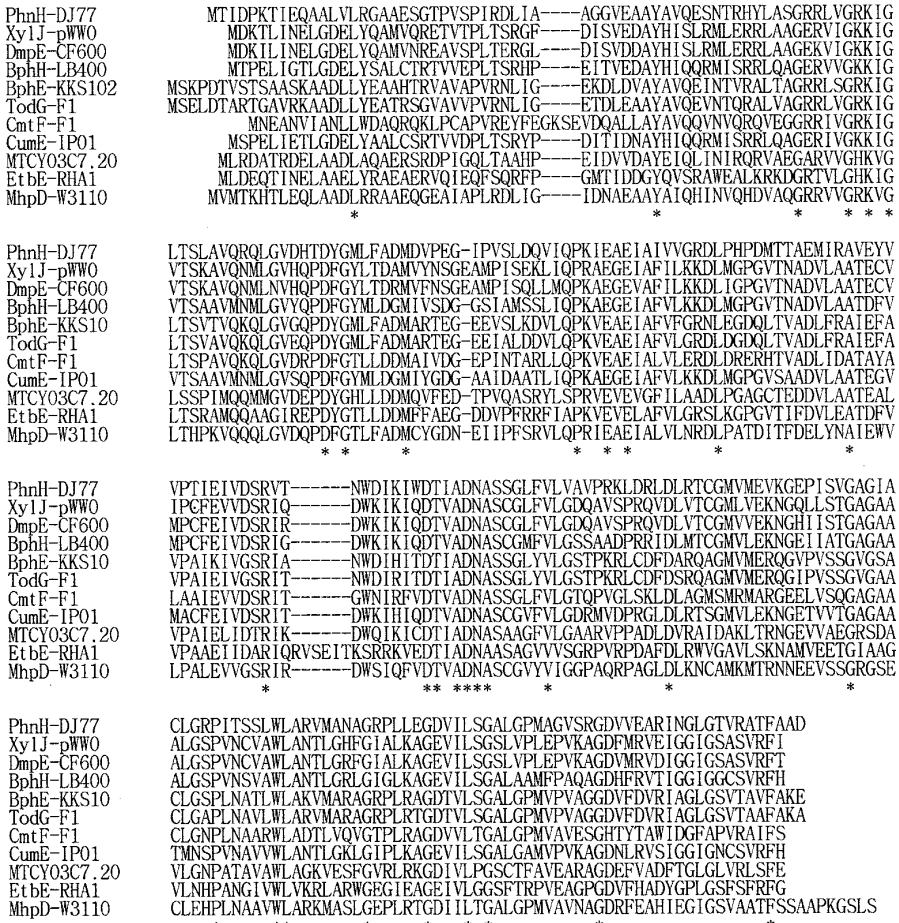

FIG. 5. Amino acid sequence alignments of PhnH and related proteins. The amino acids identical in all proteins are indicated by asterisks. References for individual sequence are given in the section 2.3. Because the amino acid sequences of three hydratases from *E. coli* W3110, MG1655 and CS520 are 99.6% identical with each other (a G residue, instead of E, has been found at position 204 of the published sequences of MG1655 and CS520), only the sequence from W3110 is shown. And the amino acid sequence of a HPDH from *P. pseudoalcaligenes* KF707 (BphX1) is also omitted as it differs from a HPDH of *Pseudomonas* sp. LB400 (BphH) at only two positions, 33 and 42 (Y and H, instead of I and Q, respectively). The full sequence of the NahL hydratase has not yet been reported.

*M*r of the PhnH polypeptide, based on the deduced sequence of 264 amino acids, is 28043. The *phnH* gene is preceded by a Shine-Dalgarno-type sequence [20]. The G + C content of the coding region of the newly identified *phnH* gene (62.4%) is slightly higher than those of the *phnE* (57.4%) [15] and *phnD* (60.5%) [14] genes encoding a C23O and a HMSH, respectively. A partial sequence of the *phnG* gene encoding a HMSD was found immediately upstream of the *phnH* gene. Therefore the gene order of the catabolic operon of pHENX7 responsible for catechol oxidation is determined to be: promoter-*phnD* (HMSH)-*phnE* (C23O)-*phnF* (unknown polypeptide)-*phnG* (HMSD)-*phnH* (HPDH).

As shown in Fig. 4, the relative position of the hydratase gene, *phnH*, in the *phn* operon is unique compared to those of the other three *meta*-cleavage operons, *dmp* [3], *xyl* [4] and *nah* [5], which are able to use both H and D branches of the pathway. The PhnH hydratase gene is located immediately 4 bp downstream of the PhnG dehydrogenase gene (in order of HMSH-C23O-ORF-HMSD-HPDH), while the equivalent hydratase gene is preceded by the hydrolase gene (in order of C23O-HMSD-HMSH-HPDH) in the *dmp* [3], *xyl* [4] and *nah* [5] operons. In the operons containing only a H branch such as *tod* (*P. putida* F1) [4,21] and *bph* (both from *Pseudomonas* sp. LB400 [9] and *Pseudomonas* sp. KKS102 [8]), which are responsible for the degradation of toluene and biphenyl, respectively, the hydratase gene is located downstream (in the *tod* operon) or upstream (in both *bph* operons) of the hydrolase gene at intervals of 3-9 genes.

Sequence comparisons. Binary sequence comparison revealed that the PhnH hydratase is more closely related to HPDHs from *P. putida* F1 (TodG, 60.9% iden-

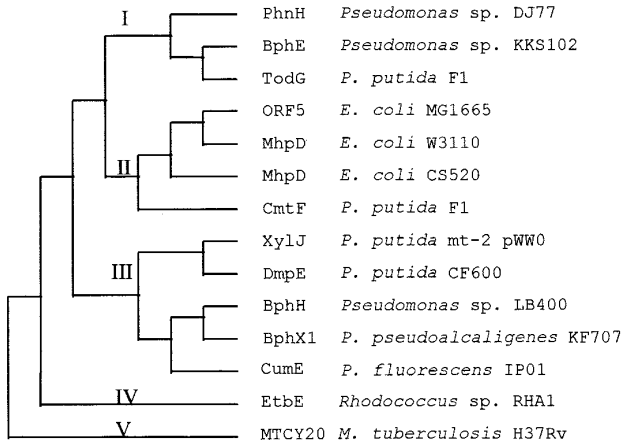

FIG. 6. Phylogram of the best tree obtained by PAUP analyses of alignment of 14 hydratase sequences. Symbols at branch points designate each subfamily as discussed in the text. MTCY20 (MTCY03C7.20) was used as an outgroup. The bootstrap analyses show that these clades are very stable (bootstrap values over 70%, results not shown).

tity) [6] and *Pseudomonas* sp. KKS102 (BphE, 56.7% identity) [8], which can use only a H branch, although the *phn* operon can utilize both H and D branches of the pathway.

The PhnH hydratase is more distantly related to the eight sequenced HPDHs from *Pseudomonas* sp. LB400 (BphH) [9], *P. pseudoalcaligenes* KF707 (BphX1), *P. putida* mt-2 TOL plasmid pWW0 (XylJ) [2], *P. putida* F1 (CmtF) [7], *P. putida* CF600 (DmpE) [3], *P. fluorescens* (CumE) [10], *Mycobacterium tuberculosis* H37Rv (MTCY03C7.20) and *Rhodococcus* sp. RHA1 (EtbE) [11], being 44.4%, 43.9%, 43.2%, 47.4%, 42.5%, 41.3%, 37.3% and 31.6% identical, respectively. The remaining three sequenced 2-keto-4-pentenoate hydratases from *E. coli* CS520 (MhpD), *E. coli* W3110 (MhpD) and *E. coli* MG1665 (ORF5) have 49.4%, 49.0% and 49.4% identity, respectively.

Amongst all 14 enzymes shown in the alignment, 34 amino acids were found to be completely conserved (Fig. 5). Recent study on the crystal structure of enoyl-CoA hydratase suggests that a glutamate serves as the catalytic acid for providing the alpha-proton and that another glutamate serves as the catalytic base for the activation of a water molecule in the hydratase reaction [22]. As shown in Fig. 5, two glutamates (E106 and E108) and four aspartates (D80, D154, D158 and D178) occur at invariant positions and one of these amino acids might represent a potential active site residue. This possibility can be tested by site-directed mutagenesis.

Phylogenetic relationships of the PhnH hydratase. A phylogram showing the relationships between the amino acid sequences of eleven HPDHs and three 2-keto-4-pentenoate hydratases is shown in Fig. 6. The phylogram reveals that HPDHs can be clustered into five subfamilies. Sequences within the same subfamily show greater than 50% identity in pairwise comparisons. The subfamily I includes the TodG, BphE and PhnD hydratases. The subfamily II includes CmtF and three 2-keto-4-pentenoate hydratases from *E. coli* W3110, *E. coli* CS520 and *E. coli* MG1665, which are involved in the *p*-cumate and 3-hydroxyphenylpropionate degradation pathway, respectively. The subfamily III includes BphH, XylJ, BphX1, DmpE and CumE. The subfamily IV and V include HPDH from *Rhodococcus* sp. RHA1 (EtbE) and *Mycobacterium tuberculosis* H37Rv (MTCY03C7.20), respectively.

The results that the PhnH hydratase belongs to subfamily I, but the DmpE and XylJ hydratases belong to subfamily III and that the *phn* operon is different from the *dmp*, *xyl* and *nah* operons in the gene organization suggest that the *phn* operon might have distinct evolutionary lineage from these three operons.

ACKNOWLEDGMENTS

This work was supported by grants BSRI-95-4432 and 96-4432 from Ministry of Education, Korea. We are thankful to Lindsay D. Eltis of University Laval, Canada, for his assistance with phylogenetic analysis of PAUP.

REFERENCES

1. Harayama, S., Mermod, N., Rekik, M., Lehrbach, P. R., and Timmis, K. N. (1987) *J. Bacteriol.* **169,** 558–564.
2. Horn, J. M., Harayama, S., and Timmis, K. N. (1991) *Mol. Microbiol.* **5,** 2459–2474.
3. Shingler, V., Powlowski, J., and Marklund, U. (1992) *J. Bacteriol.* **174,** 711–724.
4. Harayama, S., and Rekik, M. (1990) *Mol. Gen. Genet.* **221,** 113–120.
5. Platt, A., Shingler, V., Taylor, S. C., and Williams, P. A. (1995) *Microbiol.* **141,** 2223–2233.
6. Lau, P. C. K., Bergeron, H., Labbe, D., Wang, Y., Brousseau, R., and Gibson, D. T. (1994) *Gene* **146,** 7–13.
7. Eaton, R. W. (1996) *J. Bacteriol.* **178,** 1351–1362.
8. Kikuchi, Y., Yasukochi, Y., Nagata, Y., Fukuda, M., and Takagi, M. (1994) *J. Bacteriol.* **176,** 4269–4276.
9. Hofer, B., Backhaus, S., and Timmis, K. N. (1994) *Gene* **144,** 9–16.
10. Habe, H., Kimura, T., Nojiri, H., Yamane, H., and Omori, T. (1996) *J. Ferment. Bioeng.* **81,** 247–254.
11. Masai, E., Sugiyama, K., Iwashita, N., Shimizu, S., Hauschild, J. E., Hatta, T., Kimbara, K., Yano, K., and Fukuda, M. (1997) *Gene* **187,** 141–149.
12. Kim, J. W., Kim, C. K., Kim, Y. C., Yeum, J. H., and Lee, J. G. (1987) *Kor. J. Microbiol.* **25,** 122–128.
13. Kim, Y. C., Youn, K. S., Shin, M. S., Kim, H. S., Park, M. S., and Park, H. J. (1992) *Kor. J. Microbiol.* **30,** 1–7.
14. Shin, H. J., Kim, S., and Kim, Y. C. (1997) *Biochem. Biophys. Res. Commun.* **232,** 288–291.

15. Kim, Y. C., Shin, M. S., Youn, K. S., Park, Y. S., and Kim, U. H. (1992) *Kor. J. Microbiol.* **30,** 8–14.
16. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory, New York.
17. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
18. Swofford, D. L. (1993) PAUP: Phylogenetic analysis using parsimony, version 3.1, Computer program distributed by the Illinois Natural History Survey, Champaign, IL.
19. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410.
20. Shine, J., and Dalgarno, L. (1975) *Nature* **254,** 34–38.
21. Zylstra, G. J., Mccombie, W. R., Gibson, D. T., and Fimmette, B. A. (1988) *Appl. Environ. Microbiol.* **54,** 1498–1503.
22. Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K., and Wierenga, R. K. (1996) *EMBO J.* **15,** 5135–5145.