Cloning and Characterization of a 4-Hydroxyphenylacetate 3-Hydroxylase From the Thermophile *Geobacillus* sp. PA-9

J. F. Hawumba[1, 2], V. S. Brözel[1, 3] and J. Theron[1]

1. Department of Microbiology and Plant Pathology, University of Pretoria, 0002 Pretoria, South Africa
2. Biochemistry Department, Makerere University, P. O. Box 7062, Kampala, Uganda
3. Department of Biology and Microbiology, South Dakota State University, Brookings, SD 57007, USA

Abstract

A 4-hydroxyphenylacetic acid (4-HPA) hydroxylase-encoding gene, on a 2.7-kb genomic DNA fragment, was cloned from the thermophile *Geobacillus* sp. PA-9. The *Geobacillus* sp. PA-9 4-HPA hydroxylase gene, designated *hpaH*, encodes a protein of 494 amino acids with a predicted molecular mass of 56.269 Da. The deduced amino-acid sequence of the *hpaH* gene product displayed <30% amino-acid sequence identity with the larger monooxygenase components of the previously characterized two-component 4-HPA 3-hydroxylases from *Escherichia coli* W and *Klebsiella pneumoniae* M5a1. A second oxidoreductase component was not present on the 2.7-kb genomic DNA fragment. The deduced amino-acid sequence of a second C-terminal truncated open reading frame, designated *hpaI*, exhibited homology to extradiol oxygenases and displayed the highest amino-acid sequence identity (43%) with the 3,4-dihydroxyphenylacetate 2,3-dioxygenase of *Arthrobacter globiformis*, encoded by *mndD*. These results, along with catalytic activity observed in crude intracellular extracts prepared from *Escherichia coli* cells expressing *hpaH*, is in support of a role for *hpaH* in the 4-HPA degradative pathway of *Geobacillus* sp. PA-9.

Introduction

Bacteria that are able to use aromatic compounds as sources of carbon and energy for growth have long been considered attractive candidates for detoxification of a wide range of environmental pollutants [*1*, *9*]. Consequently, degradation of aromatic compounds—including phenolics, e.g., phenol, cresol, benzoate and catechol as well as substituted aromatics, e.g., 4-hydroxyphenylacetic acid (4-HPA) and 3,4-dihydroxyphenylacetic acid (3,4-DHPA)—has been studied extensively amongst mesophilic bacteria, and several degradation pathways have been elucidated [*9*]. In contrast, there is relatively little information on the degradation of these compounds by thermophilic bacteria. The degradation of 4-HPA is environmentally important because it is a product of lignin decomposition and is found as an industrial pollutant in wastewater from olive oil production [*8*]. In Gram-positive bacteria, including thermophilic *Bacilli*, 4-HPA is metabolized through a *meta*-cleavage pathway with 3,4-DHPA as the dihydroxylated intermediate and succinate and pyruvate as the final products [*1*]. The initial step in the aerobic catabolism of 4-HPA is carried out by hydroxylases or monooxygenases, which introduce a single hydroxyl group into the phenyl ring. Despite some reports on strains of thermophilic *Bacillus* spp. being capable of degrading aromatic compounds [*5*, *11*, *15*], only three aromatic hydroxylase genes, from strains of *B. stearothermophilus*, *B. thermoleovorans*, and *B. thermoglucosidasius*, have been cloned and sequenced [*10*, *11*, *17*]. In this article, we report on the cloning and characterization of a 4-HPA hydroxylase from the thermophile *Geobacillus* sp. PA-9, and its relation with other enzymes that have similar properties is discussed.

Materials and Methods

Bacterial strains, plasmids, and growth conditions

Geobacillus sp. PA-9, isolated from a hot spring in Western Uganda, was cultured at 55°C in modified Castenholtz medium, as described previously [*16*]. *Escherichia coli* strain DH5α, an *E. coli* K-12 derivative that cannot metabolize 4-HPA and 3,4-DHPA [*4*], was used as the host for cloning procedures and routinely cultured at 37°C in Luria-Bertani (LB) medium (0.5% [w/v] yeast extract, 1% [w/v] tryptone, and 1% [w/v] NaCl;

pH 7.4). Plasmid pSVBI [*19*] was used as cloning vector. To select for recombinants, ampicillin was added to the medium to a final concentration of 100 μg/ml.

Library construction and screening

Genomic DNA from *Geobacillus* sp. PA-9 was isolated according to Wilson [*23*] and partially digested with *Hind*III. DNA fragments (1 to 8 kb) were ligated with *Hind*III-digested, dephosphorylated pSVBI to generate a genomic library. Competent *E. coli* DH5α cells were prepared and transformed by the procedures described by Chung and Miller [*6*]. Transformants containing 4-HPA hydroxylase genes were identified on LB plates containing ampicillin as brown colonies caused by intracellular accumulation of 3,4-DHPA [*13*, *14*].

DNA sequence analysis

The nucleotide sequence of both strands of the cloned hydroxylase-active DNA fragment was determined by automated sequencing with an ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction mixture (Perkin-Elmer Applied Biosystems) in a Hitachi 3100 capillary array automated DNA sequencer. From the sequence obtained, new insert-specific primers were designed to determine the sequence of the full-length insert and to obtain good overlaps in both strands. Homology was analyzed by using BLAST at NCBI (http://www.ncbi.nlm.nih.gov/GenBank/); multiple-sequence alignments were performed with ClustalW (1.84) Multialign programme (http://www.ebi.ac.uk/Tools/clustalw); and the physicochemical parameters of the deduced amino-acid sequence and the presence of defined protein patterns were determined by using the ProSite database at ExPASY (http://www.expasy.org). The sequence data described in this article has been deposited in GenBank under accession number AY549312.

Preparation of cell extracts, enzyme purification, and enzyme activity assays

Cell extracts were prepared, and the 4-HPA hydroxylase was purified from clarified cell lysates (intracellular extracts) by affinity chromatography on a 4-HPA-coupled amino-agarose column (ICN Biochemicals), as described by Raju et al. [*22*]. The molecular mass of the protein was determined on Coomassie brilliant blue–stained 12% sodium

dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) gels [*18*] in the presence of molecular mass markers (ICN Biochemicals). The protein concentration of the extracts and the purified enzyme was determined according to Bradford [*3*] with bovine serum albumin as standard. 4-HPA 3-hydroxylase activity was determined by measuring the liberation of 3,4-DHPA from 4-HPA (Sigma-Aldrich) in 50 mM Tris-HCl buffer (pH 9) at 50°C, according to the method of Anrow [*2*].

Results

Cloning and sequencing the 4-HPA hydroxylase gene from *Geobacillus* sp. PA-9

A recombinant clone, *E. coli*/pSVBI-R113, was obtained by screening a genomic library of *Geobacillu*s sp. PA-9 on LB/Amp agar plates and observing for 3,4-DHPA accumulation, as indicated by the presence of a diffusible brown pigment resulting in colonies with a brown discolouration. *E. coli* DH5α, containing only the vector pSVBI, did not yield similar-coloured colonies. The nucleotide sequence of the cloned 2.7-kb DNA fragment was determined, and two colinear open reading frames (ORFs), of which one was truncated at its C-terminus, were identified. The ORFs were separated by 406 bp, and both were preceded by potential Shine-Dalgarno sequences and putative promoter sequences, thus suggesting the occurrence of independent transcription and translation. The first ORF, encompassing 1,485 bp, encodes a protein of 494 amino acids with a theoretical molecular mass of 56.269 Da, and the deduced amino-acid sequence contained an HpaB domain, which is conserved among the HpaB family of 4-HPA 3-hydroxylase enzymes. The new enzyme was designated HpaH, and the corresponding gene was designated *hpaH*. The second ORF, designated as *hpaI*, was incomplete because no TGA stop codon could be identified. The N-terminal amino-acid sequence (178 residues) derived from *hpaI* exhibited highest homology to the putative 3,4-dihydroxyphenylacetate 2,3-dioxygenase of *G. kaustophilus* HTA426 (93%; accession number YP_148886) and displayed 43% sequence identity to the characterized enzyme encoded by *mndD* of *Arthrobacter globiformis* (accession number AAA67362). Therefore, *hpaI* can be postulated to encode for an extradiol dioxygenase enzyme that forms part of the 4-HPA degradative pathway of *Geobacillus* sp. PA-9.

Deduced amino-acid sequence of HpaH and homology with other 4-HPA 3-hydroxylase sequences

Comparison of the deduced amino-acid sequence of *hpaH* by BLAST-P search with the sequences in the GenBank database showed similarity to putative 4-HPA 3-hydroxylases of several bacteria of which the genome sequences have recently been completed. Pairwise sequence alignments showed that HpaH of *Geobacillus* sp. PA-9 has the highest amino-acid sequence identity to the putative 4-HPA 3-hydroxylases enzymes of *G. kaustophilus* HTA426 (96%; accession number YP_148887) and *Oceanobacillus iheyensis* HTE831 (67%; accession number NP_693794), whereas lower identities were found with the enzymes of *Bacillus cereus* subsp. *cytotoxis* (55%; accession number ZP_01179289); *Thermus thermophilus* HB8 (48%; accession number YP_144226); *B. halodurans* C-125 (43%; accession number NP_244703); and *B. subtilis* 168 (40%; accession number NP_389743). A comparison of HpaH with previously characterized 4-HPA 3-hydroxylases showed that HpaH shares 28% sequence identity with the equivalent 4-HPA monooxygenases of *E. coli* W ATCC 11105 (encoded by *hpaB*; accession number CAA86048) [*20*] and *Klebsiella pneumoniae* M5a1 (encoded by *hpaA*; accession number Q48440) [*14*]. Furthermore, HpaH shares 31% and 29% sequence identity, respectively, with the phenol 2-hydroxylases of *B. thermoglucosidasius* A7 (encoded by *pheA1*; accession number AAF66546) [*10*] and *B. thermoleovorans* A2 (encoded by *pheA*; accession number AAC38324) [*11*]. As with the aforementioned enzymes, the FAD- and NAD-binding signature sequences—GXGXXG and [TM]XXXX[IVAL][YWF][IVAL][IVA]GD, respectively—were not detected in the HpaH sequence of *Geobacillus* sp. PA-9. Multiple alignment of the deduced amino-acid sequence from *hpaH* of *Geobacillus* sp. PA-9 with the previously characterized aromatic hydroxylases is shown in Fig. 1.

Fig. 1 Alignment of the amino-acid sequence of the HpaH hydroxylase of *Geobacillus* sp. PA-9 with the amino-acid sequences of other characterized aromatic hydroxylases. Alignment was maximized by introducing gaps, which are indicated by dashes. Identical (*), highly similar (:) and similar (.) amino acids are indicated. G. PA-9_HpaH, HpaH from *Geobacillus* sp. PA-9 (this study); G. thermogl_PheA1, phenol 2-hydroxylase from *G. thermoglucosidasius* A7 (AAF66546); G. thermoleov_PheA, phenol 2-hydroxylase from *G. thermoleovorans* A2 (AAC38324); *E. coli*_HpaB, 4-HPA 3-monooxygenase from *E. coli* W (CAA86048); *K. pneum*_HpaA, 4-HPA 3-monooxygenase from *K. pneumoniae* M5a1 (Q48440)

Purification of HpaH

Cell extracts from *E. coli*/pSVBI-R113 were prepared and tested for 4-HPA hydroxylase activity using 4-HPA as substrate. The highest activity was observed in the intracellular fraction, whilst no activity was detectable in the cell-free culture supernatant. In contrast, a control nonrecombinant *E. coli* DH5α culture did not show 4-HPA hydroxylase activity in identically prepared fractions. Purification of hydroxylase activity from intracellular

extracts of *E. coli*/pSVBI-R113 cells rendered a protein that was purified to near homogeneity (Fig. 2). The purified enzyme showed an apparent molecular mass on SDS-PAGE of 56 kDa, which is in agreement with the calculated molecular mass (56.269 kDa) of HpaH. Compared with intracellular cell extracts, the purified *Geobacillus* sp. PA-9 HpaH protein did not display detectable activity toward the 4-HPA substrate under the assay conditions. This result is in agreement with that reported for the purified HpaB protein of *E. coli* W and may be caused by its low stability and the complex purification procedure [*12*, *20*].

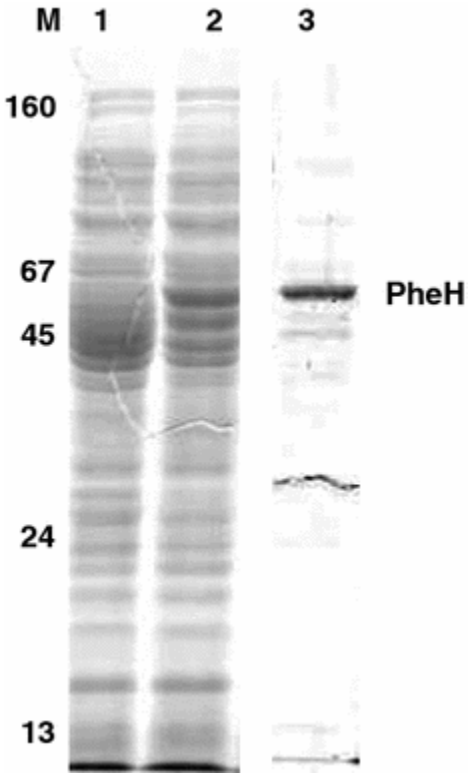

Fig. 2 SDS-PAGE analysis of the expression and purification of the HpaH protein of *Geobacillus* sp. PA-9. Lane 1, intracellular cell extract sample from nonrecombinant *E. coli* DH5α; lane 2, intracellular cell extract sample from *E. coli*/pSVBI-R113; lane 3, sample of the affinity chromatography-purified HpaH protein. The sizes of the molecular mass markers (in kDa) are shown to the left of the figure

Discussion

Sequence analysis of the 4-HPA hydroxylase from *Geobacillus* sp. PA-9, designated HpaH, demonstrated homology to the equivalent HpaB and HpaA monooxygenases from *E. coli* W and *K. pneumonia*e M5a1, respectively, and to the phenol hydroxylase PheA1 from *B. thermoglucosidasius* A7. These three proteins are each part of two-component aromatic hydroxylase enzyme systems that are comprised of a monooxygenase and a smaller flavin:NAD(P)H oxidoreductase, and the genes encoding the two components of these enzymes are located in the same operon [*10*, *14*, *20*]. No additional ORF with the potential of encoding an oxidoreductase was found in the intergenic region between the *hpaH* and *hpaI* genes in the 2.7-kb genomic DNA fragment cloned from *Geobacillus* sp. PA-9. There are, however, several exceptions to this arrangement, and the gene encoding the reductase component can thus be located on the genome far from the genes encoding the monooxygenase components [*12*, *24*]. Interestingly, a comparative analysis of the *G. kaustophilus* HTA426 genome showed an ORF (GK3021; accession number NC_006510) encoding a protein displaying 59% and 27% amino-acid sequence identity with the PheA2 and HpaC reductases of *B. thermoglucosidasius* and *E. coli* W, respectively. GK3021 is located upstream of a gene cluster containing two contiguous genes that encode for putative 4-HPA 3-hydroxylase (GK3034) and 3,4-DHPA 2,3-dioxyenase (GK3033) enzymes. These proteins share 96% and 93% sequence identity with HpaH and HpaI from *Geobacillus* sp. PA-9, respectively. It is therefore tempting to speculate that *Geobacillus* sp. PA-9 may similarly posses an oxidoreductase-encoding gene that is located distantly from the gene encoding HpaH.

Compared with the purified protein, expression of the cloned HpaH enzyme of *Geobacillus* sp. PA-9 yielded detectable 4-HPA hydroxylase activity in crude intracellular enzyme extract preparations, despite the *E. coli* host strain used not being able to metabolize 4-HPA [*4*]. It has been reported for the 4-HPA hydroxylase of *E. coli* W that the HpaB oxygenase component does not require a direct interaction with the HpaC oxidoreductase to hydroxylate 4-HPA [*12*] and that any host cell-encoded flavin reductase able to release $FADH_2$ into the cytoplasm can replace the role of HpaC [*12*, *25*]. Indeed, the hydroxylase activity initially observed in *E. coli* K-12 strains expressing

the oxygenase HpaB component alone [*20*] has since been ascribed to the presence of a host-encoded flavin reductase [*12*]. Although no biochemical data on the cofactor requirements of HpaH activity are available, the lack of a flavin:NAD(P)H reductase center in the HpaH protein suggests that it is likely to require an independent reductase enzyme that could provide the activity. Based on the existence of several reductases in *E. coli* capable of producing free reduced flavins and the reported functional interchangeability between them [*7*, *12*], the results obtained may indicate that the observed 4-HPA hydroxylase activity of HpaH in crude intracellular enzyme extracts is caused by a host cell-encoded oxidoreductase. However, this assumption must be confirmed by future biochemical assays.

Expression of both components of the *E. coli* W [*21*] and *K. pneumoniae* [*13*] 4-HPA hydroxylases in *E. coli* K-12 derivative strains has been reported to result in the production of brown to black pigments in the culture medium. Similarly, growth of *E. coli*/pSVBI-113 in LB/Amp medium, in the absence of 4-HPA, yielded a brown pigment in the medium that did not appear when untransformed *E. coli* cells were grown in the same medium. These results therefore indicated that formation of the pigment results from the catabolic activity of the cloned 4-HPA hydroxylase and suggests that HpaH might thus recognize other substrates distinct from 4-HPA. Such relaxed substrate specificity is in agreement with results reported for the 4-HPA 3-hydroxylases of *E. coli* [*21*] and *K. pneumoniae* [*13*, *14*]. The 4-HPA hydroxylase described in this article is the first from a thermophile species to be analyzed at the genetic level. Further molecular genetic studies of this new hydroxylase will contribute in obtaining new insight into the biodegradation and biotransformation of aromatic compounds by thermophilic Gram-positive bacteria.

References

1. Ali S, Fernandez-Lafuente R, Cowan DA (1998) *Meta*-pathway degradation of phenolics by themophilic Bacilli. Enzyme Microbial Technol 23:462–468

2. Anrow EL (1937) Colorimetric determination of the components of 3,4-dihydroxyphenylalanine-tyrosine mixture. J Biol Chem 118:531–537

3. Bradford M (1976) A rapid and sensitive method for quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254

4. Burlingame R, Chapman PJ (1983) Catabolism of phenylpropionic acid and its 3-hydroxy derivative by *Escherichia coli*. J Bacteriol 155:113–121

5. Buswell JA, Twomey G (1975) Utilization of phenol and cresols by *Bacillus stearothermophilus*, strain PH24. J Gen Microbiol 87:377–379

6. Chung TC, Miller HR (1993) Preparation and storage of competent *Escherichia coli* cells. Methods Enzymol 218:621–627

7. Covès J, Nivière V, Eschenbrenner M, Fontecave M (1993) NADPH-sulfite reductase from *Escherichia coli*. A flavin reductase participating in the generation of the free radical of the ribonucleotide reductase. J Biol Chem 268:18604–18609

8. Crawford DL, Barder MJ, Pometto AL, Crawford RL (1982) Chemistry of softwood lignin degradation by *Streptomyces viridosporus*. Arch Microbiol 131:140–145

9. Díaz E, Ferrández A, Prieto MA, García JL (2001) Biodegradation of aromatic compounds by *Escherichia coli*. Microbiol Mol Biol Rev 65:523–569

10. Duffner FM, Kirchner U, Bauer MP, Müller R (2000) Phenol/cresol degradation by the thermophilic *Bacillus thermoglucosidasius* A7: Cloning and sequence analysis of five genes involved in the pathway. Gene 256:215–221

11. Duffner FM, Müller R (1998) A novel phenol hydroxylase and catechol 2,3-dioxygenase from the thermophilic *Bacillus thermoleovorans* strain A2: Nucleotide sequence and analysis of the genes. FEMS Microbiol Lett 161:37–45

12. Galán B, Díaz E, Prieto MA, García JL (2000) Functional analysis of the small component of the 4-hydroxyphenylacetate 3-monooxygenase of *Escherichia coli* W: A prototype of a new flavin:NAD(P)H reductase subfamily. J Bacteriol 182:627–636

13. Gibello A, Ferrer E, Sanz J, Martín M (1995) Polymer production by a *K. pneumoniae* 4-HPA hydroxylase gene cloned in *E. coli*. Appl Environ Microbiol 61:4167–4171

14. Gibello A, Suárez M, Allende JL, Martín M (1997) Molecular cloning and analysis of the genes encoding the 4-hydroxyphenylacetate hydroxylase from *Klebsiella pneumoniae*. Arch Microbiol 167:160–166

15. Gurujeyalakshmi G, Oriel P (1989) Isolation of phenol-degrading *Bacillus stearothermophilus* and partial characterization of phenol hydroxylase. Appl Environ Microbiol 55:500–502

16. Hawumba FJ, Theron J, Brözel SV (2002) Thermophilic protease-producing *Geobacillus* from Buranga hot springs in western Uganda. Curr Microbiol 45:144–150

17. Kim IC, Oriel PJ (1995) Characterisation of a *Bacillus stearothermophilus* BR219 phenol hydroxylase gene. Appl Environ Microbiol 61:1252–1256

18. Laemmli UK (1970) Cleavage associated with the maturation of the head of bacteriophage T4. Nature 227:680–685

19. Lee J-S, Wittchen K-D, Stahl C, Strey J, Meinhardt F (2001) Cloning, expression, and carbon catabolite repression of the bam gene encoding β-amylase of *Bacillus megaterium* DSM319. Appl Microbiol Biotechnol 56:205–211

20. Prieto MA, García JL (1994) Molecular characterization of phenol 4-hydroxyphenylacetate 3-hydroxylase of *Escherichia coli*. A two-protein component system. J Biol Chem 269:22823–22829

21. Prieto MA, Pérez-Aranda A, García JL (1993) Characterization of an *Escherichia coli* aromatic hydroxylase with a broad substrate range. J Bacteriol 175:2162–2167

22. Raju GS, Kamath AV, Vaidyanathan SC (1988) Purification and properties of 4-hydroxyphenylacetic acid 3-hydroxylase from *Pseudomonas putida*. Biochem Biophys Res Commun 154:537–543

23. Wilson K (1990) Preparation of genomic DNA from bacteria. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, et al. (eds) Current protocols in molecular biology. New York, NY, Wiley Interscience, pp 2.4.1–2.4.5

24. Xi L, Squires CH, Monticello DJ, Childs JD (1997) A flavin reductase stimulates DszA and DszC proteins of *Rhodococcus erythropolis* IGTS8 *in vitro*. Biochem Biophys Res Commun 230:73–75

25. Xun L, Sandvik ER (2000) Characterization of 4-hydroxyphenylacetate 3-hydroxylase (HpaB) of *Escherichia coli* as a reduced flavin adenine dinucleotide-utilizing monooxygenase. Appl Environ Microbiol 66:481–486