Investigation of Two Evolutionarily Unrelated Halocarboxylic Acid Dehalogenase Gene Families

KATJA E. HILL, JULIAN R. MARCHESI,* AND ANDREW J. WEIGHTMAN
Cardiff School of Biosciences, Cardiff University, Cardiff, CF1 3TL, Wales, United Kingdom

Received 1 September 1998/Accepted 29 January 1999

Dehalogenases are key enzymes in the metabolism of halo-organic compounds. This paper describes a systematic approach to the isolation and molecular analysis of two families of bacterial α-halocarboxylic acid (αHA) dehalogenase genes, called group I and group II deh genes. The two families are evolutionarily unrelated and together represent almost all of the αHA deh genes described to date. We report the design and evaluation of degenerate PCR primer pairs for the separate amplification and isolation of group I and II deh genes. Amino acid sequences derived from 10 of 11 group I deh partial gene products of new and previously reported bacterial isolates showed conservation of five residues previously identified as essential for activity. The exception, DehD from a Rhizobium sp., had only two of these five residues. Group II deh gene sequences were amplified from 54 newly isolated strains, and seven of these sequences were cloned and fully characterized. Group II dehalogenases were stereoselective, dechlorinating L- but not D-2-chloropropionic acid, and derived amino acid sequences for all of the genes except dehIIb_P1 showed conservation of previously identified essential residues. Molecular analysis of the two deh families highlighted four subdivisions in each, which were supported by high-bootstrap values in phylogenetic trees and by enzyme structure-function considerations. Group I deh genes included two putative cryptic or silent genes, dehF_P1 and dehFb_P1 produced by different organisms. Group II deh genes included two cryptic genes and an active gene, dehII_P1, that can be switched off and on. All αHA-degrading bacteria so far described were Proteobacteria, a result that may be explained by limitations either in the host range for deh genes or in isolation methods.

The biosphere contains a multitude of halogenated organic compounds, more than 2,400 of which have been identified as occurring naturally; however, those constituting the bulk quantities are synthesized industrially (13). Many halo-organic compounds have been categorized as priority pollutants (8), even though a wide range of bacterial species that can degrade such substances and, in many cases, utilize them as sole sources of carbon and energy have been isolated in laboratory culture (11, 21). Notwithstanding the recalcitrance of halo-organic compounds in the biosphere, microbial catabolism is clearly a major latent route by which these compounds may be detoxified and recycled. Therefore, we need to understand much more about the process of microbial adaptation involved in order to harness this potential.

Dehalogenation is a key reaction in such recycling, and a variety of microbial enzymes which catalyze carbon-halogen bond cleavage have been described (11, 21, 56). α-Halocarboxylic acids (αHAs) were originally listed on the United Kingdom Department of the Environment’s “Red List” and have also been identified as intermediates in the biodegradation of halogenated solvents such as 1,2-dichloroethene (20). The pioneering studies of Jensen (23) and Goldman and colleagues (12) resulted in the identification and characterization of hydrolytic dehalogenases associated with the catabolism of αHAs. In the last decade, some 18 dehalogenase (deh) genes have been cloned and sequenced (24, 28, 30, 40, 41, 52, 64). However, despite the availability of microbial data, even recent attempts to classify dehalogenases have tended to focus on arbitrary characteristics such as substrate specificity, especially with optically active substrates such as 2-monochloropropionic acid (2MCPA) (16, 29, 55). Koonin and Tatusov (32) suggested that at least some αHA dehalogenases were evolutionarily related to a group of hydrolases, which they called the HAD superfamily. All of the HAD superfamily dehalogenases were active with the L- but not the D-isomer of 2MCPA, and Kurihara et al. (33) referred to them as the L-Dex family.

The evolution of αHA dehalogenases is of interest in terms of understanding the origin of novel enzyme activities and the adaptation of bacteria to degrade xenobiotic compounds. In this respect, it is important to establish the true evolutionary relationships between dehalogenase genes and to develop methods by which adaptive processes involving deh genes can be studied in the natural environment. This paper describes a genetic approach to investigate the diversity and molecular ecology of deh genes. The aim was to establish a molecular phylogenetic classification that would provide a solid framework for studies of the adaptation of bacteria to degrade halogenated aliphatic compounds. Degenerate oligonucleotide primers for PCR amplification of two distinct families of deh genes were designed and extensively evaluated. The use of these primers for isolation and characterization of active and silent (cryptic) deh genes, and their wider utility, is described.

MATERIALS AND METHODS

Bacterial strains, plasmids, media, and culture conditions. Over 50 bacterial strains were isolated independently by enrichment culture with one of several carbon sources, enrichment source materials, and isolation temperatures. Enrichments were done with either SBS (53) or Brunner’s (6) minimal enrichment medium. Burkholderia (β subclass of Proteobacteria) sp. strains G02, I11, and K13 were isolated from separate enrichment cultures on monochloroacetic acid (MCA) (0.5 g of C liter⁻¹) at 30°C by using three different bulk soil sources (natural woodland, rosebed, and cultivated woodland soils). The following strains were isolated from grass rhizosphere soil: Burkholderia sp. strain P11 (β subclass of Proteobacteria), isolated on 2,3-dichloropropionic acid (23DPA) (0.5 g of C liter⁻¹) at 30°C, and strain DA1 (β subclass of Proteobacteria) and Bradyrhizobium sp. strains DA2 and DA3 (α subclass of Proteobacteria) isolated on 2,2-dichloropropionic acid (22DPA) (0.5 g of C liter⁻¹) at 20°C. Pseudomonas sp.
obtained with primer pairs dehIfor1-dehIrev2 and dehIIfor1-dehIIrev1 were cloned,
by 36 cycles of 94°C for 45 s, 55°C for 2 min, and 75°C for 45 s, with a final genes. The PCR involved an initial denaturation step at 94°C for 10 min followed and
Cambridge, United Kingdom). This fragment contains an active dehalogenase sequence derived from an alignment of the following genes:

Pseudomonas DEX gene from PP3 genomic DNA ligated into the pBluescript II KS(1)

DehIfor1 and dehIrev2 amplified a 504-bp product (Fig. 1).

Oligo version 3.4 (National Biosciences Inc.) from a consensus sequence align-

Taq cillin ml

transformants was done on Luria-Bertani agar (39) containing 50
ligation with plasmid vector pGEM-T Easy (Promega, Southampton, United
were purified by using the Qiagen Qiaquick PCR purification kit and cloned by
pYW10 was isolated by using Qiagen (Crawley, United Kingdom) Qiaquick
strain 17a was isolated from river epilithon on MCA, also at 4°C. Once purified,
ist of 2MCPA at 15 and 4°C, respectively, and
strains K55 and 18a (subclass of

(50). Both strands were sequenced to ensure accuracy. Partial
erwise, protocols for DNA manipulation were as described by Sambrook et al.

extracts were carried out as described by Thomas et al. (59). Dehalogenase assays
assays was estimated by coulometric titration of free halide in samples by using


A. or T = C or T, N = other nucleotides identical to those of the primer, and indicated by
de used as template in PCR evaluation of group I

de used as template for PCR primer sequences, showing comparisons with binding sites in
dehr-type genes from various sources


...T.CG.C.....C..... .T..GA..T..C.....G..G.. C..C.TGG.CTT.T.G....

ACGYTNSGSGTGCCNTGGGT

AWCARRTAYTTYGGATTRCCRTA

G, C, or T; D

5

A, C, G, or T; R

5

Gene used for design of group I
dehr primers.

dehr gene used as template in PCR evaluation

Gene used as template in PCR evaluation of group I
dehr primers.

dehr primers.

DNA extraction and manipulation of group II
dehr sequences

DNA extraction and manipulation of group I
dehr sequences
obtained from EMBL (release 55). Bootstrap analysis (9) of up to 500 replicates
was performed on the neighbor-joining method (49) with TREECON for Windows (63).

RESULTS
16S rRNA gene analysis of oHA-degrading bacterial isolates. Batch enrichment cultures were used to isolate oHA-utilizing bacteria from three soils. Eight different strains were isolated and purified on solid media with one of the following oHAs as the sole source of carbon and energy: MCA, 2MCPA, 22DCPA, and 23DCPA. Identification of the bacterial isolates, based on 16S rRNA gene analysis (43), showed that they were all Proteobacteria. Strains DA2 and DA3 were Bradyrhizobium species (α subclass of Proteobacteria); strains DA1, G02, 111, K13, and P11 were members of the β subclass of Proteobacteria (all assigned to the genus Burkholderia except strain DA1, which remains unassigned); and strains 17a, 18a, and K55 were Pseudomonas sensu stricto (γ subclass of Proteobacteria).

PCR amplification of group I dehalogenase (dehI) genes. An alignment of four related sequences (dehI, dehII, dhlA, and hadD) was used to design conserved regions as potential annealing sites for PCR primers that would enable amplification of these and other related oHA dehalogenase (dehI) genes. Three such regions were identified and used to design two sets of degenerate PCR primers: dehI for1 and dehI rev2.

<table>
<thead>
<tr>
<th>Table 2. Group II deh PCR primer sequences, showing comparisons with binding sites in dehI-type genes from various sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>dehI-type gene</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>dehI for1</td>
</tr>
<tr>
<td>dehI rev2</td>
</tr>
<tr>
<td>DehH109 gene</td>
</tr>
<tr>
<td>t-DEX gene</td>
</tr>
<tr>
<td>DehH2 gene</td>
</tr>
<tr>
<td>dhlA</td>
</tr>
<tr>
<td>dehCI</td>
</tr>
<tr>
<td>dehCIF</td>
</tr>
<tr>
<td>hddIV</td>
</tr>
</tbody>
</table>

References: hadl, 24; DehH109 gene, 28; t-DEX gene, 41; DehH2 gene, 27; dhlA, 64; dehCI and dehCII, 52; hadD, 40.

a S = +. The predicted PCR product size is 422 bp for all dehI genes except dhlB (416 bp). IUPAC ambiguity code used: B = C, G, or T; D = A, G, or T; M = A or C; N = A, C, G, or T; R = A or G; S = C or G; V = A, C, or G. Nucleotides identical to those of the primer are indicated by dots.

b Gene used for design of group II deh primers.

c A plasmid-encoded dehalogenase.
with either dehIrev1 or dehIrev2 (Table 1). It should be noted that primers dehIfor1 and dehIrev1 were both 128-fold degenerate, while dehIrev2 was 2,048-fold degenerate. The larger PCR product (504 bp, obtained with dehIfor1 and dehIrev2) was used in this study because of its higher information content (approximately 56% of the dehI gene). Testing of the primers was initially carried out with template DNAs from P. putida AJ1, which contains hadD (24), and P. putida PP3, containing dehI (59). In each case PCR products of the expected sizes (Fig. 1) and sequences (see database entries noted in Materials and Methods) were obtained. It should be noted that minor errors in the published sequence of hadD (3) were identified, and the published derived amino acid sequence had to be corrected to account for these (Fig. 2).

Primers dehIfor1 and dehIrev2 were further tested on a range of bacteria enriched and isolated on αHAs, all of which were shown to produce αHA dehalogenases by native polyacrylamide gel electrophoresis (gel zymography). All PCR products were sequenced, and the sequences were compared with the published sequences of the dehI and hadD genes. The results were consistent with the hypothesis that the dehI and hadD genes are present in these bacteria.

FIG. 2. Alignment of deduced amino acid sequences of the group I Deh proteins. The conserved residues, described by Nardi-Dei et al. (42), are underlined and in boldface. The conserved amino acid motif, YGNPKY, is shown in boldface. The group I dehalogenase gene PCR priming sites dehI for1, dehIrev1, and dehIrev2 are shown in the three shaded areas. The deduced amino acid sequence for HadD is based on the resequenced version (see text).
were cloned and sequenced, and a positive result was recorded only when products from at least two replicate PCRs were shown to be identical and homologous to deh genes in this group. A small number of PCR products of the expected size were found after sequencing to be obviously unrelated to dehI, even though they were amplified consistently from some bacterial isolates. These were excluded from further consideration in the present study. Five of 20 new bacterial isolates tested positive with the dehI PCR primers: strains 17a, K55, 18a, DA1, and DA2. Analysis of the dehalogenase genes thus obtained indicated that they were evolutionarily related in terms of both nucleotide and derived amino acid sequence alignments. Thus, the dehI PCR primers were successful in amplifying a family of related deh genes, designated group I deh genes, from a variety of bacteria.

Unexpectedly, two different group I deh genes were amplified from P. putida PP3. One was expected: the dehI that resides on a transposable element previously described by Thomas et al. (59) and Topping et al. (62). The other, designated dehIppr, was located immediately upstream of dehIppr, a group II deh, and was cloned separately from dehIppr on a 12-kbp HindIII restriction fragment of P. putida PP3 genomic DNA (see below).

**Phylogenetic analysis of group I deh genes.** Figure 3 shows a phylogenetic tree illustrating the relationships between all known group I deh genes, based on a CLUSTAL W (61) nucleotide alignment over the whole region amplified with primers dehI for1 and dehIrev2 (504 bp). Almost identical tree topologies were obtained whether they were constructed from alignments done by using maximum-parsimony, maximum-likelihood, or Jukes-Cantor algorithms (25). The four subdivisions identified from this tree were supported by high bootstrap values (Fig. 3), as well as by analysis of alignments with the deduced amino acid sequences for the dehalogenase gene products (data not shown). The gene products of the DL-DEX gene (subdivision B) and dhlIV (subdivision A) have basic structural similarities in that they are both homodimers made up of polypeptides of between 32 and 35 kDa, whereas HadD (subdivision C) is a tetrameric holoenzyme (57).

Nardi-Dei et al. (42) identified five amino acids in DL-DEX (T65, E69, N117, Y120, and D194) encoded within the region amplified by dehI for1 and dehIrev2 (504 bp). Almost identical tree topologies were obtained whether they were constructed from alignments done by using maximum-parsimony, maximum-likelihood, or Jukes-Cantor algorithms (25). The four subdivisions identified from this tree were supported by high bootstrap values (Fig. 3), as well as by analysis of alignments with the deduced amino acid sequences for the dehalogenase gene products (data not shown). The gene products of the DL-DEX gene (subdivision B) and dhlIV (subdivision A) have basic structural similarities in that they are both homodimers made up of polypeptides of between 32 and 35 kDa, whereas HadD (subdivision C) is a tetrameric holoenzyme (57).
Design and evaluation of group II deh PCR amplification primers. Alignments of five known deh gene sequences, related to each other but showing no obvious homology to the group I deh family (Table 2), were used to identify four conserved regions that were analyzed as potential binding sites for universal PCR primers for this group of genes. Primers dehII_for1 and dehII_rev1 were designed to match the two most highly conserved potential primer-binding sites and were chosen for further evaluation. However, a potential concern was the degeneracy of these primers, since it was even greater than that of the group I deh primer pair and allowed 864 and 6,912 possible target sequence permutations with dehII_for1 and dehII_rev1, respectively. In Table 2 these primers’ sequences are compared with corresponding target sites on nine deh genes. The size of the PCR products predicted from all of these templates except dhlB (416 bp [54.8% of the gene]) was 422 bp, representing 60.1 to 63.4% of the gene. Initial testing of PCR primer pair dehII_for1-dehII_rev1 was carried out with genomic DNA templates from four bacteria, representing five deh genes: P. putida PP3 (dehII), P. putida AJ1 (hadL), Pseudomonas sp. strain CBS3 (dehCII and dehCII), and Burkholderia cepacia MBA4 (dhlIVA). In each case, products of the expected size were obtained (see, e.g., Fig. 1), and following their cloning and sequencing from replicate PCRs, they were confirmed as originating from the corresponding dehII-type templates. For example, two products of identical size, corresponding exactly to the DNA regions between the dehII_for1 and dehII_rev1 binding sites shown in Table 2 for dehCII and dehCII, were amplified and cloned from Pseudomonas sp. strain CBS3 template DNA. It should be noted that single-base mismatches between the group II deh primers and target sites on three of the test dehII-type genes (dehII, hadL, and dhlIVA) did not prevent the amplification of the expected DNA fragments. The complete sequence of dehII from P. putida PP3 has not been reported before, and this gene was clearly shown to be a member of this group, as was the unpublished dhlVII (19), with which it had 98.8% nucleotide sequence identity. The dehII_for1 gene was separately isolated on a 12-kbp HindIII restriction fragment, cloned to produce plasmid pYW2, and subsequently subcloned on a 3.0-kbp BamHI fragment to produce pYW10.

As might be expected from the use of such highly degenerate primers, some PCR artifacts were observed, usually as amplified DNA fragments outside the expected size range for group II deh genes. However, even when testing was extended to newly isolated αHA-degrading bacteria and mixed enrichment cultures (see below), such artifacts occurred only infrequently, were easily identified (confirmed by sequencing), and were eliminated from this study. Further comparisons between the predicted primer-binding sites in the five control dehII-type genes (Table 2) and the observed sequences of PCR product primer ends, obtained from replicate reactions, showed an unexpectedly high number of mismatches, i.e., up to 5 of 23 bp for dehII_for1 and dehII_rev1, respectively. These data suggested that optimal primers from the degenerate mixture (7,776 different oligonucleotides, providing approximately 6 million possible PCR primer pair combinations) were not necessarily being selected in the annealing step of PCR. A likely explanation would be that optimal primers were used up in the early cycles of PCR, but as they were depleted, oligonucleotides with suboptimal matching to the primer-binding sites had sufficient homology to compete and prime elongation in the final cycles of the reaction. It should be noted that unusually high concentrations (2 μM) of the group II deh primers were used in PCRs so as to reduce this problem.

Identification of group II deh genes in newly isolated αHA-degrading bacteria. Overall, 54 αHA-degrading bacteria were isolated from independent batch culture enrichments with combinations of eight soil or sediment samples and three different αHA substrates. PCR with the primer pair dehII_for1-dehII_rev1 resulted in 43 isolates (i.e., 80%) testing positive for group II
deh. Eight of these, including strains 17a, 18a, DA2, and K55 (Fig. 3), also gave a positive result with group I deh primers. Only two isolates, strains J14 and DA5, did not give rise to a PCR product with either group I or group II deh primers.

Group II deh PCR products from seven independently isolated aHA-utilizing bacteria were cloned and fully sequenced. Figure 4 shows a phylogenetic tree illustrating the relationships between the deh gene sequences isolated from these bacteria and the other group II deh genes previously reported in the literature (Table 2). The phylogenetic trees constructed by the neighbor-joining program from a similarity matrix of pairwise comparisons made by using the Jukes-Cantor algorithm (25). Bootstrap values from 100 replicate trees are shown at the dendrogram nodes. *, cryptic deh gene. The scale bar represents sequence divergence (0.1 = 10%). Accession numbers and/or references for published aHA genes not already cited in Table 2: dhlS5I0 (30) (not deposited in EMBL), dhlII, X94147 (10); cbbZp (2-phosphoglycolate phosphatase), M68905 (51). *, dehalogenase genes from reference strains for which PCR products were cloned and sequenced in this study. A, B, C, and D are subdivisions based on >55% nucleotide sequence identity and supported by high bootstrap values as shown.

**Conservation of structure and function in group II dehalogenases.** Figure 5 shows an alignment of amino acid sequences derived from all of the group II deh genes. The alignment shows the three conserved secondary-structure motifs defining the HAD superfamily (32). The group II deh PCR primers were designed to amplify a region of the deh gene containing only motif II and part of motif III; however, these are clearly conserved in all but one case, that of DhlVII. The crystal structures of L-DEX and DhlB have both been solved, and active-site amino acids have been identified (18, 33, 48). Figure 5 shows that all of these residues are conserved, except in the cases of DhlVII, which lacks R41 (cf. L-DEX [18]), and DehII\_p11, which has a stop codon instead of Y157 (cf. L-DEX). It should be noted that Burkholderia sp. strain P11 was isolated on 23DCPA, and no aHA dehalogenase activity was detected in this strain (Fig. 6, lane 9), suggesting that its group II deh gene may be cryptic. F60, which was suggested by Li et al. (35) to be associated with the substrate-stabilizing hydrophobic pocket of L-
DEX, is also conserved in all of the group II deh-derived sequences.

Dehalogenases from the newly isolated cHA-degrading bacteria were characterized by assays and gel zymography. Figure 6 shows, not unexpectedly, that some of the strains produced more than one dehalogenase, but group II Deh proteins could be distinguished from group I Deh proteins by their stereoselective dechlorination of L-2MCPA but not D-2MCPA. 

**Figure 5.** Alignment of derived amino acid sequences of bacterial group II deh genes from the GenBank-EMBL database or from a sequenced cloned gene (dehIIPP3) or partial dehalogenase sequences derived from PCR products (dehIIK13, dehIIH11, dehIIpp3, dehIIK55, dehIIK13, and dehIIK55) by using the CLUSTAL W program (where the X in DehII indicates the strain from which the sequence is derived). The regions corresponding to the group II dehalogenase gene PCR priming sites dehIIfor1 and dehIIrev1 are shown as shaded areas, and the three conserved secondary structure motifs I, II, and III (32) of the HAD superfamily are boxed. Nine residues essential for catalysis for the group II dehalogenases are shown underlined in boldface, based on L-DEX, as follows (18, 33): D10, T14, R41, S118, K151, Y157, S175, N177, and D180. Conserved amino acids forming the hydrophobic pocket (35) are shown in boldface: Y12, Q42, L45, F60, K151, N177, and W179.
tigated by gel zymography, using the approach described above and in the legend to Fig. 6 to distinguish between group I and group II dehalogenases.

As indicated in Table 3, all except one of the bacteria containing group I deh genes produced at least one dehalogenase with activity against D-2MCPA. The exception, strain 17a, thus contained a dehI-type gene, assigned to subdivision C, that may be cryptic, although it should be noted that enzyme instability may have accounted for the absence of a group I dehalogenase with activity against D-2MCPA. Two other genes, dehI°PP3 and dehI°K55, almost identical to dehI°17a (99% nucleotide sequence identity) but from different bacterial isolates, clustered with hadD in subdivision C (Fig. 3). Whereas HadD, produced by Pseudomonas sp. strain AJ1 (Fig. 6, lane 3), has activity with D- but not L-2MCPA (3, 57), the only dehalogenase produced by Pseudomonas sp. strain K55 showing activity with D-2MCPA also had activity with L-2MCPA. The organism from which dehI°PP3 was isolated, P. putida PP3, has been studied extensively and found to produce only two dehalogenases, DehIPP3 and DehIIPP3 (54, 66), as indicated in Fig. 6 (lane 2). The dehI°PP3 gene was localized in a cluster alongside dehII°PP3 (17), and following cloning of the two genes together in E. coli DH5α (pYW2), the recombinant strain produced only one dehalogenase, corresponding to DehIIPP3 (Fig. 6, lane 1). There-

FIG. 5—Continued.
work for studying dehalogenase diversity and evolution and a basis for understanding the adaptation of bacteria to degrade xenobiotic haloaliphatic compounds.

Despite the fact that they are highly degenerate, the deh PCR primers described in this paper are useful molecular tools, facilitating rapid and efficient isolation and identification of deh genes in bacteria isolated from the natural environment. We have used the group I and group II deh primers to amplify deh genes directly from soil and enrichment culture samples (38), thus identifying dehalogenase genes in environmental samples without the need for bacterial cultivation. To maximize their specificity, the group II deh PCR primers were designed on the basis of conserved regions that did not overlap extensively with regions corresponding to the HAD superfamily motifs (32), thus limiting the number of potential universal primer-binding sites. Indeed, on the basis of theoretical considerations alone, these highly degenerate primers (Table 2) might easily have been discounted. However, the utility of the group II deh primers for isolation of a range of dehII-type genes is indicated by the observation from our wider screening program that 42 of 50 newly isolated αHA-utilizing bacteria tested positive. This result also suggested that group II deh genes may predominate over other αHA deh genes in the natural environment.

Group I now comprises 11 deh genes and is split into four subdivisions. Prior to this study, all of the deh genes described in the literature that were assigned to group I were produced by Pseudomonas species. Although subdivision C group I deh genes are also contained within host species from the genus Pseudomonas (sensu stricto), subdivisions A and B contain genes from various species of the α, β, and γ subclasses of Proteobacteria. Group I deh genes do not share any obvious feature with group II deh genes in terms of DNA or deduced amino acid sequences, suggesting that they are not evolutionarily related. They also seem to be functionally distinct in that all of the noncryptic group I deh genes tested encoded dehalogenases that dechlorinated 2-MCPA, whereas all group II dehalogenases tested lacked this activity. However, consideration of substrate specificities in isolation can be misleading in terms of investigating the evolutionary relationships between the dehalogenases. For example, of all the group I deh genes, hadD is most distantly related to dehD, even though these two genes uniquely encode dehalogenases that can dechlorinate 2-MCPA but not 1−2-MCPA (3, 7). Also, dehL appears to be similar to

**FIG. 6.** Native polyacrylamide gel electrophoresis gel showing dehalogenases in crude cell extracts from the following αHA-utilizing bacteria: lane 1, E. coli (pYW10); lane 2, P. putida PP3; lane 3, P. putida AJ1; lane 4, Pseudomonas sp. strain CBS3; lane 5, B. cepacia MBA4; lane 6, Burkholderia sp. strain K13; lane 7, Burkholderia sp. strain G02; lane 8, Burkholderia sp. strain 111; lane 9, Burkholderia sp. strain P11 (no αHA dehalogenase activity detected). Enzymes were visualized by activity staining with dichloroacetic acid (lanes 1, 2, and 6 to 9) or MCA (lanes 3 to 5) according to the method of Thomas et al. (59). Unless otherwise indicated, staining bands represent group II dehalogenases identified in separate zymography experiments by their activity with L-2MCPA but not 2-MCPA. Pseudomonas sp. strain CBS3 and Burkholderia sp. strain G02 each produced two separate group II dehalogenases. P. putida PP3 and AJ1 produced both group I (arrows) and group II dehalogenases.

**TABLE 3.** Comparison of group I deh genes and their expression in selected αHA-utilizing bacteria

<table>
<thead>
<tr>
<th>Organism (dehalogenase gene)</th>
<th>Group I deh PCR product (bp)*</th>
<th>RT-PCR product (bp)#</th>
<th>Presence of dehalogenase showing activity with 2-MCPAa</th>
</tr>
</thead>
<tbody>
<tr>
<td>P. putida A31 (hadD)</td>
<td>462</td>
<td>ND</td>
<td>+</td>
</tr>
<tr>
<td>Pseudomonas sp. strain 17α (dehl17α)</td>
<td>466</td>
<td>270</td>
<td>–</td>
</tr>
<tr>
<td>Pseudomonas sp. strain 18α (dehl18α)</td>
<td>460</td>
<td>400</td>
<td>+</td>
</tr>
<tr>
<td>Pseudomonas sp. strain K55 (dehlK55)</td>
<td>466</td>
<td>270</td>
<td>–</td>
</tr>
<tr>
<td>Bradyrhizobium sp. strain DA2 (dehlDA2)</td>
<td>466</td>
<td>330</td>
<td>+</td>
</tr>
<tr>
<td>Strain DA1, β subclass of Proteobacteria (dehlβ subclass)</td>
<td>466</td>
<td>400</td>
<td>–</td>
</tr>
<tr>
<td>P. putida PP3 (dehlPP3)</td>
<td>460</td>
<td>ND</td>
<td>+</td>
</tr>
<tr>
<td>E. coli (pYW10) (dehlPP3)</td>
<td>469</td>
<td>290</td>
<td>–</td>
</tr>
</tbody>
</table>

* Size determined by sequencing but not including the primer sequences.
# RT-PCR was performed with specific primers as described in Materials and Methods.
a As determined by gel zymography.
ND, not determined.
the group II deh genes, in that it encodes a dehalogenase with activity against L- but not D-2MCPA; however, it seems to be phylogenetically unrelated to any other deh gene.

The group II deh family seems to constitute a branch of the HAD superfamily (32), which was recently proposed to be linked with the evolution of P-type ATPases (1), and contains at least four major subdivisions defining closer deh gene relationships (Fig. 4). The conservation of amino acid motifs and active-site residues derived from PCR-amplified group II deh gene sequences (Fig. 5) supports the results of previous studies and helped to validate the approach used here.

All but one of the αHA-utilizing bacteria isolated in this study were shown to produce αHA dehalogenases. The exception, Burkholderia strain P11, utilized only 23DCPA (an αβHA), and no αHA dehalogenase activity was detected in this strain, suggesting that dehPP11 is a cryptic or silent gene. This would be consistent with the amino acid sequence derived from dehPP11, which gave a stop codon in the place corresponding to the conserved Y157, which was identified as essential for the activity of L-DEX by Soda and colleagues (18, 23). Consistent with the cryptic state of dehPP11, the open reading frame of this gene did not provide any obvious explanation as to why dehPP11 was cryptic. Interestingly, our analysis of unpublished sequence data from Honnens et al. (19) showed a partial gene sequence with high homology (95% identity over 474 nucleotides) to group I deh subdivision C, upstream from dehII (a group II deh) in Pseudomonas fluorescens ABVII. Furthermore, the similar region flanking the dehII gene in P. putida P3 and the sequence reported previously (19) upstream and downstream of dehII (98% identity over 1,663 nucleotides, including dehII and dehII homologues) suggested that adaptation of the host organisms, one isolated in the United Kingdom and the other isolated in Germany, to utilize αHAs may have involved horizontal transfer of linked group I and II deh genes. To utilize αHAs, the HA-dehalogenase activity was detected in the strain, suggesting that dehPP11 is a cryptic or silent gene. This would be consistent with the amino acid sequence derived from dehPP11, which gave a stop codon in the place corresponding to the conserved Y157, which was identified as essential for the activity of L-DEX by Soda and colleagues (18, 23).

All but one of the αHA-utilizing bacteria isolated in this study were shown to produce αHA dehalogenases. The exception, Burkholderia strain P11, utilized only 23DCPA (an αβHA), and no αHA dehalogenase activity was detected in this strain, suggesting that dehPP11 is a cryptic or silent gene. This would be consistent with the amino acid sequence derived from dehPP11, which gave a stop codon in the place corresponding to the conserved Y157, which was identified as essential for the activity of L-DEX by Soda and colleagues (18, 23). Consistent with the cryptic state of dehPP11, the open reading frame of this gene did not provide any obvious explanation as to why dehPP11 was cryptic. Interestingly, our analysis of unpublished sequence data from Honnens et al. (19) showed a partial gene sequence with high homology (95% identity over 474 nucleotides) to group I deh subdivision C, upstream from dehII (a group II deh) in Pseudomonas fluorescens ABVII. Furthermore, the similar region flanking the dehII gene in P. putida P3 and the sequence reported previously (19) upstream and downstream of dehII (98% identity over 1,663 nucleotides, including dehII and dehII homologues) suggested that adaptation of the host organisms, one isolated in the United Kingdom and the other isolated in Germany, to utilize αHAs may have involved horizontal transfer of linked group I and II deh genes.

Group II was found to contain two cryptic genes, dehIS1 and dehIS1, and an active gene that can be switched on and off by appropriate environmental selection, and although the mechanism for this has not been elucidated, it is likely that it involves interaction with a transposable element, DEH, containing dehIS1, a group I deh gene (59, 60).

Clearly, further investigation is needed to uncover how genes that are silenced in these various ways contribute to the potential of bacteria to degrade xenobiotic compounds.

There was no apparent correlation between the assignments of deh genes to subdivisions of group I and group II families and the 16S rRNA based phylogenetic assignments of their bacterial hosts. In itself this was not surprising, since horizontal transfer of deh genes carried on plasmids has previously been reported (27, 28). Also, Brokamp et al. (5) have described five plasmid-encoded dehalogenases, and at least three other dehalogenase genes, dehPP11 (59), the DehH2 gene (29), and dehII (4), are known to be carried on transposon-like mobile genetic elements. As far as we know, no dehalogenase-producing organism outside the phylum Proteobacteria has so far been reported, and this may reflect a limited phylogenetic range of organisms producing αHA dehalogenases. However, given their potential for horizontal transfer and the catalytic versatility of phyla such as the low- and the high-G+C gram-positive bacteria and the Cytophagales, each of which contains many xenobiotic-degrading species, confinement of αHA deh genes within Proteobacteria appears to be anomalous. Haloalkane dehalogenases are produced by species outside the Proteobacteria, so that the restricted range of species containing group I and group II deh genes may simply reflect limitations in currently used enrichment and isolation procedures related to the growth requirements of other bacterial groups.

Almost all of the αHA dehalogenases isolated in this laboratory and described in the literature are encoded by genes now assigned to either the group I or group II deh genes. Only two exceptions were identified: the DehH1 gene from a Moraxella sp. (27) and dehL from a Rhizobium sp. (7). On the basis of derived amino acid sequence alignments, Janssen et al. (22) proposed that the DehH1 gene is related to the haloalkane-dehalogenase genes, dhlA and dhaA, and the α-hexachlorocyclohexadiene dehalogenase gene, linB. This indicates the existence of a third family of αHA deh genes encoding dehalogenases with broader substrate specificities, including activities against haloalkanes. It should be noted that DehH1 is different from all of the group I and II dehalogenases in that it is a defluorinating enzyme, showing high activity with fluoroacetic acid but much lower activity with chloro- and bromo- analogues and virtually no activity with higher chemical homologues (26).

The dehL gene from a Rhizobium sp. (7) shows the same stereospecificity as group II deh genes (i.e., defluorination of L- but not D-2MCPA) but is not a member of this family. No significant match in either nucleotide or derived amino acid alignments between dehL and group II deh genes was found, and the derived DehL amino acid sequence lacks almost all of the conserved residues of the group II dehalogenases. Since this gene shows no significant match with any other sequence currently deposited in the GenBank and EMBL databases (on the basis of FASTA searches), it might tentatively be suggested that dehL uniquely represents a fourth deh family.

ACKNOWLEDGMENTS

We gratefully acknowledge funding for the research described in this paper from the European Commission (project no. ENV4-CT95-0086; support for K.E.H.) and the Wellcome Trust (research fellowship to I.R.M.).

Many thanks are due to Lee Parry for providing expert technical assistance and important preliminary data. Thanks also go to Andrew Gane and Gareth Lewis for sequencing.

The first two authors contributed equally to this paper.
REFERENCES


ERRATUM

Investigation of Two Evolutionarily Unrelated Halocarboxylic Acid Dehalogenase Gene Families

KATJA E. HILL, JULIAN R. MARCHESI, AND ANDREW J. WEIGHTMAN

Cardiff School of Biosciences, Cardiff University, Cardiff, CF1 3TL, Wales, United Kingdom

Volume 181, no. 8, p. 2535–2547, 1999. Page 2537, column 2: The following nucleotide sequence accession numbers were inadvertently omitted: dehI_{K55}, AJ133459; dehI_{PP3}, AJ133460.