

# The Role of Spontaneous Cap Domain Mutations in Haloalkane Dehalogenase Specificity and Evolution\*

(Received for publication, February 23, 1994, and in revised form, April 22, 1994)

Frens Pries, Arjan J. van den Wijngaard, Rolf Bos, Marjan Pentenga, and Dick B. Janssen‡

From the Department of Biochemistry, Groningen Biomolecular Sciences and Biotechnology, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands

The first step in the utilization of the xenobiotic chlorinated hydrocarbon 1,2-dichloroethane by *Xanthobacter autotrophicus* is catalyzed by haloalkane dehalogenase (DhlA). The enzyme hydrolyses 1-haloalkanes to the corresponding alcohols. This allows the organism to grow also on short-chain ( $C_2$ – $C_4$ ) 1-chloro-*n*-alkanes. We have expressed DhlA in a strain of *Pseudomonas* that grows on long-chain alcohols and have selected 12 independent mutants that utilize 1-chlorohexane. Six different mutant enzymes with improved  $K_m$  or  $V_{max}$  values with 1-chlorohexane were obtained. The sequences of the mutated *dhlA* genes showed that several mutants had the same 11-amino acid deletion, two mutants carried a different point mutation, and three mutants had different tandem repeats. All mutations occurred in a region encoding the N-terminal part of the cap domain of DhlA, and it is concluded that this part of the protein is involved in the evolution of activity toward xenobiotic substrates.

Some bacterial strains of the species *Xanthobacter autotrophicus* (1, 2) and *Ancylobacter aquaticus* (3) are capable of growing on the synthetic chlorinated hydrocarbon 1,2-dichloroethane. Degradation proceeds via 2-chloroethanol, chloroacetaldehyde, and chloroacetic acid to glycolate (1). Haloalkane dehalogenase (DhlA),<sup>1</sup> a 310-amino acid cytoplasmic protein, catalyzes the first hydrolytic reaction step. The gene has been cloned and sequenced (4), and identical dehalogenases were detected in different strains that grow on 1,2-dichloroethane (3).

The three-dimensional structure of DhlA was recently solved by x-ray crystallography (5, 6). DhlA is composed of a globular main domain with an  $\alpha/\beta$ -hydrolase fold structure and a separate cap domain (5–7). The active site cavity is located between the two domains. X-ray crystallographic studies indicated that the enzyme catalyzes cleavage of halogenated compounds by nucleophilic displacement of the halogen by Asp<sup>124</sup> (8). It was shown that during incubations of DhlA with substrate in  $H_2^{18}O$  incorporation of  $^{18}O$  occurred both in Asp<sup>124</sup> and product (9). This indicates that the covalent alkyl-enzyme intermediate, formed by nucleophilic displacement of the halogen by Asp<sup>124</sup>, is hydrolyzed by nucleophilic attack of a water molecule on the carbonyl carbon atom. His<sup>289</sup> probably activates this water molecule by subtracting a proton, with assistance of Asp<sup>260</sup> (8). Two tryptophans, Trp<sup>125</sup> and Trp<sup>175</sup>, are involved in substrate and halide binding (8, 10).

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 31-50-634209; Fax: 31-50-634165.

<sup>1</sup> The abbreviations used are: DhlA, haloalkane dehalogenase; bp, base pair.

The  $\alpha/\beta$ -hydrolase fold domain of DhlA is a structurally conserved fold shared by several hydrolytic proteins (7). The overall topology of DhlA consists of an eight-stranded  $\beta$ -sheet, with connecting  $\alpha$ -helices. The nucleophilic residues, either Ser, Cys, or Asp, are positioned on a sharp bend between the fifth  $\beta$ -strand and the following  $\alpha$ -helix. The positions of the other two active site residues, His<sup>289</sup> and Asp<sup>260</sup>, are conserved as well. The proteins differ by the presence, position, and sequences of cap domains, and therefore this part of the proteins was suggested to influence substrate specificity (7).

Haloalkane dehalogenase is active with several 1-chloro-*n*-alkanes and  $\alpha,\omega$ -dichloro-*n*-alkanes, but there is hardly any activity with 1-chlorohexane or chloroalkanes of longer chain length (1). The enzyme has some activity with long-chain bromoalkanes ( $C_5$ – $C_{10}$ ), suggesting that the active site can accommodate longer substrates and that differences in binding affinities or reaction rates are responsible for the lower activity with long-chain chloroalkanes.

Since 1,2-dichloroethane is not known to occur or to be produced naturally, it is likely that an enzyme capable of hydrolyzing this compound evolved to its present form after industrial 1,2-dichloroethane production and emission started in 1922 (11, 12). No sequences that are similar to DhlA and from which the present enzyme could be derived are known, however. We decided to investigate how DhlA can evolve to convert new substrates by selecting spontaneous mutants that degrade 1-chlorohexane, an approach termed experimental enzyme evolution (13–15). The sequences and activity of the mutant enzymes that were obtained indicate that short direct repeats in the N-terminal part of the cap domain play an important role in the evolution of haloalkane dehalogenase specificity.

## EXPERIMENTAL PROCEDURES

**Strains and Plasmids**—*Pseudomonas* GJ31 is a chlorobenzene utilizing organism described earlier (4, 16). *Escherichia coli* JM101 was used as host strain for isolation of single strand DNA and sequencing. Plasmid pPJ20 is a broad host range plasmid (4) derived from pLAFR1 (17). It contains the *dhlA* gene with its own promoter and constitutively produces DhlA in several Gram-negative bacteria (4).

**Selection of Mutants**—Selection of spontaneous mutant enzymes was achieved with the recombinant bacterial strain *Pseudomonas* GJ31(pPJ20). *Pseudomonas* GJ31 is capable of growing on long-chain alcohols. The plasmid was introduced by triparental mating (17) and allowed the organism to grow on 1-chlorobutane by expression of the *dhlA* gene up to levels of 25% of the total cellular protein (4). Mutants of *Pseudomonas* GJ31(pPJ20) that could utilize 1-chlorohexane as a growth substrate were selected (see Fig. 1) at 22 °C by cultivation in 1-liter serum flasks containing 400 ml of synthetic medium (1). The medium contained 2 mM 1-chlorobutane, which did support growth of the recombinant, and 2 mM 1-chlorohexane, which is hardly hydrolyzed by the wild type enzyme and did not support growth. After 4 weeks of incubation, some cultures showed higher turbidities than a control containing no 1-chlorohexane. After three serial transfers in medium containing 2 mM 1-chlorohexane as the sole carbon source, mutants were purified on nutrient broth agar plates containing tetracycline (12 µg/ml).

**Sequencing**—The *dhlA* gene of mutant V152–153 was determined by cloning *AluI*-*HindIII* and *Sau3A* DNA fragments in phage M13mp18

and M13mp19 (18), followed by dideoxy sequencing (19). The other mutant *dhlA* genes were amplified by the polymerase chain reaction (20), followed by dideoxy sequencing of the double-stranded polymerase chain reaction DNA. The primers used for sequencing are as follows (listed from 3' to 5' with the starting nucleotide number (4) given and c and n indicating coding and non-coding strand, respectively): 999 ATGATAATGCAATTCGC (c), 1303 ACCGCAACTTCCTGCTTGC (c), 1590 CTGACCGAAGCTGAGGCC (c), 1928 TTCTGTCTCGGCAAAGTG (n), 1628 GAAAGGCGCAGCATACGC (n), 1322 TGCAAGCAGGAGTTGCG (n).

**Purification of Haloalkane Dehalogenase**—Mutants of *Pseudomonas* GJ31(pPJ20) were grown at 30 °C in a mineral medium (1) containing 4 mM 1-chlorohexane as the sole carbon source. After 4 days, cells were harvested by centrifugation, washed in 10 mM Tris-H<sub>2</sub>SO<sub>4</sub> buffer, pH 7.5, containing 1 mM EDTA and 1 mM  $\beta$ -mercaptoethanol, and sonicated after resuspension in this buffer. Mutant haloalkane dehalogenases were purified as described before (1).

**Determination of Dehalogenase Activities**—The activities of the mutant enzymes were tested with a variety of substrates by spectrophotometric measurement of the rates of halide production (1) using purified enzymes (wild type and mutants P168S,  $\Delta$ 164–174, and  $\nabla$ 152–153) or crude extracts (mutants D170H,  $\nabla$ 172–174, and  $\nabla$ 145–154). The concentration of 1-chlorohexane in the assays was 3 mM, the other substrates were used at 5 mM. The dehalogenase content of cell-free extracts was determined by Coomassie Brilliant Blue staining of the protein bands of SDS-polyacrylamide gels, and subsequent peak integration of scans of the gels was taken with a desktop scanner. Accuracy was better than 10%.

For the determination of  $K_m$ ,  $V_{max}$ , and  $k_{cat}/K_m$ , alcohol production was measured in 4.5-ml incubations containing 0.01–4 mM substrate in 50 mM Tris-H<sub>2</sub>SO<sub>4</sub> buffer, pH 8.2, and 10 milliunits of 1,2-dibromoethane dehalogenase or 100 milliunits of 1-chlorohexane or 1,2-dichloroethane dehalogenase. Samples were incubated for 15 min at 30 °C, and the amount of alcohol produced was determined on a Chrompack 438S gas chromatograph with a CPWax 52 CB column, using an ECD detector for 2-bromoethanol and a FID detector for 2-chloroethanol and 1-hexanol. The carrier gas was nitrogen (60 kPa), and the temperature program was 3-min isothermal at 45 °C followed by an increase to 200 °C at 10 °C/min.  $K_m$  and  $V_{max}$  values were calculated from the rates of alcohol production by nonlinear regression analysis using the Michaelis-Menten equation and the Enzfitter program of Leatherbarrow (21).

## RESULTS

**Isolation of 1-Chlorohexane-utilizing Mutants**—The dehalogenase-expressing strain *Pseudomonas* GJ31(pPJ20) was found to be capable of growth on 1-chlorobutane, but not on 1-chlorohexane. The specific activity of purified haloalkane dehalogenase with 1-chlorohexane was only 0.17 unit/mg of protein, which apparently was not sufficient for growth, although the enzyme was expressed at a high level (4). Mutants of strain GJ31(pPJ20) that utilized 1-chlorohexane as sole carbon source were easily obtained in batch incubations that contained a mixture of 1-chlorobutane and 1-chlorohexane as sole carbon sources (Fig. 1). Twelve independent mutants were purified and analyzed for the presence of mutations in the *dhlA* gene by DNA sequencing (Table I) and determining dehalogenase activities in crude extracts (Table II).

**Sequences of Mutated Dehalogenases**—From sequencing the *dhlA* genes of the 1-chlorohexane utilizing mutants, six different mutants were identified. Only two mutants carried a single base substitution, leading to the mutations Asp<sup>170</sup> → His, and Pro<sup>168</sup> → Ser. The other mutants had much larger changes, being deletions or insertions. A large in frame deletion of 33 bp, causing an 11-amino acid deletion of Phe<sup>164</sup>–Ala<sup>174</sup>, was detected in six independently selected mutants. The duplications that were found were in frame direct repeats, varying in length from 6 to 30 bp. A mutant that carried a 6-bp duplication leading to a tandem repeat of Met<sup>152</sup>–Thr<sup>153</sup> was found twice. The other duplications were detected only once, and caused tandem repeats of Phe<sup>172</sup>–Ala<sup>174</sup> and Ile<sup>145</sup>–Asp<sup>154</sup>, respectively. The 9-bp duplication in mutant  $\nabla$ 172–174 encompassed the last 9 bp of the 33-bp deletion in mutant  $\Delta$ 164–174 (see also Fig. 2).

The deletions and duplications caused slight changes in elec-

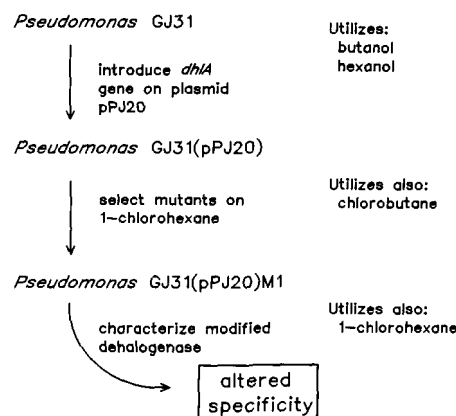


FIG. 1. Selection of mutants of haloalkane dehalogenase. Plasmid pPJ20 is a broad host range vector, expressing the *dhlA* gene product. It was introduced in *Pseudomonas* GJ31 by conjugation (4). Strains were cultivated as described under "Experimental Procedures."

TABLE I  
Mutants of haloalkane dehalogenase selected on 1-chlorohexane

| Mutant           | Mutation   | No. of mutants found |
|------------------|--|----------------------|
| D170H            | G → C substitution (Asp <sup>170</sup> → His)                      | 1                    |
| P168S            | C → T substitution (Pro <sup>168</sup> → Ser)                      | 1                    |
| $\Delta$ 164–174 | 33 bp deletion (Phe <sup>164</sup> –Ala <sup>174</sup> )           | 6                    |
| $\nabla$ 172–174 | 9 bp tandem duplication (Phe <sup>172</sup> –Ala <sup>174</sup> )  | 1                    |
| $\nabla$ 152–153 | 6 bp tandem duplication (Met <sup>152</sup> –Thr <sup>153</sup> )  | 2                    |
| $\nabla$ 145–154 | 30 bp tandem duplication (Ile <sup>145</sup> –Asp <sup>154</sup> ) | 1                    |

trophoretic mobility of some mutant dehalogenases on SDS-polyacrylamide gels (Fig. 3). The dehalogenase carrying the 11-amino acid deletion clearly had a higher mobility, whereas mutants  $\nabla$ 172–174 and  $\nabla$ 145–154 produced dehalogenases that migrated somewhat slower than the wild type.

No plasmids could be isolated from the D170H and  $\nabla$ 145–154 mutants suggesting that the plasmid had integrated into the chromosome of *Pseudomonas* GJ31. These mutants still produced haloalkane dehalogenase, but the D170H mutant expressed the enzyme at a lower level than the wild type and the other mutants (Fig. 3).

**Activities of in Vivo Mutants**—All mutants produced a dehalogenase with increased activity toward 1-chlorohexane and other C<sub>5</sub> and C<sub>6</sub> 1-chloro-*n*-alkanes (Table II). With all mutant enzymes, both the  $K_m$  and  $V_{max}$  values for 1-chlorohexane had improved, whereas these values were worse for 1,2-dichloroethane. The  $K_m$  value for 1,2-dibromoethane, which is the best substrate for the wild type enzyme, was also higher with all mutant enzymes (Table III). The individual mutants showed varying activities with several other substrates of the enzyme.

Mutant P168S had the highest  $V_{max}$  for 1-chlorohexane and 1,2-dibromoethane whereas the  $K_m$  for 1-chlorohexane was only slightly improved. The  $\Delta$ 164–174 deletion enzyme had lower activities than the wild type for all chlorinated compounds mentioned in Table II, except 1-chlorohexane, 1-chloropentane, and 1,6-dichlorohexane. The duplication mutant enzyme  $\nabla$ 172–174, which has changes in the same part of the protein as the  $\Delta$ 164–174 deletion, had similar activities, except that the specificity constant  $k_{cat}/K_m$  for 1-chlorohexane was only slightly improved and the activity for 1,2-dichloroethane was less reduced. The other two duplication mutant enzymes had the largest increase of the specificity constant for 1-chlorohexane, about 20 fold. Both  $V_{max}$  and  $K_m$  values were improved. They had a broad activity with long-chain (C<sub>4</sub>–C<sub>6</sub>) chloroalkanes. Summarizing, the general substrate range of the mutant enzymes thus had shifted from short chain substrates to 1-chloro-*n*-alkanes of longer chain length. This modification of specificity was not

TABLE II  
Activities of the mutant enzymes

Halide production in incubations containing purified dehalogenase (wild type (WT), mutants P168S,  $\Delta$ 164–174 and  $\nabla$ 152–153) or cell extract (the other mutants). Specific activities represent the amount of halide liberated (micromoles) per min per mg of dehalogenase protein added.

| Substrate          | Specific activities |       |       |                  |                  |                  |                  |
|--------------------|---------------------|-------|-------|------------------|------------------|------------------|------------------|
|                    | WT                  | D170H | P168S | $\Delta$ 164–174 | $\nabla$ 172–174 | $\nabla$ 152–153 | $\nabla$ 145–154 |
|                    | units/mg of protein |       |       |                  |                  |                  |                  |
| Dichloromethane    | 0.40                | 0.15  | <0.1  | <0.1             | <0.1             | 0.06             | <0.1             |
| 1,2-Dichloroethane | 6.60                | 1.22  | 4.69  | 0.11             | 1.03             | 0.15             | 0.43             |
| 1-Chlorobutane     | 2.05                | 0.71  | 1.21  | 0.34             | 1.07             | 1.37             | 0.92             |
| 1-Chloropentane    | 0.40                | 0.56  | 0.48  | 0.48             | 0.63             | 1.80             | 1.00             |
| 1-Chlorohexane     | 0.17                | 0.46  | 3.70  | 0.64             | 0.40             | 1.16             | 1.00             |
| 1,6-Dichlorohexane | 0.28                | 0.38  | 0.57  | 0.66             | 0.44             | 0.95             | 1.57             |
| 1-Chlorooctane     | 0.26                | 0.17  | 0.21  | <0.2             | <0.2             | 0.39             | 0.81             |
| 1,2-Dibromoethane  | 6.00                | 4.33  | 8.05  | 4.25             | 3.96             | 4.30             | 5.82             |
| 1-Bromopropane     | 1.43                | 0.51  | 1.25  | 1.12             | 0.87             | 0.56             | 1.22             |
| 1-Bromobutane      | 1.53                | 0.35  | 1.53  | 1.92             | 1.23             | 0.77             | 1.63             |
| 1-Bromopentane     | 1.94                | 0.64  | 2.13  | 1.52             | 1.50             | 0.72             | 1.75             |
| 1-Bromohexane      | 1.08                | 0.88  | 2.50  | 1.47             | 1.82             | 0.97             | 1.28             |
| 1-Bromooctane      | 1.21                | 0.37  | 1.37  | 0.84             | 1.23             | 0.72             | 1.22             |

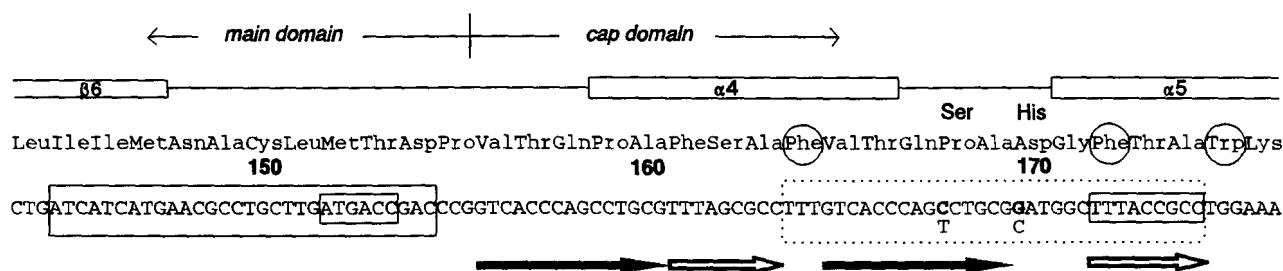


FIG. 2. **Repeats and mutations in the *dhIA* gene.** The part of the gene that encodes the N-terminal region of the cap domain is shown, together with the amino acid sequence. Residues forming the active site cavity are encircled. The boxes above the protein sequence indicate the position of strand 6 of the  $\beta$ -sheet and  $\alpha$ -helices 4 and 5. The large solid arrows mark the 15-bp direct repeat. One copy of the repeat with flanking nucleotides is deleted in the  $\Delta$ 164–174 mutant enzymes (dashed box), the formation of which is explained by misalignment of the 9-bp imperfect direct repeat at the ends of the deleted sequence (open arrows). Of this 9-bp repeat, the right-hand copy is again duplicated in mutant  $\nabla$ 172–174 (closed box). The other repeats detected in the mutants are located in the N-terminal part of the sequence encoding the cap domain (closed boxes). The bases printed in bold characters mark the position of the mutations in D170H and P168S, causing the substitutions indicated.

observed with bromoalkanes, as illustrated by the high activity of both the wild type and the mutant enzymes with 1,2-dibromoethane (Table III).

**Location of Mutations**—Attempts to crystallize purified mutant enzymes have failed so far, probably due to disturbance of the crystal contacts of the DhIA molecule at Gln<sup>167</sup> and Ala<sup>169</sup> by the mutations (6). Examination of the amino acids that are changed using the three-dimensional structure of the wild type enzyme (5) showed that none of the mutations directly affects active site residues, with the exception of the 11-amino acid deletion, in which the active site cavity forming residues Phe<sup>164</sup> and Phe<sup>172</sup> are lost. The other mutations thus have indirect effects.

The mutations were all located in a segment of the *dhIA* gene that encodes the N-terminal part of the cap domain (Fig. 2). In the mutant enzymes D170H, P168S,  $\Delta$ 164–174, and  $\nabla$ 172–174, the changes all affect the structurally important salt bridge Asp<sup>170</sup>–Lys<sup>261</sup> (Fig. 4). Although the Asp<sup>170</sup> O $\delta$ 1 is 15.6 Å away from the active site Asp<sup>124</sup> O $\delta$ 1, the loss of the Asp<sup>170</sup>–Lys<sup>261</sup> salt bridge appears to change the active site cavity in such a way that improved 1-chlorohexane conversion is achieved. Probably, the active site cavity residues Phe<sup>164</sup> and Phe<sup>172</sup> are more mobile or have a slightly different position in the mutants (Fig. 4). In the P168S enzyme, the proline which is located in a *cis* proline turn between helices 4 and 5, is substituted. This allows rotation around the C $\alpha$ –N bond of residue 168 and a change of the *cis* bond in a *trans* bond (22), which could lead to disturbance of the Asp<sup>170</sup>–Lys<sup>261</sup> salt bridge that is in its proximity. In the  $\nabla$ 172–174 mutant enzyme, Phe<sup>172</sup> is preceded by an Ala instead of Gly<sup>171</sup>, which has unusual  $\phi\psi$ -angles for Ala in the

structure of wild type DhIA and forms the N-cap of helix 5, where Ala is less favored (23, 24). This will affect the cavity geometry via Phe<sup>172</sup> or via the salt bridge by increasing the size of the *cis* proline turn. In the  $\Delta$ 164–174 mutant enzyme, Phe<sup>164</sup> and Phe<sup>172</sup>, which point to the cavity, are deleted, as is the salt bridge residue Asp<sup>170</sup>. In the other two mutant enzymes,  $\nabla$ 152–153 and  $\nabla$ 145–154, cavity size may be influenced, but this is less clear since the amino acids surrounding the duplication are not in contact with the active site.

## DISCUSSION

**Structure-Activity Relationships of Mutant Enzymes**—The mutants of haloalkane dehalogenase described here show that the cap domain of DhIA acts as an activity-modifying domain that can accommodate spontaneous mutations which can be selected *in vivo* and lead to alterations of substrate specificity. Since all mutants showed improved kinetics for 1-chlorohexane, the substrate binding site must have been altered in these enzymes such that improved binding and/or conversion is achieved. It is known that two tryptophans are involved in binding of the halide moiety of the substrate (10), but other residues are probably responsible for determining the substrate specificity. Candidates for residues that interact with the R-group of the substrate are Phe<sup>128</sup>, Phe<sup>164</sup>, Phe<sup>172</sup>, Phe<sup>222</sup>, Pro<sup>223</sup>, Val<sup>226</sup>, Leu<sup>262</sup>, and Leu<sup>263</sup> (5, 6). All mutations that influence substrate range were located in the part of the cap domain that is N-terminal of Trp<sup>175</sup>. This part is mainly stabilized by a surface located salt bridge between Asp<sup>170</sup> and Lys<sup>261</sup> of the main domain (Fig. 4) and hydrophobic interactions of 4 buried residues (Phe<sup>161</sup>, Phe<sup>164</sup>, Phe<sup>172</sup>, and Gly<sup>171</sup>). Since buried resi-





FIG. 3. SDS-polyacrylamide gel electrophoresis of crude extracts from *Pseudomonas* GJ31(pPJ20) and its 1-chlorohexane utilizing mutants. See Table II for properties of the mutants. The arrow indicates the position of the haloalkane dehalogenase protein. Note the differences in mobility of mutant proteins carrying the deletion and duplications. The different lanes indicate crude extracts from: 1, 2, and 3, three independent mutants  $\Delta 164$ –174; 4 and 5, two independent mutants  $\nabla 152$ –153; 6, mutant P168S; 7, mutant  $\nabla 172$ –174; 8, mutant  $\nabla 145$ –154; 9, mutant D170H; 10, purified wild type haloalkane dehalogenase.

TABLE III  
Kinetic constants of dehalogenase mutants

Kinetic constants were determined by measuring alcohol production. 1,2-dibromoethane is included since it is the best substrate known and the  $V_{\max}$  values are similar for the different mutants and the wild type enzyme.

| Strain            | 1,2-Dibromoethane      |             |   | 1-Chlorohexane         |             |  |
|-------------------|------------------------|-------------|---|------------------------|-------------|--|
|                   | $V_{\max}$<br>units/mg | $K_m$<br>mM | $k_{\text{cat}}/K_m$<br>$\text{s}^{-1}\text{mM}^{-1}$ | $V_{\max}$<br>units/mg | $K_m$<br>mM | $k_{\text{cat}}/K_m$<br>$\text{s}^{-1}\text{M}^{-1}$ |
| Wild type         | 6.0                    | 0.0072      | 483   | 0.25                   | 1.38        | 0.11   |
| D170H             | 4.4                    | 0.087       | 30  | 0.59                   | 0.88        | 0.40   |
| P168S             | 8.1                    | 0.031       | 152   | 5.51                   | 1.17        | 1.8  |
| $\Delta 164$ –174 | 4.5                    | 0.35        | 7.5   | 0.85                   | 1.0         | 0.5  |
| $\nabla 172$ –174 | 4.06                   | 0.092       | 26  | 0.46                   | 0.46        | 0.21   |
| $\nabla 152$ –153 | 4.5                    | 0.18        | 14.5  | 1.29                   | 0.34        | 2.2  |
| $\nabla 145$ –154 | 6.4                    | 0.51        | 7.3   | 1.06                   | 0.26        | 2.3  |

dues serve as structural anchors and strongly resist translocation (25) and because of the low number of hydrogen bonds and salt bridges, structural changes will be easily accommodated by the N-terminal part of the cap domain. If the mutations influence the Asp<sup>170</sup>–Lys<sup>261</sup> salt bridge, as indicated by examination of the position and role of the residues influenced by the mutations, then this part of the cap domain could have undergone significant changes that allow improved binding or conversion of long chain chloroalkanes.

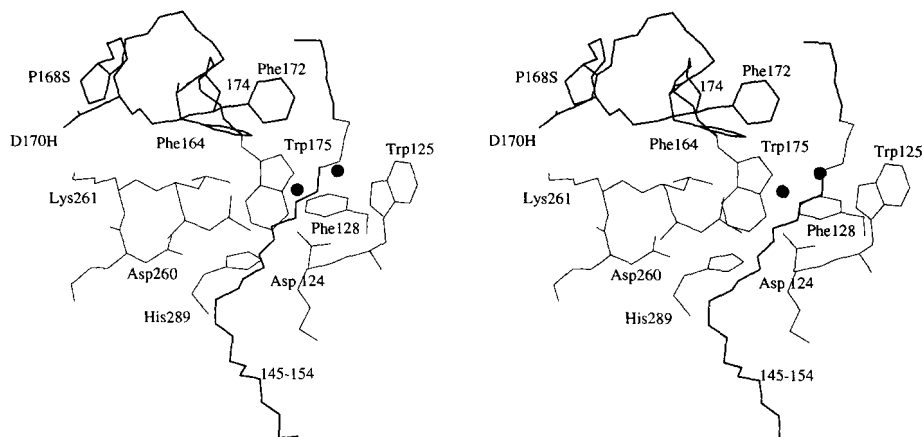
The mutations all influence the binding and/or reaction rate of various substrates (Tables II and III). The higher  $V_{\max}$  for 1-chlorohexane shows that the reaction of bound substrate is faster in the mutants. Since this effect is observed with chlorinated but not with brominated substrates, formation of the covalent alkyl-enzyme intermediate must be the rate-limiting step in the wild type that is accelerated in the mutants. This may be caused by a better average positioning of the C $\alpha$  of the bound substrate for nucleophilic attack by Asp<sup>124</sup> and stabilization of the reaction intermediate. The lower  $K_m$  could indicate that substrate binding is improved in the mutants, but it can also be explained by a higher rate of formation of the covalent

intermediate. How 1-chlorohexane is bound is unknown, but a larger or more flexible hydrophobic cavity that increases the binding energy of the larger substrate is likely. The wild type enzyme also must have considerable flexibility to allow substrate binding, since the active site cavity (5, 6) is too small for binding 1-bromohexane, which is a substrate for wild type haloalkane dehalogenase. Furthermore, the cavity is buried which requires some cap domain residues to move for substrate entrance. Large substrate-induced structural changes are observed in the mobile surface loops of lipase of *Rhizomucor miehei* (26), human pancreatic lipase (27), and lipoprotein lipase (28, 29), which are also  $\alpha/\beta$ -hydrolase fold excursions (7).

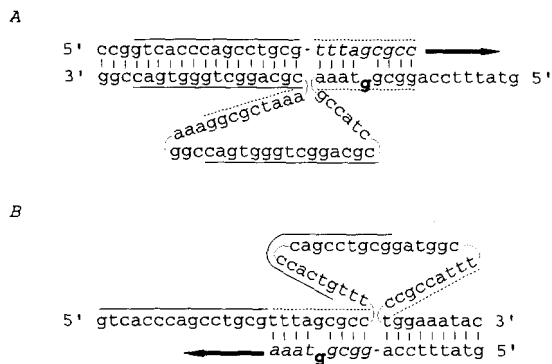
**Mechanism of in Vivo Mutagenesis**—Of the six mutations found, only two appeared to be substitutions. The high frequency with which the 11-amino acid deletion mutant was found (6 out of 12 mutants) can be explained by a strand slippage mechanism involving the 9-bp imperfect repeat that flanks in the wild type sequence the segment that is deleted in the  $\Delta 164$ –174 mutant. This can be followed by excision during mismatch repair or replication (Figs. 2 and 5). Such a process has been observed in  $\beta$ -galactosidase of *Lactobacillus bulgaricus* (15) and bacteriophage T4 (30) and was also proposed for the evolutionary divergence of benzoate dioxygenases (31). A role for replication was suggested (32, 33). Plasmid pPJ20 is derived from plasmid RK2, which has an unidirectional mode of replication (34). The orientation of the *dhla* gene in pPJ20 is such that the direction of transcription is the same as the direction of replication (4). All deletion mutants still contained the proximal copy of the 9-bp imperfect repeat encoding FSA, not FTA. This means that if the strand slippage occurred during leading strand synthesis (Fig. 5A), excision and mismatch repair occurred only in the template strand, but if strand slippage occurred during lagging strand synthesis (Fig. 5B), excision and mismatch repair took place in both strands, which seems less likely.

The occurrence of three different mutations that are direct tandem repeats of varying length, rather than insertions of random sequence, suggests a specific mechanism for their origin. Formation of direct repeats has been found to occur at sites flanked by short repeats that create DNA structures that are easily misaligned, recombined and extended (30, 35, 36). The tandem repeats in the mutated *dhla* genes are not flanked by repeated sequences, however, indicating that specific sequences are not required for their generation. The results suggest that an aspecific stuttering type of process can lead to the formation of short repeats, irrespective of existing repetitions.

**Evolution of Haloalkane Dehalogenase**—1,2-Dichloroethane is a synthetic compound. Although traces of it may be formed by haloperoxidases (37), it is unlikely that sufficient selective pressure to generate a dehalogenase for this chlorinated hydrocarbon existed until its industrial production and emission started in 1922. Dhla could have adapted to 1,2-dichloroethane by similar modifications of the cap domain as found here in the spontaneous mutants. Our results present experimental evidence that the cap domain, which is an excursion on the  $\alpha/\beta$ -hydrolase fold structure (7), is involved in the evolutionary modification of specificity, and that generation of repeats is an important mutational event during this process. The DNA coding for the N-terminal part of the cap domain of the wild type enzyme already contains two in frame direct repeats: a 15-bp perfect direct repeat encoding the sequence Val-Thr-Gln-Pro-Ala and a 9-bp repeat with one mismatch encoding Phe-Ser-Ala and Phe-Thr-Ala, respectively (Fig. 2). The two large repeats are in the same reading frame, but they are not tandem. It is highly unlikely that the repeats are required for the optimal catalytic performance of the enzyme. The direct repeats may well have been generated from an older dehalogenase during a



**FIG. 4. Part of the structure of haloalkane dehalogenase with the mutations indicated.** Stereoscopic depiction of the active site of wild type haloalkane dehalogenase. Asp<sup>124</sup> acts as the nucleophile (9), His<sup>289</sup> and Asp<sup>260</sup> are probably involved in the general-base catalyzed hydrolysis of the covalent intermediate, and Trp<sup>125</sup> and Trp<sup>175</sup> play a role in substrate binding and halide release (10). The two spheres represent water molecules in the active site cavity that occupy the substrate binding site (5). Residues affected by *in vivo* selected mutants are shown in *thick lines*. Side chains of the residues Asp<sup>124</sup>, Trp<sup>125</sup>, Phe<sup>128</sup>, Trp<sup>175</sup>, Asp<sup>260</sup>, Leu<sup>263</sup>, Leu<sup>263</sup>, and His<sup>289</sup> line the active site cavity. In mutants D170H, P168S, and V172-174 the salt bridge between Asp<sup>170</sup> and Lys<sup>261</sup> can be disturbed such that the active site cavity is larger or more flexible, allowing improved conversion of 1-chlorohexane. Probably, the active site cavity residues Phe<sup>164</sup> and Phe<sup>172</sup> have become more mobile or have slightly different positions in the mutants. These residues are both deleted in the  $\Delta$ 164-174 mutant. The V145-154 mutant is indicated by the *thick line* on the *right side* of the structure, and the V152-153 mutant by the intervening *thin part* in the sequence.



**FIG. 5. Proposed models for the formation of the  $\Delta$ 164-174 deletion during replication.** Replication of the plasmid and transcription of *dhlA* are in the same direction (4, 34). The two direct repeats which are present in wild type *dhlA* are indicated by lines. The *broken line* indicates the 9-bp imperfect repeat that can be involved in strand slippage. A loop of DNA is excised and a **g** (bold) is mismatch repaired to *c*. The result is a deletion of 33 bases. In A, leading strand synthesis during replication is indicated by the *arrow*. Mismatch repair occurs only in the template strand. In B, the *arrow* indicates lagging strand synthesis, and repair must occur in both strands. The slipped DNA is shown in *italics*.

process similar to the selection of the mutations that arose during cultivation on 1-chlorohexane, with a few additional fine tuning modifications leading to the present wild type DhIA. Thus, we hypothesize that the direct repeats in the DNA sequence encoding the N-terminal part of the cap domain are of recent evolutionary origin and were selected during the adaptation of an older dehalogenating enzyme to industrially produced 1,2-dichloroethane.

**Acknowledgments**—We thank Bauke Dijkstra and Koen Verschuieren for many stimulating discussions. We also thank Arjen Smal for experimental help.

#### REFERENCES

- Keuning, S., Janssen, D. B., and Witholt, B. (1985) *J. Bacteriol.* **163**, 635-639
- Janssen, D. B., Scheper, A., Dijkhuizen, L., and Witholt, B. (1985) *Appl. Environ. Microbiol.* **49**, 673-677
- Van den Wijngaard, A. J., Van der Kamp, K. W. H. J., Van der Ploeg, J., Pries, F., Kazemier, B., and Janssen, D. B. (1992) *Appl. Environ. Microbiol.* **58**, 976-983
- Janssen, D. B., Pries, F., Van der Ploeg, J., Kazemier, B., Terpstra, P., and Witholt, B. (1989) *J. Bacteriol.* **171**, 6791-6799
- Franken, S. M., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1991) *EMBO J.* **10**, 1297-1302
- Verschuieren, K. H. G., Franken, S. M., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1993) *J. Mol. Biol.* **232**, 856-872
- Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J., Sussman, J. L., Verschuieren, K. H. G., and Goldman, A. (1992) *Protein Eng.* **5**, 197-211
- Verschuieren, K. H. G., Seljée, F., Rozeboom, H. J., Kalk, K. H., and Dijkstra, B. W. (1993) *Nature* **363**, 693-698
- Pries, F., Kingma, J., Pentenga, M., Van Pouderoyen, G., Jeronimus-Stratingh, C. M., Bruins, A. P., and Janssen, D. B. (1994) *Biochemistry* **33**, 1242-1247
- Verschuieren, K. H. G., Kingma, J., Kalk, K. H., Janssen, D. B., and Dijkstra, B. W. (1993) *Biochemistry* **32**, 9031-9037
- Kastens, A. S. (1962) in *Chlorine, Its Manufacture, Properties and Uses* (Sconce, J. S., ed) pp. 310-333, Reinhold Pub. Corp., New York
- Hardie, D. W. F. (1964) in *Kirk-Othmer Encyclopedia of Chemical Technology* (Mark, H. F., McKetta, J. J., Othmer, D. F., and Standen, A., eds) 2nd Ed., Vol 5, pp. 148-170, John Wiley & Sons, New York
- Clarke, P. H. (1984) in *Microbial Degradation of Organic Compounds* (D. T. Gibson, ed), pp. 11-27, Marcel Dekker, New York
- Timmis, K. N., Rojo, F., and Ramos, J. L. (1990) *Adv. Appl. Biotechnol.* **4**, 61-82
- Mollet, B., and Delley, M. (1991) *Mol. Gen. Genet.* **227**, 17-21
- Oldenhuis, R., Kuijk, L., Lammers, A., and Janssen, D. B. (1988) *Appl. Microbiol. Biotechnol.* **30**, 211-217
- Friedman, A. M., Long, S. R., Brown, S. E., Buikema, W. J., and Ausubel, F. M. (1982) *Gene* **18**, 289-296
- Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) *Gene* **33**, 103-109
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463-5467
- Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. F., Higuchi, R. Horn, R. T., Mullis, K. B., and Erlich, H. A. (1988) *Science* **239**, 487-491
- Leatherbarrow, R. J. (1987) *Enzfitter: A Non-linear Regression Data Analysis Program for the IBM PC*, Elsevier Science Publishers B. V., Amsterdam, the Netherlands
- MacArthur, M. W., and Thornton, J. M. (1991) *J. Mol. Biol.* **218**, 397-412
- Serrano, L., Neira, J.-L., Sancho, J., and Fersht, A. R. (1992) *Nature* **356**, 453-455
- Serrano, L., Sancho, J., Hirshberg, M., and Fersht, A. R. (1992) *J. Mol. Biol.* **227**, 544-559
- Heinz, D. W., Baase, W. A., Dahlquist, F. W., and Matthews, B. W. (1993) *Nature* **361**, 561-564
- Derewenda, Z. S., Derewenda, U., and Dodson, G. G. (1992) *J. Mol. Biol.* **227**, 818-839
- Winkler, F. K. D'Arcy, A., and Hunziker, W. (1990) *Nature* **343**, 771-774
- Faustinella, F., Smith, L. C., and Chan, L. (1992) *Biochemistry* **31**, 7219-7223
- Dugi, K. A., Dichek, H. L., Talley, G. D., Brewer, H. B., Jr., and Santamaria-Fojo, S. (1992) *J. Biol. Chem.* **267**, 25086-25091
- Streisinger, G., and Owen, J. (1985) *Genetics* **109**, 633-659
- Harayama, S., Rekik, M., Bairoch, A., Neidle, E. L., and Ornston, L. N. (1991) *J. Bacteriol.* **173**, 7540-7548
- Gaillard, C., Strauss, F., and Bernardi, G. (1980) *Nature* **283**, 218-220
- Bron, S., Holsappel, S., Venema, G., and Peeters, B. P. H. (1991) *Mol. Gen. Genet.* **226**, 88-96
- Meyer, R. J., and Helinski, D. R. (1977) *Biochim. Biophys. Acta* **478**, 109-113
- Papanicolaou, C., and Ripley, L. S. (1991) *J. Mol. Biol.* **221**, 805-821
- Levinson, G., and Gutman, G. A. (1987) *Mol. Biol. Evol.* **4**, 203-221
- Gschwend, P. M., MacFarlane, J. K., and Newman, K. A. (1985) *Science* **227**, 1033-1036