Tetrachloroethene Dehalogenase from *Dehalospirillum multivorans*: Cloning, Sequencing of the Encoding Genes, and Expression of the pceA Gene in *Escherichia coli*

ANKE NEUMANN, GERT WOHLFARTH, AND GABRIELE DIEKERT*

*Institut für Mikrobiologie der Universität Stuttgart, Stuttgart, Germany*

Received 12 November 1997/Accepted 5 June 1998

The genes encoding tetrachloroethene reductive dehalogenase, a corrinoid-Fe/S protein, of *Dehalospirillum multivorans* were cloned and sequenced. The pceA gene is upstream of pceB and overlaps it by 4 bp. The presence of a ω70-like promoter sequence upstream of pceA and of a p-independent terminator downstream of pceB indicated that both genes are cotranscribed. This assumption is supported by reverse transcriptase PCR data. The pceA and pceB genes encode putative 501- and 74-amino-acid proteins, respectively, with calculated molecular masses of 55,887 and 8,354 Da, respectively. Four peptides obtained after trypsin treatment of tetrachloroethene (PCE) dehalogenase were found in the deduced amino acid sequence of pceA. The N-terminal amino acid sequence of the PCE dehalogenase isolated from *D. multivorans* was found 30 amino acids downstream of the N terminus of the deduced pceA product. The pceA gene contained a nucleotide stretch highly similar to binding motifs for two Fe₄S₄ clusters or for one Fe₄S₄ cluster and one Fe₃S₄ cluster. A consensus sequence for the binding of a corrinoid was not found in pceA. No significant similarities to genes in the databases were detected in sequence comparisons. The pceB gene contained two membrane-spanning helices as indicated by two hydrophobic stretches in the hydrophobic plot. Sequence comparisons of pceB revealed no sequence similarities to genes present in the databases. Only in the presence of pUBS 520 supplying the recombinant bacteria with high levels of the rare *Escherichia coli* tRNAPhe was pceA expressed, albeit nonfunctionally, in recombinant *E. coli* BL21 (DE3).

*Dehalospirillum multivorans* is a strictly anaerobic, gram-negative bacterium, which is able to grow with tetrachloroethene (PCE) as the terminal electron acceptor for the oxidation of different electron donors (14, 15, 19). The bacterium is able to be grown as previously described (14). PCE dehalogenase was isolated from the Allmandring 31, D-70550 Stuttgart, Germany. Phone: 49-(0)711-

**MATERIALS AND METHODS**

**Determination of amino acid sequences of PCE dehalogenase.** *D. multivorans* was grown as previously described (14). PCE dehalogenase was isolated from the organism as described elsewhere (16). The protein was treated with trypsin (22). The peptides obtained were separated by high-pressure liquid chromatography with a Grom-Sil 300 octyldecyl silane column, and the N-terminal amino acid sequence of four peptides and of PCE dehalogenase were determined by H. Weber at the Fraunhofer-Institut für Grenzflächen und Bioverfahrenstechnik (Stuttgart, Germany) or V. Nödinger at the Institut für Technische Biochemie (University of Stuttgart, Stuttgart, Germany).

**Cloning of pceA.** The isolation of DNA from *D. multivorans*, restriction, DNA ligation, and other standard techniques were performed as described elsewhere (2). Plasmid DNA for cloning and sequencing was prepared with the Flexi Prep kit (Pharmacia, Freiburg, Germany). Properties of plasmids used in this study are summarized in Table 1.

A homologous probe for pceA was generated by PCR with genomic DNA from *D. multivorans* as the template. The oligonucleotides (GGI GAG GTI AAG CCI TGG TT and GTC CCA IAC YTC IGX DAT RTT) were derived from the internal peptides GGVKPVWFLXYD and NITEVWDGK (Fig. 1). PCR mixtures (50 μl) for the amplification of genomic DNA contained 50 pmol of each primer, 0.1 μg of chromosomal template DNA, a 0.1 mM concentration of each deoxynucleotide triphosphate, Goldstar DNA polymerase reaction buffer, and 1 mM MgCl₂. The PCR program started with initial denaturing (3 min, 96°C). The addition of 0.5 U of Goldstar DNA polymerase (Eurogentec, Cologne, Germany) was followed by 29 cycles of polymerization (1 min, 45°C; 1.5 min, 72°C; 0.5 min, 95°C) and a final cycle with prolonged polymerization time (15 min, 72°C). A 1.2-kb fragment was amplified and cloned into a T-tailed vector (11) prepared from pBluescript II SK+ (Stratagene, Heidelberg, Germany). The resulting plasmid, named pW3, was partially sequenced. The identity of the fragment was confirmed by comparison of the deduced amino acid sequence with the peptide sequences of PCE dehalogenase. Genomic DNA was digested with several restriction endonucleases. The DNA fragments generated were separated by agarose gel electrophoresis, transferred to a nylon membrane by the capillary transfer method (21), and hybridized at 68°C with the 1.2-kb PCR product labeled with digoxigenin (DIG) by using the DIG DNA labeling and detection kit (nonradioactive) as indicated by the supplier (Boehringer, Mannheim, Germany). Genomic EcoRI fragments were isolated from agarose gels (Gene Clean II; Bio 101, La Jolla, Calif.), ligated into pBluescript II SK+, and transformed into *E. coli* DH5α cells (9). Positive clones were identified by Southern hybridization with the DIG-labeled 1.2-kb PCR product by using the DIG DNA detection kit (nonradioactive). One clone, named pY179, containing a 6-kb EcoRI fragment was used for further analyses. DNA sequencing and analyses. The nucleotide sequence of the 6-kb EcoRI fragment was determined by sequencing pY179 and subclones of pY179. For the sequencing reactions, an Applied Biosystems Prism kit ( Weiterstadt, Germany) was used, with subsequent electrophoresis and analyses in an Applied Biosystems 373A sequencer. Oligonucleotides (about 30 bases) were used for sequencing the remaining gaps. Both strands were independently and completely sequenced.
Computer-assisted DNA and protein sequence analyses, alignments, and hydrophilic plots were performed with the software package PCGene (IntelliGenetics Inc., Geneva, Switzerland). Database searches were performed with BLAST (1). Recognition of DNA and protein sequences was based on the method of Needleman and Wunsch (7) and the program package FASTA (8). The alignment of deduced protein sequences was performed with CLUSTAL W (9). Nucleotide sequence accession number. The sequences described in the present manuscript have been deposited in GenBank under accession no. AF022812.

RESULTS

Cloning and sequencing the PCE dehalogenase gene. PCE dehalogenase from D. multivorans was purified from the cytoplasmic fraction of pyruvate- and fumarate-grown cells as described elsewhere (16). The N terminus of the purified protein was sequenced. It started with the amino acid glycine, indicating a posttranslational processing of the protein. In addition, the PCE dehalogenase was digested with trypsin and the amino acid sequences of four of the resulting peptides were determined (Fig. 1). Two degenerated oligonucleotides were derived from the amino acid sequences of peptides 1 and 4 (Fig. 1). A 1.2-kb fragment was amplified by PCR with genomic DNA of D. multivorans as the template. The PCR product was cloned in pBluescript, and resulting plasmid pW3 was partially sequenced. All four peptides were found in the deduced amino acid sequence encoded by the 1.2-kb fragment. Given the apparent molecular mass of 58 kDa for the PCE dehalogenase

TABLE 1. Plasmids used in this study

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Relevant characteristics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pBluescript II SK+</td>
<td>Derived from pUC19; pPCEA; pPCEAB</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pET 11d</td>
<td>Expression vector, Ap r</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pLyS</td>
<td>Km r, T7 lysozyme; origin of replication</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pUBS 520</td>
<td>Derived from pACYC 177; Km r</td>
<td>R. Matsie</td>
</tr>
<tr>
<td>pW3</td>
<td>1.2-kb PCR product in EcoRI site of pBluescript II SK+</td>
<td>This study</td>
</tr>
<tr>
<td>pY179</td>
<td>6-kb EcoRI fragment of D. multivorans DNA in EcoRI site of pBluescript II SK+</td>
<td>This study</td>
</tr>
<tr>
<td>pCEA</td>
<td>PCR-generated pceA between NcoI and BamHI sites of pET 11d</td>
<td>This study</td>
</tr>
<tr>
<td>pCEA′</td>
<td>PCR-generated pceA′ (without the leader sequence) between NcoI and BamHI sites of pET 11d</td>
<td>This study</td>
</tr>
<tr>
<td>pCEAB</td>
<td>PCR-generated pceA′ between NcoI and BamHI sites of pET 11d</td>
<td>This study</td>
</tr>
</tbody>
</table>

* Ap r, ampicillin resistance gene; Km r, kanamycin resistance gene.

Cloning and sequencing the PCE dehalogenase gene. PCE dehalogenase from D. multivorans was purified from the cytoplasmic fraction of pyruvate- and fumarate-grown cells as described elsewhere (16). The N terminus of the purified protein was sequenced. It started with the amino acid glycine, indicating a posttranslational processing of the protein. In addition, the PCE dehalogenase was digested with trypsin and the amino acid sequences of four of the resulting peptides were determined (Fig. 1). Two degenerated oligonucleotides were derived from the amino acid sequences of peptides 1 and 4 (Fig. 1). A 1.2-kb fragment was amplified by PCR with genomic DNA of D. multivorans as the template. The PCR product was cloned in pBluescript, and resulting plasmid pW3 was partially sequenced. All four peptides were found in the deduced amino acid sequence encoded by the 1.2-kb fragment. Given the apparent molecular mass of 58 kDa for the PCE dehalogenase.

FIG. 1. DNA sequence of the operon containing the genes for the PCE dehalogenase from D. multivorans, deduced amino acid sequence, and N-terminal sequence. Putative E. coli σ70 promoter sequences (~35 and ~10 regions) are overlined. A possible terminator sequence is underlined with broken arrows, and ribosome binding sites are marked with asterisks. The leader peptide is in boldface. Solid lines below the deduced amino acid sequence are nucleotides encoded by pceA indicate sequences also determined by Edman degradation.
Taking into consideration the fact that PCE dehalogenase contains a protein with an apparent molecular mass of 61 kDa (Fig. 1), which possibly acts as a signal peptide, the sequences have been aligned relative to the consensus sequence RRXFXK (in boldface). Cleavage sites are marked by arrows.

**Sequence analysis of the PCE dehalogenase gene.** An open reading frame (ORF) coding for a protein (product of pceA) which harbors the N terminus and all four internal peptides of PCE dehalogenase was found in the 6-kb EcoRI fragment (Fig. 1). The sequence of the product of this ORF started 30 amino acids (corresponding to 90 bp of the ORF) upstream of the N terminus of PCE dehalogenase isolated from *D. multivorans*. In the N-terminal part of the deduced protein encoded by pceA, a putative signal sequence, RRXFXK, followed by a stretch of hydrophobic amino acids was detected (5) (Fig. 2). These findings support the assumption of a processing of the protein. During the first steps of purification of PCE dehalogenase, a protein with an apparent molecular mass of 61 kDa was copurified (Fig. 6, lane 6, band B). The N-terminal sequence of this protein was identical to the N-terminal amino acid sequence of the protein deduced from pceA, indicating that this protein was the unprocessed PCE dehalogenase.

The molecular masses of the deduced 501-amino-acid protein (nonprocessed) and of the truncated 471-amino-acid protein were calculated to be 55,887 Da and 52,674 Da, respectively. During the first steps of purification of PCE dehalogenase, a protein with an apparent molecular mass of 61 kDa was copurified (Fig. 6, lane 6, band B). The N-terminal sequence of this protein was identical to the N-terminal amino acid sequence of the protein deduced from pceA, indicating that this protein was the unprocessed PCE dehalogenase.

**B12-binding proteins performed with the software package PC/ Genet.**

The start codon of a further ORF (pceB) overlaps by four bases the C terminus-encoding region of pceA (Fig. 1). The 74-amino-acid protein encoded by this gene (225 bp) has a calculated molecular mass of 8,354 Da. Two hydrophobic regions were detected in the hydropathic plot of this protein, indicating the presence of two membrane-spanning helices (Fig. 3). Cysteine and histidine residues were not detected in the deduced amino acid sequence of pceB. In addition, sequence comparisons revealed no significant similarities to genes present in the databases. Each of the start codons of pceA and pceB is preceded by a putative ribosome binding site (Fig. 1). As shown in Fig. 1, there are two stretches resembling the −10 and −35 regions of an *E. coli* σ70 promoter, indicating a transcription start between positions 3160 and 3170. The primer extension method, with total RNA isolated from *D. multivorans* as the template, was used to determine that the transcription start site of the PCE dehalogenase was at approximately position 3180 (Fig. 4). In addition, RT-PCR with two different oligonucleotide pairs revealed no significant similarities to genes present in the databases.

**Genes upstream of the pceAB genes.** Upstream of the pceAB genes, one ORF (ORF1) with no significant sequence similarities to known B12-binding proteins performed with the software package PC/Genet was found in sequence comparisons. In addition, no significant similarities were revealed by amino acid sequence comparisons of the pceA gene product with vitamin B12- and coenzyme B12-binding proteins performed with the software package PC/Genet.
Moreover, two AGA tandems are present in (data not shown). Expected DNA fragments indicating the transcription of the gene of an IPTG-inducible T7/"promoter and transformed into E. coli BL21 (DE3)/pLysS. After induction of the bacteria with (positions 3302 to 3307 and 3776 to 3781 in Fig. 1). In E. coli, AGA codons are translated by the rare tRNA"AGA encoded by argU (18).

The plasmid pUBS 520 containing argU was transformed in E. coli BL21 (DE3) harboring pceA, pceA', or pceAB (see Table 1). After induction of the bacteria with 0.4 mM IPTG, proteins with molecular masses of 61 ± 1 kDa (pceA product) and 57 ± 1 kDa (pceA' product) were expressed in crude extracts of the respective recombinant bacteria as determined by SDS-PAGE (shown for pceA in Fig. 6). PCE dehalogenase activity could not be detected in crude extracts of recombinant E. coli. Growth of the recombinant bacteria under anaerobic conditions or the addition of vitamin B12 or of autoclaved crude extract of D. multivorans to the medium did not result in functional expression of PCE dehalogenase.

**DISCUSSION**

We show here for the first time cloning, sequencing, and expression of a PCE reductive dehalogenase from gram-negative, strictly anaerobic D. multivorans. Evidence for the presence of two ORFs, designated pceA and pceB, on the PCE dehalogenase operon is presented. From the finding that a putative E. coli σ70 promoter region precedes pceA and a p-independent terminator structure follows the stop codon of pceB, it was concluded that the two genes form one operon. This assumption was further supported by RT-PCR experiments indicating that both ORFs are cotranscribed. ORF pceA encodes the PCE dehalogenating protein. The finding that pceB encodes a highly hydrophobic protein with two transmembrane helices suggests that the gene product might be a membrane-anchoring subunit for the attachment of the pceA gene product to the cytoplasmic membrane. This is feasible, since the PCE dehalogenase is involved in a respiratory process (13, 17). No significant similarities of the pceB product to other proteins was found in sequence comparisons. Since the amino acids histidine and cysteine were lacking in the pceB gene product, a binding of heme to the protein as in cytchromes or of Fe/S clusters is not feasible, indicating that this "subunit" is probably not involved in the electron transport chain.

In the upstream region of pceA, three ORFs (ribABH) were identified as encoding putative enzymes of riboflavin biosynthesis; one ORF could not be identified. No ORF could be detected in the 0.9-kb downstream region of pceB. Putative ribosome binding sites were detected upstream of the start codon of each gene, indicating that the genes could be expressed in E. coli.

From the finding that the N terminus of PCE dehalogenase was found downstream of the N terminus of the deduced pceA protein, it is concluded that the protein was modified by truncation of the first 30 amino acids in D. multivorans. The modification signal was probably the peptide RRXFXK followed by a hydrophobic stretch, which was mainly reported for periplasmic, cofactor-binding proteins (5). This is surprising, since PCE dehalogenase was recovered exclusively in the cytoplasmic fraction of D. multivorans. The only other protein containing this leader sequence and facing the cytoplasmic side of the membrane is the dimethyl sulfoxide reductase of E. coli. This enzyme was reported to contain a hydrophobic subunit, which obviously hampers the catalytically active subunits from being excreted into the periplasm (24). Hence, it is feasible that the product of pceB serves a similar function for PCE dehalogenase. Usually, the cleavage sites for these and the Sec signal peptides are preceded by two small amino acids at positions −1 and −3 (23). In the pceA gene product, this site is preceded by phenylalanine at position −1 (Fig. 2).

The deduced amino acid sequence of corrinoid-iron/sulfur protein PCE dehalogenase exhibits no significant similarities to those of other proteins, including other cobalamin-containing enzymes. In addition, the cobalamin-binding site DXHXXG described for, e.g., the cobalamin-dependent methionine synthase (3) as well as for several adenosylcobalamin-containing...
mutases (4, 12), could not be detected. Since there are corrinoid-containing enzymes, which also lack this binding site, e.g., the corrinoid-iron-sulfur protein of Clostridium thermosacchari- 
cum (10), the corrinoid is probably noncovalently bound to the 
PCE dehalogenase. Evidence based on electron paramagnetic 
resonance (EPR) data is available for the PCE dehalogenase of 
Dehalobacter restrictus and shows that the cobalamins in 
this enzyme is a base-off corrinoid (20), suggesting that the 
cofactor is not coordinated with a histidine of the apoenzyme. 
Assuming a similar structure for the enzyme of D. multivorans, 
the lack of a cobalamin binding site is not surprising.

A consensus sequence very similar to that for Fe₄S₄ 
ferredoxins (two CXXXXCCXXCP stretches; 7) was detected in 
the deduced amino acid sequence encoded by pceA. However, 
the first cysteine of the second stretch is replaced by glycine. 
There are three possible explanations for this finding: (i) one 
iron atom of an Fe₄S₄ cluster is not bound to the sulfur of 
cysteine; (ii) one of the clusters contains only three Fe atoms; (iii) 
one iron atom of an Fe₄S₄ cluster is bound to the sulfur atom of 
cysteine, which is not located in the cluster. EPR experi-
ments performed with the PCE dehalogenase of D. restrictus 
indicated the presence of 8 Fe atoms (20). Until now, the 
absence of cysteine at the first position of the second stretch 
has never been reported for Fe₄S₄ proteins (examples are given in 
Fig. 7). Database research revealed a consensus of both 
stretches with only three other sequences, probably those of 
unknown proteins (Fig. 7). The elucidation of the structure of 
the iron-sulfur clusters will have to await EPR studies on the 
PCE dehalogenase of D. multivorans.

The high PCE concentrations detected in groundwater and 
soils of contaminated sites is due to extensive use of this com-
 pound as a solvent during the last 50 years. It is surprising that 
PCE dehalogenase has no significant homology to other pro-
 teins known so far, although it has to be assumed that the en-
zyme was developed from ancestors within this short period of 
time. The evolutionary origin of the enzyme remains to be un-
 raveled.

The codon usage of D. multivorans differs from that of E. coli especially with respect to the arginine tRNA codons. In the 
genesis of D. multivorans known so far, the frequency of the 
codon AGA is about the same as in closely related Wolinella succinogenes and about 10 times higher than in E. coli. Since 
AGA tandems, which often result in the formation of truncated 
gene products (18), are present in pceA, it was not sur-
prising that the gene was only expressed in E. coli in the 
presence of pUBS 520, which supplied the recombinant bac-
teria with high levels of the rare E. coli tRNA₅⁸⁸ for the AGA 
codon. Crude extracts of the recombinant E. coli did not ex-
hbit PCE-dehalogenating activity. The protein expressed in 
E. coli has an apparent molecular mass of 61 kDa, which was 
about 4 kDa larger than that of the PCE dehalogenase isolated from D. multivorans. Therefore, it is feasible that the 61-kDa 
protein product is an unprocessed PCE dehalogenase. This 
might be one of the reasons why the 61-kDa PCE dehalogenase 
was not enzymatically active. Although the apparent signal peptide 
RRXFKX should be processed by E. coli under anaerobic 
conditions (17), no modification of the pceA protein occurred.

A possible explanation may be that no cofactors were bound to 
the dehalogenase in E. coli, while only proteins containing 
the processed protein in E. coli with bound cofactors are 
processed in the signal peptide-carrying proteins (5). We also tried to express the pceA’ gene encoding the 
processed protein in E. coli. Although the protein was 
expressed, no PCE dehalogenase activity could be detected. 
Possibly the corrinoid and/or the Fe/S clusters have not been 
in- 
corporated in the protein. Since E. coli is not able to synthesize 
cobalamins, vitamin B₁₂ was added to the growth medium of the 
recombinant bacteria. The finding that no activity could be 
detected in the crude extracts of the supplemented bacteria 
may be due to the possibility that D. multivorans uses corrinoid 
derivatives other than vitamin B₁₂. To test this hypothesis, 
autoclaved crude extracts of D. multivorans were added to 
the growth medium supplying the recombinant E. coli with 
D. multivorans corrinoids. Supplementation of the medium 
with D. multivorans corrinoids did not result in functional 
expression. It is feasible that the cobalamins were not taken up by 
E. coli under the experimental conditions applied.

ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungs-
gemeinschaft, the BMBFT, and the Fonds der Chemischen Industrie.

We gratefully acknowledge helpful discussions with J. Altenbuchner 
and R. Mattes (Stuttgart, Germany). The plasmid pUBS 520 was 
kindly provided by R. Mattes (Stuttgart, Germany).

REFERENCES


Greene Publishing Associates, Inc., and John Wiley & Sons, Inc., New York, 
N.Y.

1989. Cloning and sequence analysis of the Escherichia coli metH gene 
encoding cobalamin-dependent methionine synthase and isolation of a trypt-
tic fragment containing the cobalamin-binding domain. J. Biol. Chem. 
264:13888–13895.

sion of the gene encoding the coenzyme B₁₂-dependent 2-methyleneglut- 


characterization of a 3-chlorobenzoxazole degrading pseudomonad. Arch. 
Microbiol. 99:61–70.


