

# Nucleotide Sequence and Expression of *clcD*, a Plasmid-Borne Dienelactone Hydrolase Gene from *Pseudomonas* sp. Strain B13

BETSY FRANTZ,<sup>1</sup> KA-LEUNG NGAI,<sup>2</sup> DEB K. CHATTERJEE,<sup>1</sup> L. NICHOLAS ORNSTON,<sup>2</sup>  
AND A. M. CHAKRABARTY<sup>1\*</sup>

*Department of Microbiology and Immunology, University of Illinois College of Medicine at Chicago, Chicago, Illinois 60612,<sup>1</sup> and Department of Biology, Yale University, New Haven, Connecticut 06511<sup>2</sup>*

Received 28 July 1986/Accepted 31 October 1986

The *clcD* structural gene encodes dienelactone hydrolase (EC 3.1.1.45), an enzyme that catalyzes the conversion of dienelactones to maleylacetate. The gene is part of the *clc* gene cluster involved in the utilization of chlorocatechol and is carried on a 4.3-kilobase-pair *Bgl*III fragment subcloned from the *Pseudomonas* degradative plasmid pAC27. A 1.9-kilobase-pair *Pst*I-*Eco*RI segment subcloned from the *Bgl*III fragment was shown to carry the *clcD* gene, which was expressed inducibly under the *tac* promoter at levels similar to those found in 3-chlorobenzoate-grown *Pseudomonas* cells carrying the plasmid pAC27. In this study, we present the complete nucleotide sequence of the *clcD* gene and the amino acid sequence of dienelactone hydrolase deduced from the DNA sequence. The NH<sub>2</sub>-terminal amino acid sequence encoded by the *clcD* gene from plasmid pAC27 corresponds to a 33-residue sequence established for dienelactone hydrolase encoded by the *Pseudomonas* sp. strain B13 plasmid pWR1. A possible relationship between the *clcD* gene and *pcaD*, a *Pseudomonas putida* chromosomal gene encoding enol-lactone hydrolase (EC 3.1.1.24) is suggested by the fact that the gene products contain an apparently conserved pentapeptide neighboring a cysteinyl side chain that presumably lies at or near the active sites; the cysteinyl residue occupies position 60 in the predicted amino acid sequence of dienelactone hydrolase.

Dienelactone hydrolase (EC 3.1.1.45) catalyzes the hydrolysis of dienelactone to maleylacetate (26), a step in the metabolic utilization of halogenated aromatic compounds via 3-chlorocatechol (Fig. 1). The enzyme has been purified to homogeneity from *Pseudomonas* sp. strain B13 and shown to be a monomer with a molecular weight of approximately 28,000 (20). The hydrolase acts on both *cis* and *trans* isomers of dienelactone but exhibits no activity when incubated with the substrate analogs muconolactone or  $\beta$ -keto adipate enol-lactone (20, 26).

The structural gene *clcD* encodes dienelactone hydrolase and is carried with other genes associated with catabolism of chlorocatechols (*clc* genes) on the 110-kilobase-pair (kbp) *Pseudomonas* plasmid pAC25 (10). Genes for these functions also are carried on the plasmid pWR1 in an independently isolated *Pseudomonas* sp. strain B13 (22). Comparison of pWR1 with pAC25 revealed nearly identical restriction patterns; with the exception of a 6-kbp deletion in pWR1, the two plasmids appear to be indistinguishable (8).

The plasmid pAC27, derived from pAC25, yielded a 4.3-kbp *Bgl*III fragment containing the *clc* gene cluster that is essential for the utilization of chlorocatechol (Fig. 1; 15). This cluster includes *clcA*, the gene for catechol oxygenase II (EC 1.13.11.1 or EC 1.13.11. – [chlorocatechol-1,2-dioxygenase or chlorocatechol oxygenase, respectively]); *clcB*, the gene for muconate cycloisomerase II (EC 5.5.1.7 [chloromuconate cycloisomerase, muconate lactonizing enzyme II]); and *clcD*, the gene for dienelactone hydrolase (Fig. 2).

In this report, we present the complete nucleotide sequence of the 1.9-kbp *Pst*I-*Eco*RI segment of pDC100 which encodes a dienelactone hydrolase that, consonant with the structural similarity of plasmids pWR1 and pAC25, resembles the *Pseudomonas* sp. strain B13 dienelactone hydrolase in every discernible respect. When placed under the *tac*

promoter both in *Pseudomonas putida* and *Escherichia coli*, the *clcD* gene is expressed inducibly at levels comparable to those found in 3-chlorobenzoate-grown cultures of *P. putida* carrying the plasmid pAC27.

## MATERIALS AND METHODS

**Bacteria, plasmids, and phage.** The bacteria, plasmids, and phages used in this study are listed in Table 1. Cloning, transformation, and plasmid purification procedures have been described (13, 17). Screening of strains for recombinant plasmids was carried out by the modified Birnboim and Doly alkaline lysis method (17).

**Media, chemicals, and enzymes.** *E. coli* strains were grown in L broth (GIBCO Diagnostics). Solid media contained 1.5% agar (GIBCO) except for plaque lawns, which contained 0.65% agar. When required, ampicillin (Sigma Chemical Co.) was added at 100  $\mu$ g/ml. Isopropyl- $\beta$ -D-thiogalactosylpyranoside (IPTG; Sigma) was added to a final concentration of 1 mM. Protein assay kits and the bovine serum albumin standard were from Bio-Rad Laboratories. Restriction enzymes were purchased from New England BioLabs or Bethesda Research Laboratories. Terminal deoxynucleotidyl transferase, T4 polymerase, and subcloning primers for deletion cloning were from International Biotechnologies, Inc. All nucleotides, the Klenow fragment of DNA polymerase I, 5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside (X-Gal), and 15- and 17-mer M13 sequencing primers were purchased from Bethesda Research Laboratories. 2'-Deoxyinosine 5'-triphosphate was purchased from Pharmacia, Inc. Radiochemicals were purchased from Amersham Corp. *cis*- and *trans*-diene-lactones were kindly supplied by H.-J. Knackmuss.

**HPLC.** High-performance liquid chromatography (HPLC) was conducted with a Waters system (Waters Associates, Inc.) with the following accessories: Toya Soda TSK DEAE-5PW anion-exchange columns (21.5 mm by 15 cm) from

\* Corresponding author.

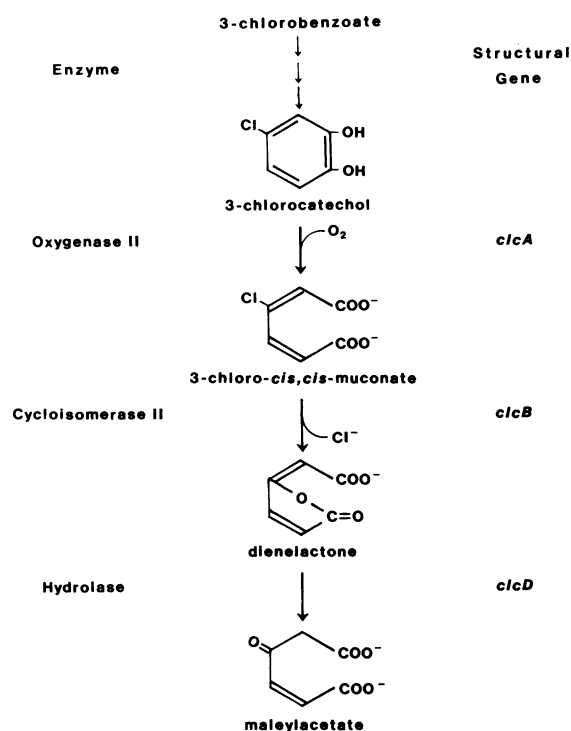


FIG. 1. Metabolic steps in the conversion of chlorocatechol to maleylacetate. 3-Chlorobenzoate is metabolized via 3-chlorocatechol and is used as a growth substrate to induce the *clc* genes in *Pseudomonas* cells with the plasmid pAC27 or pWR1. Specific designations for *clc* genes and for the enzymes that they encode are shown with the corresponding metabolic steps.

Bio-Rad and Zorbax Bio-Series GF-250 (9.4 mm by 24 cm; molecular separation range of 4,000 to 400,000) and GF-450 (9.4 mm by 25 cm; molecular separation range of up to 900,000) size-exclusion columns from E. I. du Pont de Nemours & Co. HPLC-grade water was from a Milli-Q water system (Millipore Corp.). The Centricon-10 (10,000-molecular-weight cutoff) ultrafilter was from Amicon Corp., and the 0.45- $\mu$ m (pore size) cellulose acetate membrane unit was provided by Nalgene Labware Div., Nalge/Sybron Corp.

**Maxicell labeling of plasmid-encoded protein.** Plasmid-encoded proteins were labeled with L-[<sup>35</sup>S]methionine in *E. coli* maxicells as described by Sanchar et al. (24). The proteins were separated by electrophoresis on sodium dodecyl sulfate-polyacrylamide gels and visualized by autoradiography.

**Growth of cells, extraction, enzyme purification, and pro-**

**tein characterization.** *P. putida* cultures were grown at 30°C in minimal medium containing 5 mM 3-chlorobenzoate. *E. coli* cultures were grown at 37°C in L broth containing 100  $\mu$ g of ampicillin per ml and, where indicated, 1 mM of the inducer IPTG. Cells were harvested by centrifugation and stored at -20°C.

The buffer used for extraction of cells for comparison of enzyme levels contained 50 mM sodium ethylenediamine, 1 mM MnCl<sub>2</sub>, and 1 mM 2-mercaptoethanol (pH 7.3). Extracts were maintained at or below 4°C. Frozen cells from 1-liter cultures were suspended in 1 ml of buffer and extracted by sonication. The suspension, after centrifugation at 15,000  $\times$  *g* for 15 min, was used to determine enzyme levels and protein concentration in induced and uninduced cells. Catechol oxygenase activity (14), dienelactone hydrolase activity (26), and protein concentrations (5, 28) were determined by published procedures.

EDA buffer (10 mM sodium ethylenediamine [pH 7.3] with 1 mM MnCl<sub>2</sub> and 1 mM 2-mercaptoethanol) was used throughout the HPLC purification of dienelactone hydrolase. HPLC was performed at room temperature; with this exception, enzyme preparations were maintained at or below 4°C. Frozen cells were suspended in EDA buffer (2.5 g of cells per 10 ml of buffer) and extracted by sonication. The extract was clarified by centrifugation, treated with 0.25 M NaCl to dissociate enzyme aggregates, and dialyzed against several changes of 1 liter of EDA buffer. The dialysate was centrifuged at 100,000  $\times$  *g* for 1 h, and the supernatant liquid was filtered through a 0.45- $\mu$ m (pore size) cellulose acetate membrane unit. About 10 ml of the filtrate containing 200 to 250 mg of protein was injected into a TSK DEAE-5PW column (19) (Bio-Rad technical bulletin no. 1150) which had been equilibrated with EDA buffer. The column was washed for 10 min, and enzyme was eluted with the same buffer containing 0 to 0.3 M sodium chloride in a linear gradient over 120 min at a flow rate of 5 ml/min. Fractions containing dienelactone hydrolase activity were pooled and concentrated to 100  $\mu$ l with a Centricon-10 unit. The concentrate was eluted through GF-250 and GF-450 columns in series (30) (Dupont Zorbax user guide). Elution was conducted at a flow rate of 0.5 ml/min with EDA buffer containing 0.25 M NaCl. Fractions of 0.5 ml were monitored for protein by measurement of A<sub>280</sub>. Hydrolase activity was determined in fractions containing protein. Fractions with hydrolase activity were pooled and concentrated to 100  $\mu$ l with a Centricon-10 unit. The concentrate was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (31) to check for purity.

**Determination of amino acid compositions.** HPLC-purified protein was desalted in an Amicon ultrafilter, lyophilized, and hydrolyzed in 6 M HCl containing 1% (vol/vol) liquified

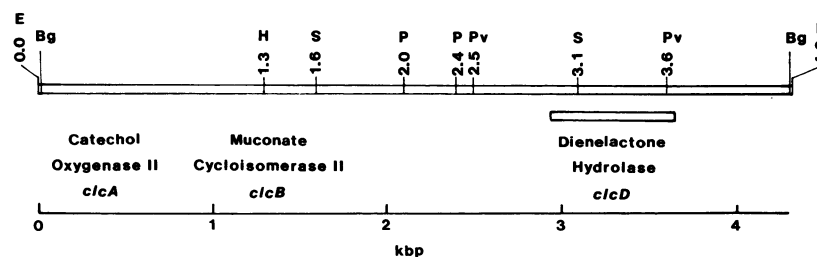


FIG. 2. Restriction map and organization of *clc* degradative genes from the 4.3-kbp *Bgl*II fragment cloned into pDC100. Genes have been assigned on the map by correlation of nucleotide sequence with the amino acid composition and NH<sub>2</sub>-terminal amino acid sequence of the purified enzymes. Restriction sites: Bg, *Bgl*II; E, *Eco*RI; H, *Hind*III; P, *Pst*I; Pv, *Pvu*II; S, *Sal*I.

phenol in vacuo at 115°C for 20 h. Amino acid compositions were then determined with a Waters PICO-TAG amino acid analysis system. Amino acids from the hydrolysate were derivatized with phenylisothiocyanate for separation and quantitation by reversed-phase HPLC against external phenylthiocarbonyl amino acid standards (16; Waters PICO-TAG amino acid analysis system operator manual).

**DNA sequence analysis.** DNA sequences were determined by the dideoxy chain termination procedure described by Sanger et al. (25). The 1.9-kbp *PstI-EcoRI* segment from the 4.3-kbp *BglII-EcoRI* fragment of pDC100 was cloned into both M13mp18 and M13mp19. Deletion subcloning of the fragment in both phages by the method of Dale et al. (11) gave the fully overlapped sequence of both strands. Ligation mixtures of M13 recombinants were transformed into *E. coli* strain JM103. We have found better yields, however, of both mature and replicative form M13 DNAs for both analytical determination and template isolation if plaques have been inoculated for 6 to 7 h into overnight cultures of *E. coli* C600F<sup>+</sup> cells rather than into JM103 cells. Sequencing reactions with [<sup>32</sup>P]dCTP (800 Ci/mmol; Amersham) were carried out by the methods described in the Bethesda Research Laboratories M13 Cloning/Dideoxy Sequencing Instruction Manual with the following modifications: molar ratios of dGTP and dCTP were increased twofold to allow for the high G+C content of *Pseudomonas* DNA. Polymerization reactions were performed at 37°C. When required, the following dITP reaction mixtures were used to alleviate GC compression: (i) G, 0.02 mM dITP-0.005 mM dGTP-0.075 mM dATP-0.075 mM dTTP-0.005 mM dCTP-0.05 mM ddGTP; (ii) A, 0.1 mM dITP-0.035 mM dGTP-0.005 mM dATP-0.075 mM dTTP-0.005 mM dCTP-0.5 mM ddATP; (iii) T, 0.1 mM dITP-0.035 mM dGTP-0.075 mM dATP-0.005 mM dTTP-0.005 mM dCTP-0.6 mM ddTTP; (iv) C, 0.1 mM dITP-0.035 mM dGTP-0.075 mM dATP-0.075 mM dTTP-0.005 mM dCTP-0.2 mM ddCTP.

Urea gels containing 7% polyacrylamide (1:20 *N,N'*-methylene-bisacrylamide linkage) were subjected to electrophoresis with 100 mM Tris-borate-2 mM EDTA buffer (pH

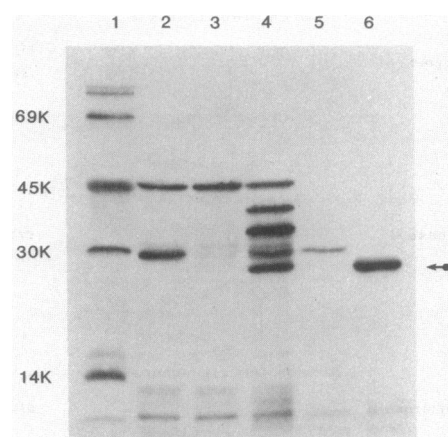


FIG. 3. Autoradiogram of L-[<sup>35</sup>S]methionine-labeled proteins after UV irradiation of *E. coli* maxicells carrying various plasmids. Lanes: 1, <sup>14</sup>C-labeled molecular weight protein standards; 2, pMMB22 induced with IPTG; 3, pDC100 uninduced; 4, pDC100 induced with IPTG; 5, pKK223-3 induced with IPTG; 6, pKK223-3 carrying the 1.9-kbp *PstI-EcoRI* fragment induced with IPTG. The three major polypeptides produced by pDC100 in response to IPTG and their apparent molecular weights are dienelactone hydrolase (28,000) (arrow), catechol oxygenase II (33,000), and muconate cycloisomerase II (40,000). The functions of minor products with apparent molecular weights of 29,000 and 32,000 have not been defined.

8.3) at a constant power of 55 W. Gels were fixed in 10% acetic acid, dried on Whatman no. 1 paper, and exposed to XAR-5 or SB-5 film (Eastman Kodak Co.) overnight at room temperature.

## RESULTS

**Cloning and expression of genes for catabolism of chlorocatechol.** The recombinant plasmid pDC25 carries the 4.3-kbp *BglII* fragment (fragment E) from the 3-chlorobenzoate degradative plasmid pAC27 (9, 15). This fragment was cloned into the broad-host-range expression vector pMMB22 (3) after addition of *EcoRI* linkers to produce the recombinant plasmid pDC100. Expression of pDC100 in *E. coli* CSR603 maxicells produced three major polypeptides that exhibited apparent *M<sub>r</sub>*s of 28,000, 33,000, and 40,000 (Fig. 3, lane 4). In a separate investigation, the larger polypeptides were shown to be catechol oxygenase II and muconate cycloisomerase II, respectively. The molecular weight of dienelactone hydrolase, encoded by plasmid pWR1 in *Pseudomonas* sp. strain B13, is known to be approximately 28,000 (20), which suggested that the 28,000-*M<sub>r</sub>* polypeptide might be dienelactone hydrolase, the product of the *clcD* structural gene. Evidence presented below demonstrates that this was indeed the case. The 1.9-kbp *PstI-EcoRI* fragment from pDC100 was subcloned in pKK223-3 (Pharmacia), and expression of this plasmid in maxicells produced only the 28,000-*M<sub>r</sub>* polypeptide (Fig. 3, lane 6). This evidence places the coding region for *clcD* toward the 3' end of the 4.3-kbp *BglII* fragment.

**Expression of *clc* genes in *E. coli* and *P. putida*.** When placed under control of the *tac* promoter of recombinant plasmid pDC100 in *E. coli* AC80, the *clcA* and *clcD* genes were expressed in response to IPTG. Respective uninduced levels of 0.0015 and 0.05 U/mg of protein increased about 40-fold to respective levels of 0.06 and 2.1 U/mg of protein when cells were grown in the presence of inducer IPTG. This

TABLE 1. Bacterial strains, plasmids, and phages

| Strain, plasmid, or phage | Genotype/phenotype   | Source or reference |
|---------------------------|--|---------------------|
| <b>Strains</b>            |  |                     |
| <i>P. putida</i> PRS2015  | <i>catB</i>  | 32                  |
| <i>E. coli</i> C600       | F <sup>+</sup> <i>leuB6 fhuA21 thi-1 thr-1 lacY1 supE44</i>                  | 2                   |
| JM103                     | Δ( <i>lac-pro</i> ) <i>thi supE rpsL endA sbcB15 hsdR4/F'</i>                | 18                  |
|                           | <i>traD36 proAB lacI<sup>q</sup>ZΔM15</i>                                    |                     |
| CSR603                    | <i>recA1 uvrA6 phr-1</i>   | 24                  |
| AC80                      | <i>thr leu met hsdR</i>  | 7                   |
| <b>Plasmids</b>           |  |                     |
| pKK223-3                  | Ap <sup>r</sup> <i>p<sub>tac</sub></i>                                       | Pharmacia           |
| pMMB22                    | IncQ Ap <sup>r</sup> Sm <sup>r</sup> <i>p<sub>tac</sub> lacI<sup>q</sup></i> | 3                   |
| pAC27                     | 3Cba <sup>+</sup>  | 15                  |
| pDC25                     | 3Cba <sup>+</sup> , Tc <sup>r</sup>  | 15                  |
| pDC100                    | 3Cba <sup>+</sup> Ap <sup>r</sup> <i>p<sub>tac</sub> lacI<sup>q</sup></i>    | This study          |
| <b>Phages</b>             |  |                     |
| M13mp18                   | <i>p<sub>tac</sub> lacZ</i>  | 21                  |
| M13mp19                   | <i>p<sub>tac</sub> lacZ</i>  | 21                  |

| Pst I  |   | Pvu II |  |      |
|--|---|--------|--|------|
| CTGCAGGCATGGGACGTGATCCATCTGGGGCCGCTCTGCTGGCCGCTCAGGCCGGGGTTGACCTGCTCCAAGTGCCTACAAAGGGAGCACGCCAGCTGCCCGACCTGATCGCGG                 |   |        |  | 120  |
| GGGACTCTCCATGATGGTGGATTGATCACCCGACAGCAGTGGTTTATCAAATCTGGACGGTGGCTGGCTGGTGTCCACCAGTTGCAACCCCGCCCTCATTGCCTGGGATTCCTC                 |   |        |  | 240  |
| CCCTGGCCGAAGCAGCCGACCAACCGAATCTGGAGTCTGACATGGTTGGCTTATTTGGCCGTCACGTACCTCCGCGGACATCGTCAAGGTGCTCAACACAGCAATGAATGAAGCAC               |   |        |  | 360  |
| TGAAGACCCCGGGATACAAAAGCGCTGGCGGATATCCGGCCCTCAGCCGACGGGGGAAATCGCAAGCTCTGGCCGCTTATGGGACAACGAGATCGACCGATGGGGCCAGTTGATCG               |   |        |  | 480  |
| TCCACCATCGACTCAACCAATGAATCGGAGAGATGGGAAG   | ATG TTG ACT GAA GGG ATA TCG ATT CAA TCG TAT GAC GGG CAT ACA TTC GGC |        |  | 580  |
|  | Met Leu Thr Glu Gly Ile Ser Ile Gln Ser Tyr Asp Gly His Thr Phe Gly |        |  |      |
|  | 1   |        |  |      |
| GGG CTC GTG GGC TCG CCG GCC AAA GGG CCC GCT CCC GTG ATT GTG ATC GCT CAA GAA ATA TTT GGT GTG AAC GCG TTC ATG CGA GAA ACG            |   |        |  | 670  |
| Ala Leu Val Gly Ser Pro Ala Lys Ala Pro Ala Pro Val Ile Val Ile Ala Gln Glu Ile Phe Gly Val Asn Ala Phe Met Arg Glu Thr            |   |        |  |      |
|  | 20  |        |  |      |
|  | Sal I   |        |  |      |
|  | 30  |        |  |      |
| GTG TCA TGG CTG GTC GAC CAG GGG TAT GCG GCA GTT TGC CCT GAT CTG TAC GCG CGC CAG GCG CCA GGT ACA GCA CTC GAT CCG CAG GAT            |   |        |  | 760  |
| Val Ser Trp Leu Val Asp Gln Gly Tyr Ala Ala Val Cys Pro Asp Leu Tyr Ala Arg Gln Ala Pro Gly Thr Ala Leu Asp Pro Gln Asp            |   |        |  |      |
|  | 50  |        |  |      |
| GAG CGC CAG AGA GAG CAA GCC TAC AAG CTC TGG CAG GCC TTC GAC ATG GAG GCC GGC GTG GGC GAT CTG GAG GCT GCT ATC CGC TAT GCG            |   |        |  | 850  |
| Glu Arg Gln Arg Glu Gln Ala Tyr Lys Leu Trp Gln Ala Phe Asp Met Glu Ala Gly Val Gly Asp Leu Glu Ala Ala Ile Arg Tyr Ala            |   |        |  |      |
|  | 80  |        |  |      |
| CGA CAC CAA CCC TAC AGC AAC GGC AAG GTG GGA TTG GTG GGG TAT TGC CTG GGC GGT GCG CTT GCC TTT CTA GTG GCC GCC AAA GGA TAC            |   |        |  | 940  |
| Arg His Gln Pro Tyr Ser Asn Gly Lys Val Gly Leu Val Gly Tyr Cys Leu Gly Gly Ala Leu Ala Phe Leu Val Ala Ala Lys Gly Tyr            |   |        |  |      |
|  | 110   |        |  |      |
| GTG GAT CGC GCC GTA GGC TAC TAC GGT GTT GGA CTG GAG AAG CAG CTC AAG AAG GTC CCG GAA GTC AAG CAT CCG GCG TTG TTT CAC ATG            |   |        |  | 1030 |
| Val Asp Arg Ala Val Gly Tyr Tyr Gly Val Gly Leu Glu Lys Gln Leu Asn Lys Val Pro Glu Val Lys His Pro Ala Leu Phe His Met            |   |        |  |      |
|  | 140   |        |  |      |
|  | Pvu II  |        |  |      |
| GGC GGC CAA GAC CAC TTC GTG CCC GCG CCA AGC CGC CAG CTG ATT ACT GAA GGC TTC GGT GCC AAT CCA TTG CTG CAA GTG CAC TGG TAC            |   |        |  | 1120 |
| Gly Gly Gln Asp His Phe Val Pro Ala Pro Ser Arg Gln Leu Ile Thr Glu Gly Phe Gly Ala Asn Pro Leu Leu Gln Val His Trp Tyr            |   |        |  |      |
|  | 170   |        |  |      |
| GAA GAG GCC GGA CAC TCG TTC GCC AGG ACG AGC AGT TCG GGC TAT GTG GCG AGT GCC GCG GCG TTG GCC AAC GAA CGT AGA CTG GAT TTC            |   |        |  | 1210 |
| Glu Glu Ala Gly His Ser Phe Ala Arg Thr Ser Ser Ser Gly Tyr Val Ala Ser Ala Ala Ala Leu Ala Asn Glu Arg Thr Leu Asp Phe            |   |        |  |      |
|  | 200   |        |  |      |
| CTG GCG CCC TTG CAG AGC AAG AAG CCA TGA ATTTTCATTCACGACTATCGTTCCCGCGGTGATTTTGGACCCGACAGCTCGCCCGATTGCCGAGGAAGTGGAACTCT              |   |        |  | 1320 |
| Leu Ala Pro Leu Gln Ser Lys Lys Pro *  |   |        |  |      |
|  | 230   |        |  |      |
| CGGCATTGACCGGGCGCTCGTGTGACCACGCTGACGAGGGCCCTGGGACGGCAAGTAGCCGAGCCGCTCATCGGCCAGTGGCAGCCTTGTACGATGGTGGCCATGCATGTACC                  |   |        |  | 1440 |
| CGCTTGGTGGCTGAGGAGCCCTGCAAGATTGCCCGCACGAGTGAAGGCTAATCGCCGGTCAATCGCGGATTCGCGATTGGTGGGGGATCTACTATCGGACTCGGCTAAGATCGTGGCCCTCCGTACAGAG |   |        |  | 1560 |
| TTGCCATCGTGGCTGTGCCACGACATATCAGGTAGCGAGATGACGTCATCTCGGTATCACAGAGGGTGGCGTCAAGAAAACCGCCGCGACGCGCGTGCATGCCCGTGCAGTT                   |   |        |  | 1680 |
| ATCTACGAGCCCGGCTGACCTAGAGCTACCCCTTAGCATCAGCGTGACCAAGCGGATCAATGCAATAGCCCATGCAGTTGAAGCCCTGTATGCCCCAGATGCCACGCGCTGCTTACC              |   |        |  | 1800 |
|  | 210   |        |  |      |
|  | Bgl II  |        |  |      |
| ATCATGCCCCAGGAAGAAATCGCAGCCAGCGTACCGCGATTTCAGAAATGTATCAATCACCAGAGATCT  |   |        |  | 1871 |

FIG. 4. Nucleotide sequence of the *clcD* gene coding and flanking regions and deduced amino acid sequences of dienelactone hydrolase. Numbering of the amino acid sequence begins with the first methionine amino acid residue determined by the NH<sub>2</sub>-terminal amino acid sequence of the enzyme purified from *Pseudomonas* sp. strain B13 (20). Major restriction sites are indicated over the DNA sequence. Overlined upstream from the structural gene is a putative Shine-Dalgarno sequence, and downstream is a putative rho-independent transcription termination sequence.

evidence indicates that the genes may be controlled as a transcriptional unit. Comparable specific activity for the *clcD* gene product (1.8 U/mg of protein) was observed in extracts of *P. putida* PRS2015 containing the plasmid pAC27 after growth with chlorobenzoate. Thus, the *tac* promoter in *E. coli* can be as effective as the natural promoter for the *clc* gene cluster on the plasmid pAC27. Purification of about 35-fold from induced cultures of either *E. coli* AC80 with recombinant plasmid pDC100 or *P. putida* PRS2015 with plasmid pAC27 yielded a homogeneous preparation of dienelactone hydrolase. Properties of the purified hydrolases were compared with those of dienelactone hydrolase encoded on plasmid pWR1 in *Pseudomonas* sp. strain B13 (20). The enzymes from the three biological sources were indistinguishable in every respect determined: the specific activities of the purified enzymes, the pH optima, the relative activities with *cis* and *trans* substrates, and the molecular weights of 28,000 (20).

**Nucleotide sequence of the *clcD* structural gene.** Sequencing of the 1.9-kbp *Pst*I-*Eco*RI fragment of the recombinant plasmid pDC100 revealed an open reading frame of 711 bases (Fig. 4), which can be assigned to the *clcD* structural gene based on the following evidence. The nucleotide sequence allows us to deduce a 236-residue amino acid sequence with an amino acid composition that closely resembles that determined for the protein purified from *E. coli* (Table 2). The first 34 residues of the amino acid sequence from the nucleotide sequence are identical to the NH<sub>2</sub>-terminal amino acid sequence of dienelactone hydrolase purified from *Pseu-*

*domonas* sp. strain B13 (20) (Fig. 4). The calculated molecular weight of the *clcD* gene product is 25,489, somewhat lower than the *M<sub>r</sub>* of 28,000 that is consistently observed after sodium dodecyl sulfate-polyacrylamide gel electrophoresis of dienelactone hydrolase.

## DISCUSSION

**Properties of the *clcD* gene.** The G+C content of the *clcD* gene is 59.8%, similar to the 60 to 62.5% G+C content of the *P. putida* chromosome (29). Thus the chromosomal G+C content of the host appears to be maintained even at the level of a structural gene carried on a plasmid. As would be predicted from the high G+C content of the *clcD* gene, its codon usage is highly biased (68.6%) in favor of G or C in the wobble base as determined for other *Pseudomonas* genes (12). The 5' region flanking the *clcD* gene contains the putative Shine-Dalgarno sequence (4, 27) GGAGAGA 8 base pairs upstream from the initiation codon ATG (Fig. 4). A putative rho-independent transcription termination sequence (1, 6, 23) was found in the 3' flanking region 20 base pairs downstream from the termination codon TGA of the *clcD* gene. The putative termination sequence contains a region of GC-rich dyad symmetry followed by the sequence TTTTT (Fig. 4).

**Properties of dienelactone hydrolase.** Dienelactone hydrolase encoded by the plasmid pAC27 *clcD* structural gene in *P. putida* is a single polypeptide chain with 236 amino acid residues and an *M<sub>r</sub>* of 25,489. The physical and



FIG. 5. Similar amino acid sequences lying near cysteinyl residues in dienelactone hydrolase and enol-lactone hydrolase. The cysteinyl residue occupies position 60 in the dienelactone hydrolase sequence. The sequence from enol-lactone hydrolase was determined with a tryptic peptide, and its location within the enzyme is not known.

chemical properties of the protein did not appear to be altered when the gene was expressed in *E. coli*. Because the protein can be readily purified from *E. coli* organisms, it is a promising subject for analysis of the structural basis of its substrate specificity by directed mutagenesis.

The dienelactone hydrolases from two *Pseudomonas* strains, isolated on separate continents, appear to be identical in every discernible respect. It is noteworthy that the NH<sub>2</sub>-terminal amino acid sequence determined for the hydrolase from *Pseudomonas* sp. strain B13 is identical to the amino acid sequence deduced from the DNA sequence of the plasmid pAC27 *clcD* structural gene.

The available evidence indicates that the dienelactone hydrolase and enolactone hydrolase of *P. putida* have diverged widely. The NH<sub>2</sub>-terminal amino acid sequences of the enzymes are dissimilar, even though the proteins share similar molecular weights and amino acid compositions (20). Both hydrolases are inactivated by stoichiometric amounts of *p*-chloromercuribenzoate (20, 33), suggesting that each hydrolase contains a cysteinyl side chain lying at or near the active site (20). The complete amino acid sequence of the enol-lactone hydrolase is not yet known, but Wu-Kuang Yeh has determined the amino acid sequence of cysteine-containing tryptic peptides for the enzyme and has generously shared this information with us. The amino acid sequence surrounding one of the cysteine residues in enol-lactone

hydrolase closely resembles the amino acid sequence neighboring Cys-60 in dienelactone hydrolase (Fig. 5). Conservation of these sequences suggests that they have been constrained against divergence, probably because they play a critical role in the function of the hydrolases.

Further information on mechanisms that underlie divergence of the hydrolase structural genes awaits determination of the complete sequence of the *Pseudomonas* enol-lactone hydrolase structural genes, and this investigation is in progress.

#### ACKNOWLEDGMENTS

This work was supported by the National Science Foundation grant DMB-8514671 and Public Health Service grants ES04050 from the National Institute of Environmental Health Sciences (to A.M.C.) and GM-33377 from the National Institutes of Health (to L.N.O.).

We thank Juan V. Parales for growing cells at the Fermentation Facility at Yale University and Han-Joachim Knackmuss for providing us with *Pseudomonas* sp. B13 strain and a generous supply of the dienelactones. We are grateful to Wu-Kuang Yeh for providing peptide amino acid sequences of *Pseudomonas* enol-lactone hydrolase and to Michael Schlömann, John Houghton, and David Ollis for valuable discussions.

#### LITERATURE CITED

- Adhya, S., and M. Gottesman. 1978. Control of transcription termination. *Annu. Rev. Biochem.* **47**:967-996.
- Appleyard, R. K. 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K-12. *Genetics* **29**:440-445.
- Bagdasarian, M. M., E. Amann, R. Lurz, B. Ruckert, and M. Bagdasarian. 1983. Activity of the hybrid *trp-lac(tac)* promoter of *Escherichia coli* in *Pseudomonas putida*: construction of broad host-range, controlled-expression vectors. *Gene* **26**:273-282.
- Bassel, B. A., and D. R. Mills. 1979. Initiation of translation with *Pseudomonas aeruginosa* phage PP7 RNA: nucleotide sequence of the coat cistron ribosome binding site. *Nucleic Acids Res.* **6**:2033-2036.
- Bradford, M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**:248-254.
- Brendl, V., G. H. Hamm, and E. N. Trifonov. 1986. Terminators of transcription with RNA polymerase from *Escherichia coli*: what they look like and how to find them. *J. Biomol. Struct. Dyn.* **3**:705-723.
- Chakrabarty, A. M., D. A. Friello, and L. H. Bopp. 1978. Transposition of plasmid DNA segments specifying hydrocarbon degradation and their expression in various microorganisms. *Proc. Natl. Acad. Sci. USA* **75**:3109-3112.
- Chatterjee, D. K., and A. M. Chakrabarty. 1983. Genetic homology between independently isolated chlorobenzoate-degradative plasmids. *J. Bacteriol.* **153**:532-534.
- Chatterjee, D. K., and A. M. Chakrabarty. 1984. Restriction mapping of a chlorobenzoate degradative plasmid and molecular cloning of the degradative genes. *Gene* **27**:173-181.
- Chatterjee, D. K., S. T. Kellogg, S. Hamada, and A. M. Chakrabarty. 1981. Plasmid specifying total degradation of 3-chlorobenzoate by a modified *ortho* pathway. *J. Bacteriol.* **146**:639-646.

TABLE 2. Comparison of determined and predicted amino acid composition of *Pseudomonas* plasmid pAC27 dienelactone hydrolase

| Amino acid       | No. of residues         |                        |
|------------------|-------------------------|------------------------|
|                  | Determined <sup>a</sup> | Predicted <sup>b</sup> |
| Cys              | ND                      | 2                      |
| Asn <sup>c</sup> |                         | 10                     |
| Asp <sup>c</sup> | 16.3                    | 5                      |
| Thr              | 7.7                     | 7                      |
| Ser              | 10.8                    | 12                     |
| Gln <sup>d</sup> |                         | 14                     |
| Glu <sup>d</sup> | 26.9                    | 13                     |
| Pro              | 17.4                    | 14                     |
| Gly              | 24.5                    | 24                     |
| Ala              | 32.6                    | 33                     |
| Val              | 15.6                    | 19                     |
| Met              | 3.4                     | 4                      |
| Ile              | 5.9                     | 7                      |
| Leu              | 21.4                    | 21                     |
| Tyr              | 8.0                     | 12                     |
| Phe              | 8.9                     | 10                     |
| His              | 4.8                     | 7                      |
| Lys              | 9.2                     | 9                      |
| Arg              | 9.4                     | 10                     |
| Trp              | ND                      | 3                      |

<sup>a</sup> Based on a total of 227 residues, with predicted rather than experimentally determined values used for Cys and Trp; ND, not determined.

<sup>b</sup> The total number of predicted residues from the DNA sequence was 236.

<sup>c</sup> The predicted value for Asx (Asn plus Asp) was 15 residues.

<sup>d</sup> The predicted value for Glx (Gln plus Glu) was 27 residues.

11. Dale, R. M. K., B. A. McClure, and J. P. Houchins. 1985. A rapid single-stranded colony strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* **13**:31–40.
12. Darzins, A., B. Frantz, R. I. Vanags, and A. M. Chakrabarty. 1986. Nucleotide sequence analysis of the phosphomannose isomerase gene (*pmi*) of *Pseudomonas aeruginosa* and comparison with the corresponding *Escherichia coli* gene *manA*. *Gene* **42**:293–302.
13. Davis, R. W., D. Botstein, and J. R. Roth. 1980. A manual for genetic engineering: advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. Dorn, E., and H.-J. Knackmuss. 1978. Chemical structure and biodegradability of halogenated aromatic compounds: two catechol 1,2-dioxygenases from a 3-chloro-benzoate-grown pseudomonad. *Biochem. J.* **174**:73–84.
15. Ghosal, D., I.-S. You, D. K. Chatterjee, and A. M. Chakrabarty. 1985. Genes specifying degradation of 3-chlorobenzoic acid in plasmids pAC27 and pJP4. *Proc. Natl. Acad. Sci. USA* **82**:1638–1642.
16. Henrikson, R. L., and S. C. Meredith. 1984. Amino acid analysis by reverse-phase high-performance liquid chromatography: precolumn derivatization with phenylisothiocyanate. *Anal. Biochem.* **136**:65–74.
17. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. Messing, J. 1979. A multi-purpose cloning system based on the single-stranded DNA bacteriophage M13. *Recomb. DNA Tech. Bull.* **2**:43–48.
19. Nakamura, K., and Y. Kato. 1985. Preparative high-performance ion-exchange chromatography. *J. Chromatogr.* **333**:29–40.
20. Ngai, K.-L., M. Schlömann, H.-J. Knackmuss, and L. N. Ornston. 1986. Dienelactone hydrolase from *Pseudomonas* sp. strain B13. *J. Bacteriol.* **169**:699–703.
21. Norrander, J., T. Kempe, and J. Messing. 1983. Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. *Gene* **26**:101–106.
22. Reineke, W., and H.-J. Knackmuss. 1979. Construction of haloaromatics-utilising bacteria. *Nature (London)* **277**:385–386.
23. Rosenberg, M., and D. Court. 1979. Regulatory sequences involved in the promotion and termination of RNA transcription. *Annu. Rev. Genet.* **13**:319–353.
24. Sancar, A., A. M. Hack, and W. D. Rupp. 1979. Simple method for identification of plasmid-encoded proteins. *J. Bacteriol.* **137**:692–693.
25. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
26. Schmidt, E., and H.-J. Knackmuss. 1980. Chemical structure and biodegradability of halogenated aromatic compounds. *Biochem. J.* **192**:339–347.
27. Shine, J., and L. Dalgarno. 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature (London)* **254**:34–38.
28. Spector, T. 1978. Refinement of Coomassie blue method of protein quantitation. *Anal. Biochem.* **86**:142–146.
29. Stanier, R. Y., N. J. Palleroni, and M. Doudoroff. 1966. The aerobic pseudomonads: a taxonomic study. *J. Gen. Microbiol.* **43**:159–271.
30. Stout, R. W., and J. J. DeStefano. 1985. A new, stabilized, hydrophilic silica packing for the high-performance gel chromatography of macromolecules. *J. Chromatogr.* **326**:63–78.
31. Weber, K., and M. Osborn. 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Biol. Chem.* **244**:4406–4412.
32. Wheelis, M. L., and L. N. Ornston. 1972. Genetic control of enzyme induction in the  $\beta$ -ketoacid pathway of *Pseudomonas putida*: deletion mapping of *cat* mutations. *J. Bacteriol.* **109**:790–795.
33. Yeh, W.-K., and L. N. Ornston. 1984. *p*-Chloromercuribenzoate specifically modifies thiols associated with the active sites of  $\beta$ -ketoacid enol-lactone hydrolase and succinyl CoA: $\beta$ -ketoacid CoA transferase. *Arch. Microbiol.* **138**:102–105.