JOURNAL OF BACTERIOLOGY, Mar. 1998, p. 1082–1094 Vol. 180, No. 5
0021-9193/98/$04.00+0
Copyright © 1998, American Society for Microbiology

Evolutionary Relationship between Chlorocatechol Catabolic Enzymes from *Rhodococcus opacus* 1CP and Their Counterparts in Proteobacteria: Sequence Divergence and Functional Convergence

DIRK EULBERG,[1] ELENA M. KOURBATOVA,[2]† LUDMILA A. GOLOVLEVA,[2] AND MICHAEL SCHLÖMANN[1]*

Institut für Mikrobiologie, Universität Stuttgart, D-70550 Stuttgart, Germany,[1] and Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, Pushchino, Russia[2]

Received 8 September 1997/Accepted 12 December 1997

Biochemical investigations of the muconate and chloromuconate cycloisomerases from the chlorophenol-utilizing strain *Rhodococcus opacus* (*erythropolis*) 1CP had previously indicated that the chlorocatechol catabolic pathway of this strain may have developed independently from the corresponding pathways of proteobacteria. To test this hypothesis, we cloned the chlorocatechol catabolic gene cluster of strain 1CP by using PCR with primers derived from sequences of N termini and peptides of purified chlorocatechol 1,2-dioxygenase and chloromuconate cycloisomerase. Sequencing of the clones revealed that they comprise different parts of the same gene cluster in which five open reading frames have been identified. The *clcB* gene for chloromuconate cycloisomerase is transcribed divergently from a gene which codes for a LysR-type regulatory protein, the presumed ClcR. Downstream of *clcR* but separated from it by 222 bp, we detected the *clcA* and *clcD* genes, which could unambiguously be assigned to chlorocatechol 1,2-dioxygenase and dienelactone hydrolase. A gene coding for a maleylacetate reductase could not be detected. Instead, the product encoded by the fifth open reading frame turned out to be homologous to transposition-related proteins of IS*1031* and Tn*4811*. Sequence comparisons of ClcA and ClcB to other 1,2-dioxygenases and cycloisomerases, respectively, clearly showed that the chlorocatechol catabolic enzymes of *R. opacus* 1CP represent different branches in the dendrograms than their proteobacterial counterparts. Thus, while the sequences diverged, the functional adaptation to efficient chlorocatechol metabolization occurred independently in proteobacteria and gram-positive bacteria, that is, by functionally convergent evolution.

Although various chloroaromatic compounds are known to be degraded slowly or incompletely under certain conditions (23), numerous bacteria are able to utilize some representatives of this class of chemicals as sole sources of carbon and energy (26, 54). In many of the cases where growth occurs, the chloroaromatic substrates are initially converted to chlorocatechols. These can be cleaved in *ortho* position, i.e., between the hydroxyl groups, to chlorosubstituted *cis*,*cis*-muconates. Cycloisomerization of the latter and concomitant chloride elimination will yield dienelactones (4-carboxymethylenebut-2-en-4-olides) which may still carry the remaining chlorine substituents (Fig. 1). Subsequent hydrolysis of the lactone and reduction of the resulting maleylacetate, in most cases, generate 3-oxoadipate, an intermediate also of the catechol and the protocatechuate branch of the ubiquitous 3-oxoadipate pathway (Fig. 1).

As initially shown especially for the 3-chlorobenzoate-utilizing *Pseudomonas* sp. strain B13 and for the 2,4-dichlorophenoxyacetate-degrading strain *Ralstonia eutropha* (*Alcaligenes eutrophus*) JMP134(pJP4), different sets of enzymes are induced for catechol catabolism on the one hand and for chlorocatechol catabolism on the other (13, 39, 53, 59, 60). Thus, chlorocatechol 1,2-dioxygenase (EC 1.13.11.-), chloromuconate cycloisomerase (EC 5.5.1.7), and dienelactone hydrolase (EC 3.1.1.45) catalyze reactions analogous to those of catechol 1,2-dioxygenase (EC 1.13.11.1), muconate cycloisomerase (EC 5.5.1.1), and 3-oxoadipate enol-lactone hydrolase (EC 3.1.1.24). They differ from the latter, however, in having higher relative turnover numbers and/or lower K_m values for chlorosubstituted substrates or the metabolites generated from them by chloride elimination. The fourth enzyme characteristic for chlorocatechol catabolism, maleylacetate reductase (EC 1.3.1.32), has no equivalent among catechol catabolic enzymes and compensates for the different oxidation states of the metabolites of the two pathways.

Genetic studies, which have so far focused on gram-negative bacteria, have shown that the enzymes of chlorocatechol pathways are often encoded by catabolic plasmids such as pJP4 of *R. eutropha* JMP134 (11), pAC27 of the 3-chlorobenzoate-utilizing *Pseudomonas putida* AC866 (6), pP51 of the 1,2,4-trichlorobenzene-mineralizing *Pseudomonas* sp. strain P51 (68), and pEST4011 isolated from the 2,4-dichlorophenoxyacetate-degrading *P. putida* EST4021 (42). Several years ago, sequencing of chlorocatechol gene clusters from the former three plasmids (20, 52, 69) showed the chlorocatechol 1,2-dioxygenases and chloromuconate cycloisomerases to be homologous to catechol 1,2-dioxygenases and muconate cycloisomerases, respectively. More interestingly, the chlorocatechol 1,2-dioxygenases of these plasmids were derived from a common ancestor not shared by the catechol 1,2-dioxygenases, as the chloromuco-

* Corresponding author. Mailing address: Institut für Mikrobiologie, Universität Stuttgart, Allmandring 31, D-70569 Stuttgart, Germany. Phone: (49)-711-6855489. Fax: (49)-711-6855725. E-mail: imbms@po.uni-stuttgart.de.

† Present address: Pushchino State University, Genetic Engineering Plant Biotechnology Department, Pushchino, Russia.

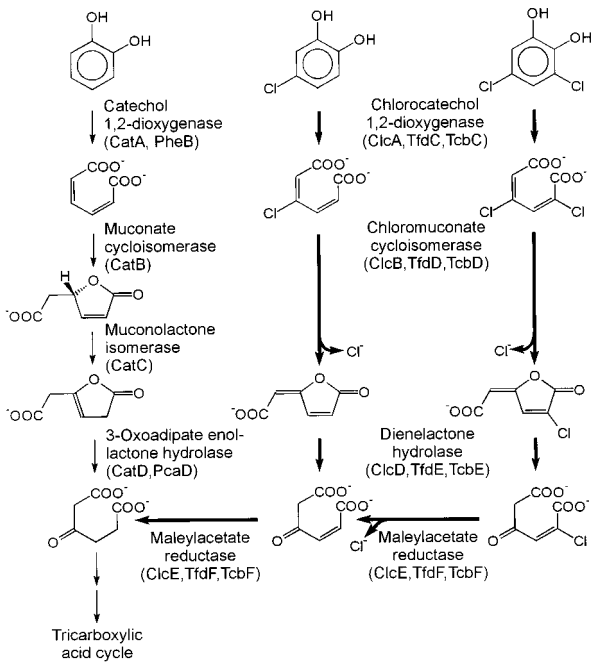

FIG. 1. Degradative pathways for catechol, 4-chlorocatechol, and 3,5-dichlorocatechol. Designations as gene products, as previously assigned for gram-negative or gram-positive bacteria, are given in parentheses. Enzymes with a specific function in chlorocatechol catabolism are set off by heavy arrows.

TABLE 1. Plasmids used in this study

Plasmid	Relevant characteristics[a]	Source (reference)
pBluescript II KS (+)	Phagemid derived from pUC19; P_{lac} *lacZ′* Ap^r; f1(+) origin; ColE1 origin	Stratagene (2)
pRERD1	Ap^r; ca. 220-bp PCR product obtained with dioxygenase-derived primers in *Eco*RV site of pBluescript II KS (+)	This study
pRERC1	Ap^r; ca. 60-bp PCR product obtained with cycloisomerase-derived primers in *Eco*RV site of pBluescript II KS (+)	This study
pRER5	Ap^r; ca. 2.9-kb *Pst*I fragment of *R. opacus* 1CP DNA in pBluescript II KS (+)	This study
pRER6	Ap^r; ca. 10-kb *Hin*dIII fragment of *R. opacus* 1CP DNA in pBluescript II KS (+)	This study
pRER61	Ap^r; ca. 2.4-kb *Sac*II fragment of pRER6 in pBluescript II KS (+)	This study
pRER7	Ap^r; ca. 3.7-kb *Eco*RI fragment of *R. opacus* 1CP DNA in pBluescript II KS (+)	This study

[a] Ap^r, plasmid carries information for ampicillin resistance.

nate cycloisomerases diverged from an ancestral enzyme not in the line of descent of known muconate cycloisomerases (58). The dienelactone hydrolases are evolutionarily so remote from the 3-oxoadipate enol-lactone hydrolases that their common evolutionary background could be identified only by their three-dimensional structure or sequence signatures related to it (29, 50, 58). These observations, as well as conserved structures of the gene clusters, suggested that the proteobacterial chlorocatechol pathways, as represented by those on pAC27, pJP4, and pP51, diverged from the same ancestral pathway (58). The catabolic genes from many other strains were shown by hybridization to be related to those of the plasmids mentioned above (4, 21, 36, 42, 65). In other cases, no hybridization was found under the conditions used (4, 21, 34, 36).

The gram-positive bacterium *Rhodococcus opacus* (*erythropolis*) 1CP has previously been isolated with 2,4-dichlorophenol as the carbon source (25). Purification and characterization of its chlorocatechol 1,2-dioxygenase, chloromuconate cycloisomerase, and dienelactone hydrolase revealed unusual prop-

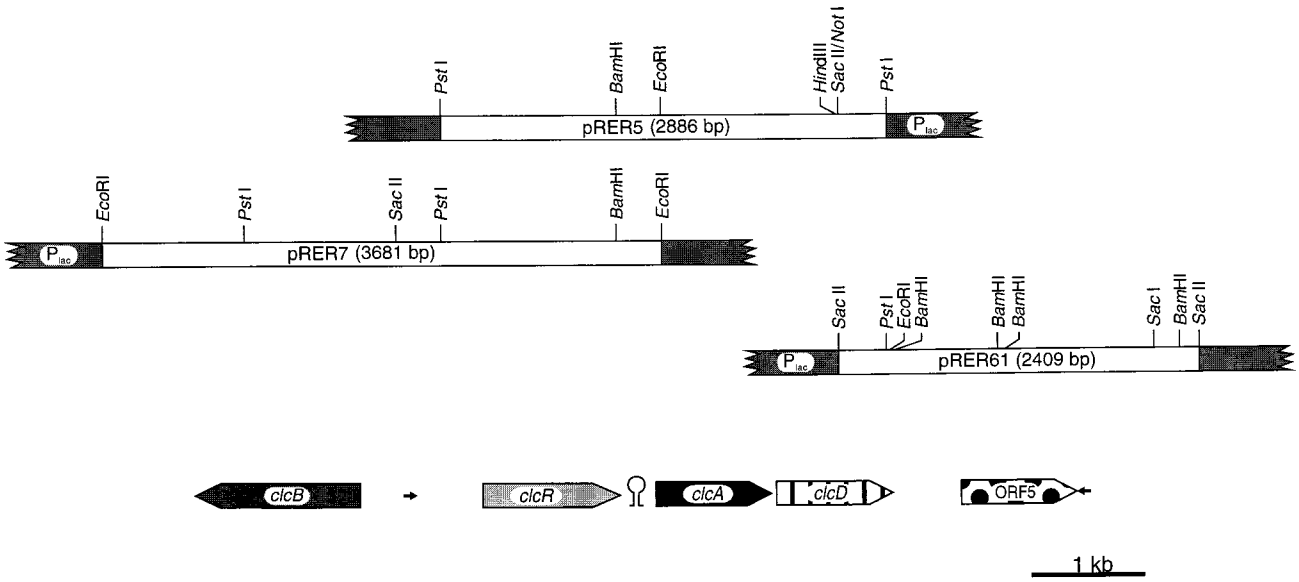

FIG. 2. Restriction map of the inserts of pRER5, pRER7, and pRER61 carrying 7,240 bp of *R. opacus* 1CP DNA in the multiple cloning site of pBluescript II KS (+). The lengths and orientations of open reading frames are indicated by large arrows. Two small arrows indicate the locations of the imperfect inverted repeats. Between the presumed *clcR* and *clcA* was found a hairpin structure that possibly represents a transcription terminator.

```
   1 GAATTCAATGCCATCCCGATCGAGAAATACGCCGTCACGCGAGCGGTCAACGACCTACTGCGCTGCTCGGTCCGCCCGCATTCCGCGATCACCTCATCGACCAGGGACGCGGGAAACACC
 121 CTGGTCAACAACCCCACCGACACCACATCCGACAACCGCCGATCCGAGACCGGTTTCACCCATCCGCTCCGAGGCACACCTCATCGTAACGCAACCCGCACTAACTGAACGGTATTGCAC
 241 CTAAACGACCACAGCACTGCCTGCGGCCAGTTCATCGTGTACGTATCGTCTGCCGCCGGGCGACTGTGGTGTCCGGTGCACGGCGTAGCCCATGAACCGTCGGGTCGATCGGCTCCGCAC
 361 AGCAAATACTGCACTTTAGCGCTCCCGGAAGGTGCAAGGTGATGGGAGCACGGGCTCCCATGCTGCAAGTTGCGCTGTTCGTCGGGGCTCGTCAGGATCAATTCGACTACGTCGTCCGCG

 481 ATCTTGTTCTTCTCGGAGACGACAAGTTCGCGGCGGGTGCCGGGTTTGTGCATGGATACGTAACCTGTCACGAGTGGAAGGTTGGGCGGCGTAGGCCGTCAGCGTCTGCGGTAGCGGTCG
                                                                                                                  AGTCGCAGACGCCATCGCCAGC
                                                                                                                 STOP  R  R  R  Y  R  D

 601 ACCTTGTCGTGGTCCAGCCGCACCCCCAGACCGGGACCCGTTGGAACCGTGAGTATTCCGTCAGTATAGGCGAGCGGCTCGGTGATCACCGAGTCGGCCAGCAGCAACGGGCCCACCAGC
     TGGAACAGCACCAGGTCGGCGTGGGGGTCTGGCCCTGGGCAACCTTGGCACTCATAAGGCAGTCATATCCGCTCGCCGAGCCACTAGTGGCTCAGCCGGTCGTCGTTGCCCGGGTGGTCG
      V  K  D  H  D  L  R  V  G  L  G  P  G  T  P  V  T  L  I  G  D  T  Y  A  L  P  E  T  I  V  S  D  A  L  L  L  P  G  V  L

 721 TCGCAGCCGTGGAGCAGCCTCGGGAATGTGGCGAACAGGTGCGCGGAGGCCGCCGTACCGACCGGGCTTTCCAGCGCCGTACCGCCGTAAGGACCGATCGAAGCGGTGTCGAGTATGGCC
     AGCGTCGGCACCTCGTCGGAGCCCTTACACCGCTTGTCCACGCGCCTCCGGCGGCATGGCTGGCCCGAAAGGTCGCGGCATGGCGGCATTCCTGGCTAGCTTCGCCACAGCTCATACCGG
      E  C  G  H  L  L  R  P  F  T  A  F  L  H  A  S  A  A  T  G  V  P  S  E  L  A  T  G  G  Y  P  G  I  S  A  T  D  L  I  A

 841 GCGATCAACGCCGCAGCCCGCAAGCCTCCGGCCTTTCCGGGTTTGACCGCGACCGAGTCGCAGGCTCGTTCGCGTACGACCTCGAGCGCGTCTGCAGGCACGCATACCGACTCGTCGGCC
     CGCTAGTTGCGGCGTCGGGCGTTCGGAGGCCGGAAAGGCCCAAACTGGCGCTGGCTCAGCGTCCGAGCAAGCGCATGCTGGAGCTCGCGCAGACGTCCGTGCGTATGGCTGAGCAGCCGG
      A  I  L  A  A  A  R  L  G  G  A  K  G  P  K  V  A  V  S  D  C  A  R  E  R  V  V  E  L  A  D  A  P  V  C  V  S  E  D  A

 961 ATGATGTCGATGCGATCGGCTCGCGACCGTAGTTCGGCAAGCCCGCGCACGTCGCGGCGCTCGACCGGCTGCTCGACCGCGCTGAGGCCGATGGCCTCGAGTTCTCGCACGGCCCAGGTC
     TACTACAGCTACGCTAGCCGAGCGCTGGCATCAAGCCGTTCGGGCGCGTGCAGCGCCGCGAGCTGGCCGACGAGCTGGCGCGACTCCGGCTACCGGAGCTCAAGAGCGTGCCGGGTCCAG
      M  I  D  I  R  D  A  R  S  R  L  E  A  L  G  R  V  D  R  R  E  V  P  Q  E  V  A  S  L  G  I  A  E  L  E  R  V  A  W  T

1081 GCAGTCCGCAGATCCCACGTGCCGTTGGGGTCGCCAAGGTAGTCGACGCCGTCGCCGATCTTGTCGACCAGTTCCCCGACGCGGTCGATATCCTCGCGCGGGGCAAGCGCCCCCATCTTG
     CGTCAGGCGTCTAGGGTGCACGGCAACCCCAGCGGTTCCATCAGCTGCGGCAGCGGCTAGAACAGCTGGTCAAGGGGCTGCGCCAGCTATAGGAGCGCGCCCCGTTCGCGGGGGTAGAAC
      A  T  R  L  D  W  T  G  N  P  D  G  L  Y  D  V  G  D  G  I  K  D  V  L  E  G  V  R  D  I  D  E  R  P  A  L  A  G  M  K

1201 AACTTGATCCTGCGGTAGCCAGCCGCCACTCGTTCCGCAGCCTCTTCGAGTGTCCCCCCGGTCCCGGCGCCGCTGAGCGGCCAGCGCAGCGGCAGTTCGGTACGGGTGGCACCGCCTAGC
     TTGAACTAGGACGCCATCGGTCGGCGGTGAGCAAGGCGTCGGAGAAGCTCACAGGGGGGCCAGGGCCGCGGCGACTCGCCGGTCGCGTCGCCGTCAAGCCATGCCCACCGTGGCGGATCG
      F  K  I  R  R  Y  G  A  A  V  R  E  A  A  E  E  L  T  G  G  T  G  A  G  S  L  P  W  R  L  P  L  E  T  R  T  A  G  G  L

1321 AGGGTCGACAGGGCAACGTTGTGTGCCTGCCCGATCGCATCGAGCAACGCGGTCTCGAGCGCGGCCTTGGCGAACAGGTTGGCGAACGCAATCCGGTCTAGTTGTCGACGGGTCCCGGCG
     TCCCAGCTGTCCCGTTGCAACACACGGACGGGCTAGCGTAGCTCGTTGCGCCAGAGCTCGCGCCGGAACCGCTTGTCCAACCGCTTGCGTTAGGCCAGATCAACAGCTGCCCAGGGCCGC
      L  T  S  L  A  V  N  H  A  Q  G  I  A  D  L  L  A  T  E  L  A  A  K  A  F  L  N  A  F  A  I  R  D  L  Q  R  R  T  G  A

1441 AGGTCTTCGACGTCGCTGCCGATCAGTGCCGGGGCAAGGTAGCAGTCGATGATCTGCTTTTGCCCTTCGATACTTTCTCCGCCCCACCACGGTCCGCCTGGTGATACCCCTTCGCCGAGG
     TCCAGAAGCTGCAGCGACGGCTAGTCACGGCCCCGTTCCATCGTCAGCTACTAGACGAAAACGGGAAGCTATGAAAGAGGCGGGGTGGTGCCAGGCGGACCACTATGGGGAAGCGGCTCC
      L  D  E  V  D  S  G  I  L  A  P  A  L  Y  C  D  I  I  Q  K  Q  G  E  I  S  E  G  G  W  W  P  G  G  P  S  V  G  E  G  L

1561 CCGACGAATCCGGCGTCTGTGACGACCTCGACGAGGAGGTAAGTTTGTGCGTCGATGCTGTGGTTCGCGAATCGGTGTGGGCGCAGGATCGGTAGGTCGACGATCGTGGTCCGCACACCG
     GGCTGCTTAGGCCGCAGACACTGCTGGAGCTGCTCCTCCATTCAAACACGCAGCTACGACACCAAGCGCTTAGCCACACCCGCGTCCTAGCCATCCAGCTGCTAGCACCAGGCGTGTGGC
      G  V  F  G  A  D  T  V  V  E  V  L  L  Y  T  Q  A  D  I  S  H  N  A  F  R  H  P  R  L  I  P  L  D  V  I  T  T  R  V  G

1681 CTGACGGTGAGGTTCGGCATGGCCGCTCGCTCTCGTTGTGTAGTTAGGGACAAAAGGACCGGTGGGTCAGATCGGTGTGACAAGCAACGTTTGGTCCCGACTCGCAATCGCACCTCGGTG
     GACTGCCACTCCAAGCCGTA
      S  V  T  L  N  P  M   ← clcB

1801 GCGGTCACATCCGCGTCGGTGTCAGTCGGTCCATCCAGGCTTGGGAGTTGTCTCCGGTCTGCCAAACGGGTGGGACGGCCACGGCGTCCAACCCATTGCGGTGGGCCCGGGCCATGTGAC
1921 GACGGTAGGACCGCGGTGAGTGAGACCGCCAATACCCATGCAGTTGCCCGGGCCCCTGATCTCGGTATCGGCTGACATCCCAACAGCGGCCGTGTGCTTGACTGCGACCATGACATCGCC
2041 GATATCGGCGGCACCACTCGACGCCCGCAAAACCCGATGGCGGGGTGTAGTTGGGGCGGCGTGTTTCCTTCCCGTCTACCACCACGCACCATGGCGCTTTGTCCGACCTGAGGGATTCGA
2161 CCACGCGCTGATCAGCATCGGCGTCGACATTCGCTCGACCCCCGCCCTCGAGTACCTCGACCAGCTACCGCTGCAGTTCGTTCGGATCGTTGCACCCGACGAGGGCTCGGAAGGGTACTG
2281 CGATCTCGACTTCGCATACGCAGCGGTGCTGGCCAGCCATCACCGCCAGGCTGCCCGGGTGTTCCCCGACGGGCGGATCGCCGGCTCCGTCGCCGCCCCTGCTCCTGATTGCCAGCCTGC

2401 GCGCCGAGCTGCTCGTCGGGCTGAGATTCTCCTCGGCGGCACCGAGACCCAATCGGTACCGACATCTCATTCGCGCCAGGCTCCGGCCGGTGCGCTTACTATCCGATCATGATTCCTCTC
                                                                                                                clcR →  M  I  P  L

2521 GACGTACGCCAGTTGGAGTGTTTCATCGCTGTCGCCGAAGAACGCCACGTCGGCCATGCCGCAGCCCGGCTGCACATGACCCAGCCGCCTTTGACCCGCCGAATCTCGCGCCTCGAACGC
     D  V  R  Q  L  E  C  F  I  A  V  A  E  E  R  H  V  G  H  A  A  A  R  L  H  M  T  Q  P  P  L  T  R  R  I  S  R  L  E  R

2641 GAAATCGGCGCACAGCTGTTCGTGCGCACTCTGACCGGGGTCGAACTGACCGCCGCCGGGCTCGTACTGCTCGACCGCGCCTACCGAATTGTGGCGCTGGCCGAGAACACCCTCGAACGC
     E  I  G  A  Q  L  F  V  R  T  L  T  G  V  E  L  T  A  A  G  L  V  L  L  D  R  A  Y  R  I  V  A  L  A  E  N  T  L  E  R

2761 GTGCGCCTGGCCGACGCCGGTGAAGCGGGCGCGCTGGCCGTCGGCTACTTCGGCTCGACGATCTTCGGACTCGTCCCCCGCCTACTGCGTCTCTTCCTCGACACTCGCCCCCGAAGTGCGG
     V  R  L  A  D  A  G  E  A  G  A  L  A  V  G  Y  F  G  S  T  I  F  G  L  V  P  R  L  L  R  L  F  L  D  T  R  P  E  V  R

2881 ATCGTCATCGAACGCGCGACCAGGGACATGCAGGCCCAGGCCATTCGCGATGGCCGCATGCACATTGGATTCAGTCGCCAATACCCCGCCGAACCTGACCTGGCGGTGCGCAACATCGCC
     I  V  I  E  R  A  T  R  D  M  Q  A  Q  A  I  R  D  G  R  M  H  I  G  F  S  R  Q  Y  P  A  E  P  D  L  A  V  R  N  I  A

3001 TCCGAGCCGCTGTATGTGGCACTGCCCGAAACACATCCACTACTGAGCAAGCAGAAGGTGCTGCTCGCCGACATCCAGAACGAACCGCTGGTGCTGTTCCCTTCGGCACCCAGGCCAGGT
     S  E  P  L  Y  V  A  L  P  E  T  H  P  L  L  S  K  Q  K  V  L  L  A  D  I  Q  N  E  P  L  V  L  F  P  S  A  P  R  P  G

3121 TTCGCCGACGAAGTCGTTCAATTACTCATGCGCACCGGCGGTGAACCGCGCATCGAAGCCGAAGCAAGCGACGTCGTCACCGCCTTAGCCTACGTGGCCATCGCACAATTGTGCGCGGTG
     F  A  D  E  V  V  Q  L  L  M  R  T  G  G  E  P  R  I  E  A  E  A  S  D  V  V  T  A  L  A  Y  V  A  I  A  Q  L  C  A  V

3241 GTCCCCCGTTCCGCCAGCAACATTGCACTGCCCGGCGTGGTGTACGTCCCTCTGGCCGACGCCCCGCCGGAATCGCTGAGCTGCATCTACCGGGCAATCGAGACACCGACGTTGGTGCGT
     V  P  R  S  A  S  N  I  A  L  P  G  V  V  Y  V  P  L  A  D  A  P  P  E  S  L  S  C  I  Y  R  A  I  E  T  P  T  L  V  R

3361 GCCTTCCTCGACCATCTTGACCTCCTGGATCCGGCGACGCTCGACCTATACGAGGACTGACCGATACCCACACCGACGTGACGACGGTGCGGTGTTGGTATTGGCCCACCCGAGCTGCGG
     A  F  L  D  H  L  D  L  L  D  P  A  T  L  D  L  Y  E  D  STOP

3481 CTGGCTACGCTTCACCCAGGCGCATCAAAGCCCCTCCGTCCCCCGCAGCCCCTTCCGACGCGAGGCCCGTCGCGACAGCAACGGATTCGCCTTTCCGCGCGTGCAGTGCCTCTTTGTCGC
                                                                                                                ────────────►

3601 CGGCAAAGAGGCACTGCGTAAGAGATATCCAAGGAGTTAAACATGGCCAATACACGCGTGATCGAGTTGTTCGACGAATTCACCGATCTCATCCGTGATTTCATCGTCCGTCACGAGATC
     ◄────────────────────────────────────────────────────────────────────────────────────────────────────────────────────────
                                                clcA →  M  A  N  T  R  V  I  E  L  F  D  E  F  T  D  L  I  R  D  F  I  V  R  H  E  I
```

FIG. 3. Sequence of the 7,240 bp of *R. opacus* 1CP DNA cloned in pRER5, pRER7, and pRER61. The predicted amino acid sequences of the open reading frames are shown below the DNA sequence. Amino acid sequences obtained by direct sequencing of proteins or tryptic peptides are underlined. For *clcB*, which is divergently transcribed from all other open reading frames, the complementary (in this case, coding) strand is also given. A segment between *clcR* and *clcA*, which forms a hairpin structure and possibly represents a transcription terminator, is indicated by two arrows of opposite orientation. An imperfect inverted repeat is represented by underlined and italicized bases between *clcB* and *clcR* as well as downstream of ORF5.

```
3721 ACCACACCGGAGTACGAAACGATCATGCAATACATGATCAGCGTCGGCGAGGCCGGCGAGTGGCCGCTGTGGCTCGACGCGTTCTTCGAAACCACCGTCGACTCCGTGAGCTACGGCAAG
      T  T  P  E  Y  E  T  I  M  Q  Y  M  I  S  V  G  E  A  G  E  W  P  L  W  L  D  A  F  F  E  T  T  V  D  S  V  S  Y  G  K

3841 GGAAACTGGACCTCCAGTGCGATCCAAGGACCGTTCTTCAAAGAAGGTGCCCCGCTGCTGACCGGCAAACCGGCCACCCTGCCGATGCGGGCAGACGAGCCTGGCGACCGGATGCGGTTC
      G  N  W  T  S  S  A  I  Q  G  P  F  F  K  E  G  A  P  L  L  T  G  K  P  A  T  L  P  M  R  A  D  E  P  G  D  R  M  R  F

3961 ACCGGGAGCGTGCGCGACACGTCGGGAACTCCGATCACCGGCGCGGTCATCGATGTGTGGCACTCCACCAACGACGGCAACTACAGCTTTTTCTCCCCCGCTCTGCCCGACCAGTACCTG
      T  G  S  V  R  D  T  S  G  T  P  I  T  G  A  V  I  D  V  W  H  S  T  N  D  G  N  Y  S  F  F  S  P  A  L  P  D  Q  Y  L

4081 CTGCGCGGCAGGGTCGTCCCCGCCGAAGACGGCAGCATCGAGTTCCACAGCATCCGACCGGTGCCCTACGAAATCCCGAAAGCTGGCCCCACGGGCCAGCTCATGAACTCATACCTCGGC
      L  R  G  R  V  V  P  A  E  D  G  S  I  E  F  H  S  I  R  P  V  P  Y  E  I  P  K  A  G  P  T  G  Q  L  M  N  S  Y  L  G

4201 CGGCACTCCTGGCGCCCCGCACACATCCACATCCGCATCACCGCCGACGGGTACCGGCCACTGATCACCCAGCTCTACTTCGAGGGCGACCCGTACCTGGACAGCGACTCCTGCTCGGCA
      R  H  S  W  R  P  A  H  I  H  I  R  I  T  A  D  G  Y  R  P  L  I  T  Q  L  Y  F  E  G  D  P  Y  L  D  S  D  S  C  S  A

4321 GTCAAGAGCGAACTGGTTCTGCCGGTCAACAAGATCGACATCGACGGCGAAACATGGCAGCTAGTCGACTTCAACTTCATTCTCCAGCACAACTGATCGCGAGACATCATGTGCCACAAC
      V  K  S  E  L  V  L  P  V  N  K  I  D  I  D  G  E  T  W  Q  L  V  D  F  N  F  I  L  Q  H  N  STOP  clcD →  M  C  H  N

4441 AAATCATCCGCGCCTCCCACACCCGCACACATATCGATACAGCAAAACCGGACTCCCGGCACCGACCCCGTTCCCTACCTGCACGCGTCCTCGGACACCGGGAGCGGTGACACGATCGTC
      K  S  S  A  P  P  T  P  A  H  I  S  I  Q  Q  N  R  T  P  G  T  D  P  V  P  Y  L  H  A  S  S  D  T  G  S  G  D  T  I  V

4561 CTCCTCACCGACATCTTCGGTGTGACCCCGTTCTATCGACACCTGGCTGCCATGCTCGCCGAGAAAGGCCACGACGTCGTCATTCCCGACGTATTCCACCGAGTAGGCCACGCCACCGAC
      L  L  T  D  I  F  G  V  T  P  F  Y  R  H  L  A  A  M  L  A  E  K  G  H  D  V  V  I  P  D  V  F  H  R  V  G  H  A  T  D

4681 CCGGGCCGCGACGCGGCCTTGGCCAGACGCCGCCAATTGGACGACCGTCTCGCGATCGAGGACATCGAACGCACCGTCGCCCACACCGTCGATGATCAACAGACATTCGGTGTACTCGGA
      P  G  R  D  A  A  L  A  R  R  R  Q  L  D  D  R  L  A  I  E  D  I  E  R  T  V  A  H  T  V  D  D  Q  Q  T  F  G  V  L  G

4801 TTCTGCCTCGGCGGAAGCTTCGCCCTGCTGACCGCGGCCGCACACCCCAATCAGGTCACCGCCACCTACTACGCCTTCCCTAAAGGCGCGCCAGGAGCGAAAGTGCCTGTCAAACCGCCA
      F  C  L  G  G  S  F  A  L  L  T  A  A  A  H  P  N  Q  V  T  A  T  Y  Y  A  F  P  K  G  A  P  G  A  K  V  P  V  K  P  P

4921 CTTGAGGCTGCCGATGCGATCGACGGGCCCGTCCTGTGCCACTGGGGGCGAGACGACTACATCGATCACGAGGAAATTGACCAACTGGAACAGATACTCGCGTCGGCTCCCGGTCCCAGC
      L  E  A  A  D  A  I  D  G  P  V  L  C  H  W  G  R  D  D  Y  I  D  H  E  E  I  D  Q  L  E  Q  I  L  A  S  A  P  G  P  S

5041 GAGATCCGGTGGTACGACAACGCCGGGCACAGCTTCCTCGCCGGACTAACAGAACCGAACCACCCCAGCACTGCAGCCGCCCACGACTCCTGGCAGCGCACAGTGGAATTCTTCGAACGG
      E  I  R  W  Y  D  N  A  G  H  S  F  L  A  G  L  T  E  P  N  H  P  S  T  A  A  A  H  D  S  W  Q  R  T  V  E  F  F  E  R

5161 TACCTCACCGTAGGATCCGCCAACTGAACGGAACGGCGCCCTGCCGGCGACCACACAGGCCCATTTGCCCATACCATTCCTGGTATACGTTGCAGAGCAACGGACTGATCACTGCCTTGT
      Y  L  T  V  G  S  A  N  STOP

5281 GGCGTGCCCGTCACCGGGCGGCGCAGCTGCTGTCACCTCTCTCCGGCTTACCCGCTGAGTCTGCTCACAGCACCAACGCGCAACCTCGAGCGTCGGCTCCTCGTGCCATTCACGACCGGC
5401 CCCGGTATTCAGAACCCGCCTGCGGACCAGTCCGCCTGCGGTTGGGTGGTTGTTGTCGTCGGGTGGAGAACTCGCGAGAACTGGGCCCCACTCAGCGCCGGGGTCGGGAAACATCGTTGT

5521 CGGAGTCACGCGTTCGGCACTCGCCGGTCGACCAGGCGCACTCAGTAGGTGTTTGGAAAGCTGTGCGATTTCGGCCCCGAAGCGGATCGGTGACATCATGATGGGGCATGGGTTCTCGTA
                                                                                                                ORF5 →  M  G  S  R  K

5641 AGCAGCATCGCAAGCCGAAGCCGTATCCGTCCGATTTGACCGATGCGCAGTGGGAGTTGATCGCACCGATGGTCCCTGTATCTTCGGCTGGTCGCCCGCCCATTCACTCGCGCCGCCGGA
      Q  H  R  K  P  K  P  Y  P  S  D  L  T  D  A  Q  W  E  L  I  A  P  M  V  P  V  S  S  A  G  R  P  P  I  H  S  R  R  R  I

5761 TCGTGGACGCGATCTTGTATGTCAACCGCACCGGCTGTTCCTGGCGCCAGTTGCCGCATGATTTTCCGAAGTGGGAGACGGTGTACTGGTACTTCAAGGTGTGGAACGAGTCTGGGGTGA
      V  D  A  I  L  Y  V  N  R  T  G  C  S  W  R  Q  L  P  H  D  F  P  K  W  E  T  V  Y  W  Y  F  K  V  W  N  E  S  G  V  T

5881 CCGACCGGATCCACGACGCACTGCGGGCGCGGGCGCGCGATCGCGGTGGTCGGGATCCGATGGCCTCGGCGGGGATGGTCGACGCCCAGGCGGTGAAGGGTGCCGACACCGTCGGGACGG
      D  R  I  H  D  A  L  R  A  R  A  R  D  R  G  G  R  D  P  M  A  S  A  G  M  V  D  A  Q  A  V  K  G  A  D  T  V  G  T  D

6001 ATACGCGCGGTTGGGACGGCGCGAAACGAGTCAATGGGCGCAAAAGGCACTTGGTGACCGACACGCTCGGCCTACTCGTGGTGGTGTTGGTGACGGCGGGGAGTGTGCAGGACCGTGACG
      T  R  G  W  D  G  A  K  R  V  N  G  R  K  R  H  L  V  T  D  T  L  G  L  L  V  V  V  L  V  T  A  G  S  V  Q  D  R  D  G

6121 GCGGACGCCGGGTGCTGGCGCGGGCGAAGACGGCGATGCCGTCGATGGTCGTGGTGTGGGCCGACGGCGGCTACGCCGGGAAGCTGGTCGCCTGGGTCGACACGCTGCCGCATCGCAGTG
      G  R  R  V  L  A  R  A  K  T  A  M  P  S  M  V  V  V  W  A  D  G  G  Y  A  G  K  L  V  A  W  V  D  T  L  P  H  R  S  G

6241 GATATCGTTCGGAAACCCAAAGGGCAGAGTGGGTTTGCTGTCCTACCGCACCGCTGGGTGATCGAACGGACGCTGTCGTGGCTGGTGCGCTGTCGCCGCCTCGACCACGACTATGAGCGC
      Y  R  S  E  T  Q  R  A  E  W  V  C  C  P  T  A  P  L  G  D  R  T  D  A  V  V  A  G  A  L  S  P  P  R  P  R  L  STOP

6361 CTGCCCGCACACTCGGAGGCGTTCATCAAGTGGGCCATGATTGGGCTCATGACACGACGACTCGCACCGACACCGGGACGCCGACCATGGCAATCGACCCGCGCATCGTGACCATTTCCA
6481 AACACTTAGTCAGAGGCATCTAGCGGCGCTCCGGCATGTACTCGGGATACTCAGCCGGTGGTGCAATCTAGAAGACAGCTTGGTGAAGCCCTGGGGCCATCCGATATGGACCTGGTCTAC
6601 CTGTACTGCCACGGAGAGCGGTACAACCCCCCGGGAGCATCCGTGTCCGACCCCCTCCTGGTCCTCGGGTCCGGCGAATCGTTCACACCCCAGGACGTCAGCGGATGGTCCCAGCTCTAC
6721 GATTGGCCCGATCCGCACTGGCGCGACCGCCCGCCCCTCGTCGTGCTAAACGCCTGCCACTCCGGAGCGAGCCTGGCGACAACCCTCGCTGATTTCGTAGGAGTCTTCACGCAAACCGCC
6841 GGCGCCGCTGGGGTCATCGCCACCGAGGTCGCGCTTGAGCAGAAAGTGGCCGGACTCGCCATGGAAACCTTCCTGACGGAGCTCGTCATCCCCCGCATCTCCGTCGGCGAGGCCATGCGC
6961 CGCATGCGCTGGACGCTCCTCAAGCGTGGAAATCTCATGGGTCTCGCCTACTCGCCGTACTGCGCAGCAGACCTGCTGCTGTGGCCTGCAACGACTCTGCAACACAGAGCCGACGAAGCG
7081 CTAAGCCTGGATCCGTGGAGCTACGCCAGGCCGTAACGACGTGTGACTAACGAAAGCGCCTCCCGTAGTGAGACAATCAAGATTCTCACGTCACAATTGTCCACATACGGAAGGGCACTT
7201 TCAGTGCAATCATCGCACACGTTCACGGCTGGGTCCGCGG
```

FIG. 3—*Continued.*

erties (43, 44, 63). Especially interesting was the observation that with respect to the conversion of 2-chloro-*cis*,*cis*-muconate and the chloromuconolactones derived from it, the muconate, and the chloromuconate cycloisomerase of *R. opacus* 1CP were more similar to each other than to proteobacterial muconate or chloromuconate cycloisomerases characterized thus far (63). If these phenotypic similarities of the *Rhodococcus* enzymes resulted from a closer evolutionary relationship, then this would mean that the adaptation to chlorosubstituted substrates occurred independently among the proteobacteria and the gram-positive bacteria. Thus, on the basis of a sequence divergence, a functional convergence (12) would have to be assumed. Proof of this hypothesis would require sequence information from both the enzymes involved in catechol catabolism and those involved in chlorocatechol catabolism of *Rhodococcus*. We recently reported on the characterization of the catechol gene cluster (17), and in the accompanying report (18) we present sequences of enzymes for protocatechuate degradation, some of which also contribute to the catechol branch of the 3-oxoadipate pathway. In this communication, we show that the chlorocatechol enzymes of *R. opacus* 1CP do indeed represent different branches than the previously characterized enzymes encoded on plasmids pAC27, pJP4, and pP51.

(Some of the results reported here have recently been presented in a preliminary communication [16].)

TABLE 2. Sequences of N termini, tryptic peptides, and deduced primers

Protein or peptide	Amino acid sequence[a]	Deduced primer sequence
Chlorocatechol 1,2-dioxygenase		
CCD67.8	VVPAEDG(V)(I)(P)F(H)SIR	
CCD72.0	GNWT(D)(L)AIQGPF	AASGGBCCYTGKATNGC
CCD82.6	VIELFDEF(T)DL(R)	GTSATCGAGYTNTTYGAYGA
Chloromuconate cycloisomerase		
N terminus[b]	PNLTVSGVRTTIVDLPIL(R)P	ACSACSATCGTSGAYYTNCC
CMCI65.74	ELEAIGLSAVPVEEQR	
CMCI76.33	QIIDEYLAPALIGSDVEDLAGTR	
CMCI79.76	VEVVT(D)AGFVGLGEGV	
CMI85.33	FANHSIDAQTYLLVEVVTDAGFVGL	TCKATBSWGTGGTTBGCRAA
Dienelactone hydrolase		
DLH49.82	SXAPPTPAHISIQQNR	
DLH71.46	XYDNAGHSFLAGLTEPNHPSTAAAH	
DLH84.65	TPGTDPVPYLHASSDTG	

[a] Segments used for the design of oligonucleotides for cloning are underlined. Letters in parentheses indicate uncertain amino acids.
[b] From reference 63. The published N-terminal sequence had an aspartate in the second position instead of the asparagine shown here.

MATERIALS AND METHODS

Strains, plasmids, and cultivation conditions. *R. opacus* 1CP was isolated with 2,4-dichlorophenol as the carbon source (25, 66). To obtain biomass for the enzyme purifications, the strain was cultivated in a mineral medium with 4-chlorophenol as described previously (63). *Escherichia coli* DH5α (5), obtained from Gibco BRL, was grown aerobically at 37°C with constant shaking in LB (Luria-Bertani) medium (56) supplied with ampicillin (100 μg/ml), if appropriate. The plasmids used in this study are described in Table 1.

Enzyme assays and estimation of protein concentration. The activity of chloromuconate cycloisomerase was measured spectrophotometrically at 260 nm as described by Schlömann et al. (59), using 0.1 mM 2-chloro-*cis*,*cis*-muconate as the substrate (extinction coefficient, 17,100 $M^{-1} \cdot cm^{-1}$ [13]). The latter was available from a previous preparation (72). Dienelactone hydrolase was assayed at 280 nm, using a modification of the procedure of Schmidt and Knackmuss (60) in the presence of 50 mM Tris-HCl (pH 7.5) and 0.05 mM *cis*-dienelactone (*cis*-4-carboxymethylenebut-2-en-4-olide) (extinction coefficient, 17,000 $M^{-1} \cdot cm^{-1}$ [59]; obtained from S. R. Kaschabek and W. Reineke, Wuppertal, Germany).

Protein purifications. Chlorocatechol 1,2-dioxygenase was purified basically as described previously (43).

For the purification of chloromuconate cycloisomerase, a cell extract was prepared in 50 mM BisTris-HCl (bis-[2-hydroxyethyl]-imino-tris[hydroxymethyl-]methane-HCl; pH 6.5) plus 1 mM dithiothreitol (DTT) by grinding the cells with glass beads and clarifying the crude extract as described previously (17). The extract (volume, 250 ml; protein, 800 mg; total activity, 14.6 U; specific activity, 0.018 U/mg) was subjected to a purification scheme which differed from that of Solyanikova et al. (63) especially in the first step and in the additional reversed-phase chromatography. From a Pharmacia Q Sepharose High Performance HR 16/10 column (bed volume, 20 ml), used with 50 mM BisTris-HCl (pH 6.5), 1 mM DTT, and a linear NaCl gradient (0 to 0.5 M NaCl over 300 ml), the most active fractions eluted at ca. 0.35 M NaCl. $(NH_4)_2SO_4$ was added to 1.6 M (final concentration), and the preparation was chromatographed on a Pharmacia Phenyl Superose HR 10/10 column. Elution with a linear gradient of $(NH_4)_2SO_4$ (1.6 to 0 M over 104 ml) in 50 mM Tris-HCl (pH 7.5), 2 mM $MnSO_4$ yielded cycloisomerase-containing fractions at about 0.35 to 0.4 M $(NH_4)_2SO_4$. The preparation was subjected to gel filtration on a Pharmacia Superdex 200 prep-grade HiLoad 16/60 column (bed length, 60 cm), using 50 mM Tris-HCl (pH 7.5)–2 mM $MnSO_4$–0.2 M NaCl for elution. The resulting preparation (volume, 9 ml; protein, 12.3 mg; total activity, 8.25 U) had a specific activity of 0.67 U/mg, equivalent to a 37-fold purification. The purified enzyme was prepared for tryptic digestion by reversed-phase chromatography. A ca. 0.2-mg aliquot of the purified enzyme was loaded onto a Pharmacia ProRPC HR 5/10 column with 0.1% (vol/vol) trifluoroacetic acid in H_2O and eluted with a linear gradient (over 18 ml) to 0.1% (vol/vol) trifluoroacetic acid in acetonitrile. The absorption peak at ca. 65% (vol/vol) acetonitrile was collected and concentrated in a vacuum centrifuge.

For purification of dienelactone hydrolase, an extract (volume, 65 ml; protein, 124 mg; total activity, 51.9 U; specific activity, 0.42 U/mg) was prepared in 50 mM Tris-HCl (pH 7.2) with 0.1 mM DTT. Compared to a previous purification (44), more advanced chromatographic media were used. Chromatography with Q Sepharose High Performance HR16/10 and a gradient from 0 to 0.75 M NaCl over 400 ml was followed by Phenyl Sepharose Fast Flow HR16/10 chromatography using a gradient from 2 to 0 M $(NH_4)_2SO_4$ over 280 ml. The final Superdex 75 HR10/30 (bed length, 30 cm) step yielded a preparation (volume, 1.5 ml; protein, 0.28 mg; total activity, 8.3 U) with a specific activity of 29.7 U/mg, equivalent to a 71-fold purification.

The purity of enzyme preparations was investigated by discontinuous sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis as described previously (17). The gels were stained with a Pharmacia silver-staining kit.

Protein cleavage, isolation of peptides, and sequencing of peptides and N termini. Trypsin digestion of chlorocatechol dioxygenase, chloromuconate cycloisomerase, and dienelactone hydrolase as well as subsequent separation of tryptic peptides by reversed-phase high-pressure liquid chromatography were performed by the procedure of Stone et al. (64), with details as described by Eulberg et al. (17). The amounts of chlorocatechol dioxygenase, chloromuconate cycloisomerase, and dienelactone hydrolase after precipitation, washing, and drying were 66, 90, and 30 μg, respectively.

Purification and general in vitro manipulations of DNA. Genomic DNA of *R. opacus* 1CP was obtained as described previously (17). Plasmid DNA from *E. coli* DH5α was isolated with Pharmacia Flexiprep kits. For other general methods, see the accompanying report (18).

Amplification of segments of the chlorocatechol 1,2-dioxygenase gene and of the chloromuconate cycloisomerase gene. Oligonucleotides were custom synthesized according to the sequences of an N terminus and of several internal peptides (Table 2). To reduce the degeneracy of the oligonucleotides, a codon usage table for the previously sequenced genes *catA*, *catB*, and *catC* of strain 1CP (17) was used. The 50-μl PCR mixtures contained 50 pmol of each primer, ca. 0.25 μg of genomic template DNA, 0.1 mM each deoxynucleotide triphosphate, Goldstar DNA polymerase buffer [75 mM Tris-HCl (pH 9), 20 mM $(NH_4)_2SO_4$, 0.01% (wt/vol) Tween 20], 1.5 mM $MgCl_2$, and 0.5 U of Goldstar DNA polymerase (Eurogentech). A denaturing agent (dimethyl sulfoxide or formamide) (71) was added to the reaction mixtures to enhance the specificity of hybridization (2 to 8% [vol/vol] dimethyl sulfoxide or formamide for amplification of a fragment of the dioxygenase gene; 3 and 6% [vol/vol] dimethyl sulfoxide for the amplification of a cycloisomerase gene fragment). The thermocycling parameters for amplification of the chlorocatechol 1,2-dioxygenase gene fragment were as follows: 30 cycles of denaturing (95°C, 30 s), annealing (48°C, 2 min), and polymerization (72°C, 45 s), with an additional 2.5 min of denaturing before addition of the polymerase during the first cycle (10) and an additional 4.25 min of polymerization during the last cycle. For amplification of the chloromuconate cycloisomerase fragment, a thermocycling program without a separate polymerization step was used as appropriate for the small size of the product expected: 30 cycles of denaturing (95°C, 30 s) and annealing (48°C, 1 min), with an additional 2.5 min of denaturing before addition of the polymerase during the first cycle and 20 s of polymerization at 72°C after the last cycle.

Cloning strategy and hybridization procedures. PCR products were ligated into a T-tailed vector (45). The plasmids containing the ca. 220-bp product obtained with the chlorocatechol 1,2-dioxygenase-derived primers and the ca. 60-bp product obtained with the chloromuconate cycloisomerase-derived primers were designated pRERD1 and pRERC1, respectively. The insert of pRERD1 was digoxigenin labeled by using a DIG DNA Labeling and Detection Kit Nonradioactive (Boehringer), while the insert of pRERC1 was labeled, after denaturation (100°C, 10 min) and immediate chilling on ice, by tailing with digoxigenin-dUTP using a DIG Oligonucleotide Tailing Kit (Boehringer). On Southern blots of genomic DNA of strain 1CP digested with *Pst*I or *Eco*RI, we could identify a ca. 2.9-kb *Pst*I fragment by hybridization with the labeled insert of pRERD1 and a ca. 3.7-kb *Eco*RI fragment by hybridization with the labeled

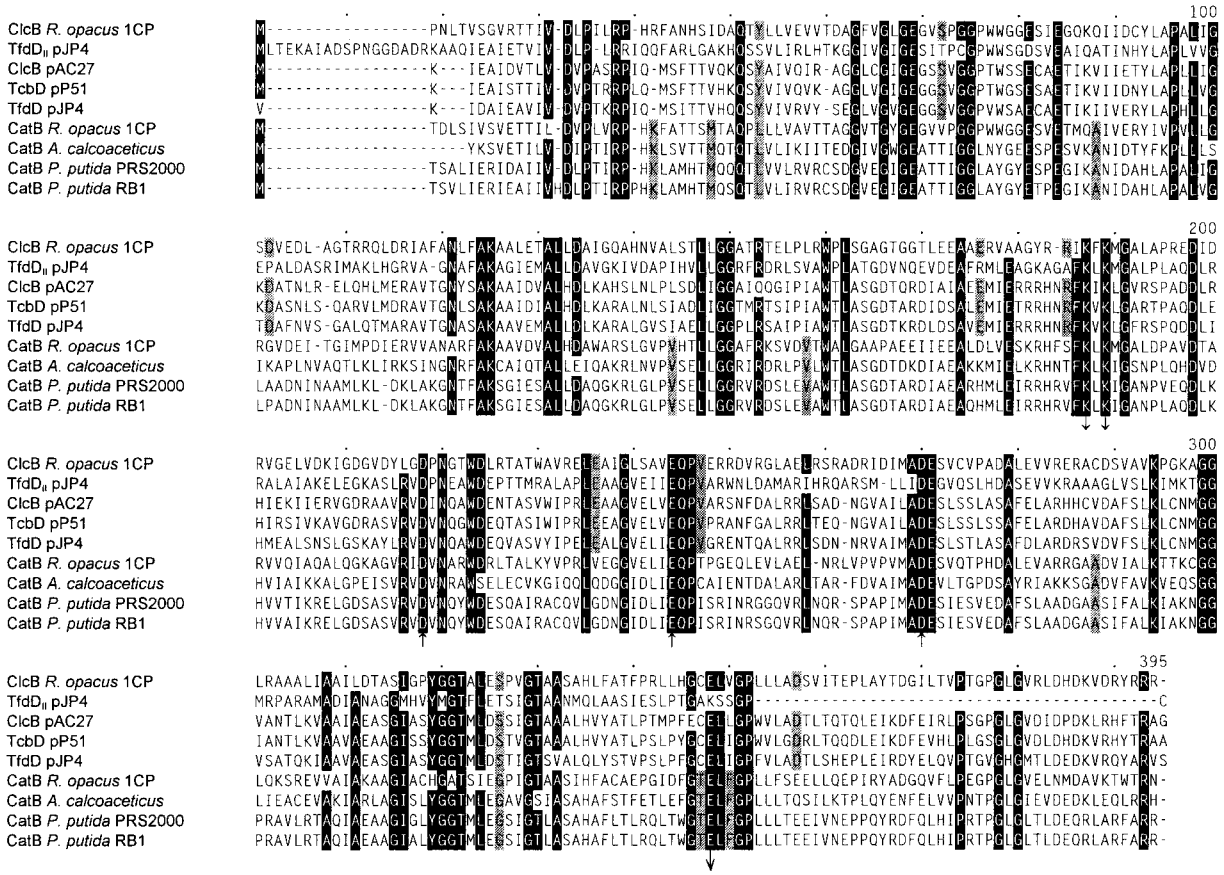

FIG. 4. Sequence alignment of muconate cycloisomerases (CatB), chloromuconate cycloisomerases (TfdD, ClcB, and TcbD), and a cycloisomerase of unknown biochemistry ($TfdD_{II}$) as calculated by the CLUSTAL program. Numbering above the sequences refers to positions in the alignment, not in single sequences. Amino acids in positions in which at least eight of the nine aligned sequences are identical are highlighted. Amino acids that are identical in all muconate but different in all chloromuconate cycloisomerases (or vice versa) are shaded. ($TfdD_{II}$ was not included because of the lack of biochemical data for this hypothetical protein.) Aspartate and glutamate residues involved in manganese coordination (arrows pointing upward), as well as lysine and glutamate residues directly involved in the enzyme mechanism (arrows pointing downward), are indicated (22, 30, 33). Accession numbers and references for the published sequences: $TfdD_{II}$, U16782 (67); ClcB of pAC27, M16964 (20); TcbD, M57629 (69); TfdD, M35097 (52); CatB *R. opacus* 1CP, X99622 (17); CatB *A.* (*Acinetobacter*) *calcoaceticus*, M76991 (62); CatB *P. putida* PRS2000, U12557 (35) with two corrections in positions 158 and 263 of the alignment (57); CatB *P. putida* RB1, M19460 (1). The N-terminal amino acid of TfdD has been determined to be methionine, although the associated codon is GTG (27).

insert of pRERC1 (procedure as described in the Boehringer manual). Identical digests were loaded onto second gels, and slices that corresponded in size to the positions of the respective hybridization signals were excised. The DNA was eluted from the gel slices and ligated into the dephosphorylated *Pst*I or *Eco*RI site of pBluescript II KS (+). The resulting gene banks were transformed into cells of *E. coli* DH5α. The clones which carried the DNA of interest were then identified with the corresponding probe (performed by a procedure analogous to that described in reference 17). The plasmid hybridizing to the labeled insert of pRERD1 was designated pRER5; the one which hybridized with pRERC1 was designated pRER7. An adjacent clone, pRER6, which completed the open reading frame for dienelactone hydrolase on pRER5, was obtained by digoxigenin labeling a ca. 300-bp *Not*I-*Pst*I fragment of the open reading frame and by using that to identify a ca. 10-kb *Hin*dIII fragment. A 2,409-bp *Sac*II subclone of pRER6 which overlaps with pRER5 was also identified by hybridization with the labeled ca. 300-bp *Not*I-*Pst*I fragment and was designated pRER61.

DNA sequence analysis. The nucleotide sequences of the overlapping inserts of pRER5, pRER7, and a part of pRER6 were determined partly by sequencing nested deletions or subclones and partly by sequencing with custom-synthesized primers. For the sequencing reactions, an Applied Biosystems Prizm kit was used, with subsequent electrophoresis and analysis in an Applied Biosystems A373 sequencer. Computer-based sequence analyses were done with the PC/GENE program package (Intelligenetics Inc., Mountain View, Calif.). Database searches were performed by the programs BLASTX (3, 24) and FASTA (51). Multiple sequence alignments were performed by the PC/GENE version of CLUSTAL (31), with open and unit gap costs set to 10. Dendrograms based on the Fitch-Margoliash algorithm were obtained by using the programs PROTDIST, FITCH, and DRAWTREE, and bootstrap data were obtained by additionally using the programs SEQBOOT and CONSENSE, all from the PHYLIP program package (19).

Nucleotide sequence accession number. The 7,240-bp sequence of the overlapping inserts of pRER7, pRER5, and pRER61 shown in Fig. 3 will appear in the EMBL, GenBank, and DDBJ nucleotide sequence databases under accession no. AF003948.

RESULTS

Cloning of the *clc* gene cluster from *R. opacus* 1CP. As a basis for the cloning of the *clc* gene cluster, the chlorocatechol 1,2-dioxygenase, the chloromuconate cycloisomerase, and the dienelactone hydrolase from strain 1CP, which previously had been purified and characterized biochemically (43, 44, 63), were purified again from 4-chlorophenol-grown *R. opacus* 1CP cells. Aliquots of the purified proteins were subjected to trypsin digestion, and selected isolated peptides were sequenced (Table 2). Some of these peptide sequences, as well as the previously sequenced N terminus of the chloromuconate cycloisomerase (63), served for the design of degenerate primers for PCR experiments (Table 2). PCR products were obtained independently for the chlorocatechol 1,2-dioxygenase gene (a ca. 220-bp fragment) and for the chloromuconate cycloisomer-

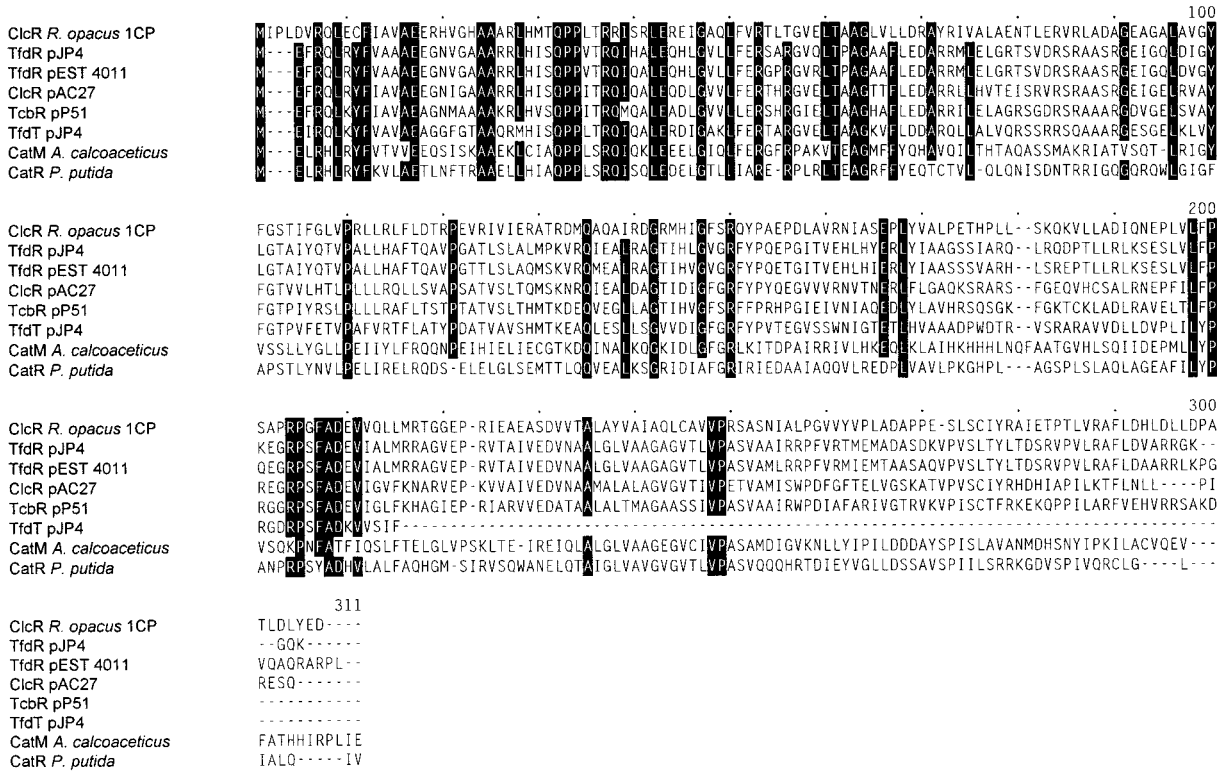

FIG. 5. Sequence alignment of the presumed ClcR of *R. opacus* to known LysR-type regulatory proteins that govern operons for catechol catabolism (CatM and CatR) or chlorocatechol catabolism (ClcR, TfdR, TcbR, and TfdT). Numbers above the sequences refer to positions in the alignment, not in single sequences. Amino acids in positions in which at least seven of the eight sequences are identical are highlighted. Accession numbers and references for the published sequences: TfdR of pJP4, M98445 (46); ClcR of pAC27, L06464 (8); TcbR, M57629 (70); TfdT, U16782 (40); CatM, M76991 (55); CatR, U12557 (35). The amino acid sequence of TfdR from pEST4011 was obtained by translating the DNA sequence after generating a frameshift by insertion of one nucleotide behind position 726 of the submitted sequence (U32188 [38]).

ase gene (a ca. 60-bp fragment). The plasmids that resulted from the ligation of the PCR products into pBluescript II KS (+), designated pRERD1 and pRERC1, respectively, were checked by sequencing to contain the fragment of interest. The inserts of pRERD1 and pRERC1 were labeled with digoxigenin and were used to identify a ca. 2.9-kb *Pst*I clone and a ca. 3.7-kb *Eco*RI clone (designated pRER5 and pRER7, respectively). By DNA sequence analysis, it was shown that the inserts of the two plasmids overlap, which means that the chlorocatechol 1,2-dioxygenase gene and the chloromuconate cycloisomerase gene are part of the same cluster of chlorocatechol catabolic genes. A major clone with an insert overlapping that of pRER5, pRER6, and its subclone, pRER61, were constructed to allow the complete sequencing of this gene cluster.

Sequence analysis of pRER5, pRER7, and pRER61. The sequenced 7,240 bp of the overlapping inserts of pRER5, pRER7, and pRER61 comprised five complete open reading frames (Fig. 2 and 3). The first open reading frame (position 1,700 to 579) is transcribed in the opposite direction of the other open reading frames. It was identified as *clcB*, the gene of the chloromuconate cycloisomerase, by showing that the sequence of its predicted gene product contains all determined peptide sequences as well as the reported N-terminal sequence (63) in which one position turned out to be wrong (Table 2; Fig. 3). In addition, the predicted subunit molecular mass of the protein encoded by this open reading frame (39,532 Da) fits well with the 40 to 42 kDa determined by SDS-polyacrylamide gel electrophoresis of purified chloromuconate cycloisomerase (63). The product of the *clcB* gene shows high sequence similarities to known bacterial muconate and chloromuconate cycloisomerases: 41.8% identical positions with *R. opacus* CatB (in the alignment of Fig. 4), 36.2 to 38.3% with muconate cycloisomerases of γ-proteobacteria, and 36.5 to 40.3% identical positions with chloromuconate cycloisomerases of proteobacterial plasmids (excluding $TfdD_{II}$, for which no biochemical data have been reported [33.9%]). The residues which are involved in manganese coordination, as well as those which are directly involved in the enzyme mechanism (22, 30, 33), are conserved throughout the known muconate and chloromuconate cycloisomerases, including ClcB from *Rhodococcus* (but excepting $TfdD_{II}$) (Fig. 4).

Divergently transcribed from *clcB*, a second open reading frame, coding for a protein with a molecular mass of 33,261 Da, was found to extend from positions 2509 to 3420. It is similar to transcriptional activators of the LysR type, especially to those regulating chlorocatechol catabolic gene clusters on proteobacterial plasmids (36.8 to 39.0% identical positions in the alignment of Fig. 5) and to γ-proteobacterial regulators of catechol catabolism (29.8 and 26.7% identical positions, respectively, with CatR and CatM). Thus, this open reading frame was considered to represent the *clcR* gene, coding for a regulatory protein.

The product predicted from the third open reading frame, positions 3643 to 4416, contains all of tryptic peptides that were sequenced of chlorocatechol 1,2-dioxygenase (Table 2; Fig. 3). The gene was thus designated *clcA*, although the deduced N-terminal amino acid sequence differed in 10 of 26 positions from the previously published, obviously incorrect

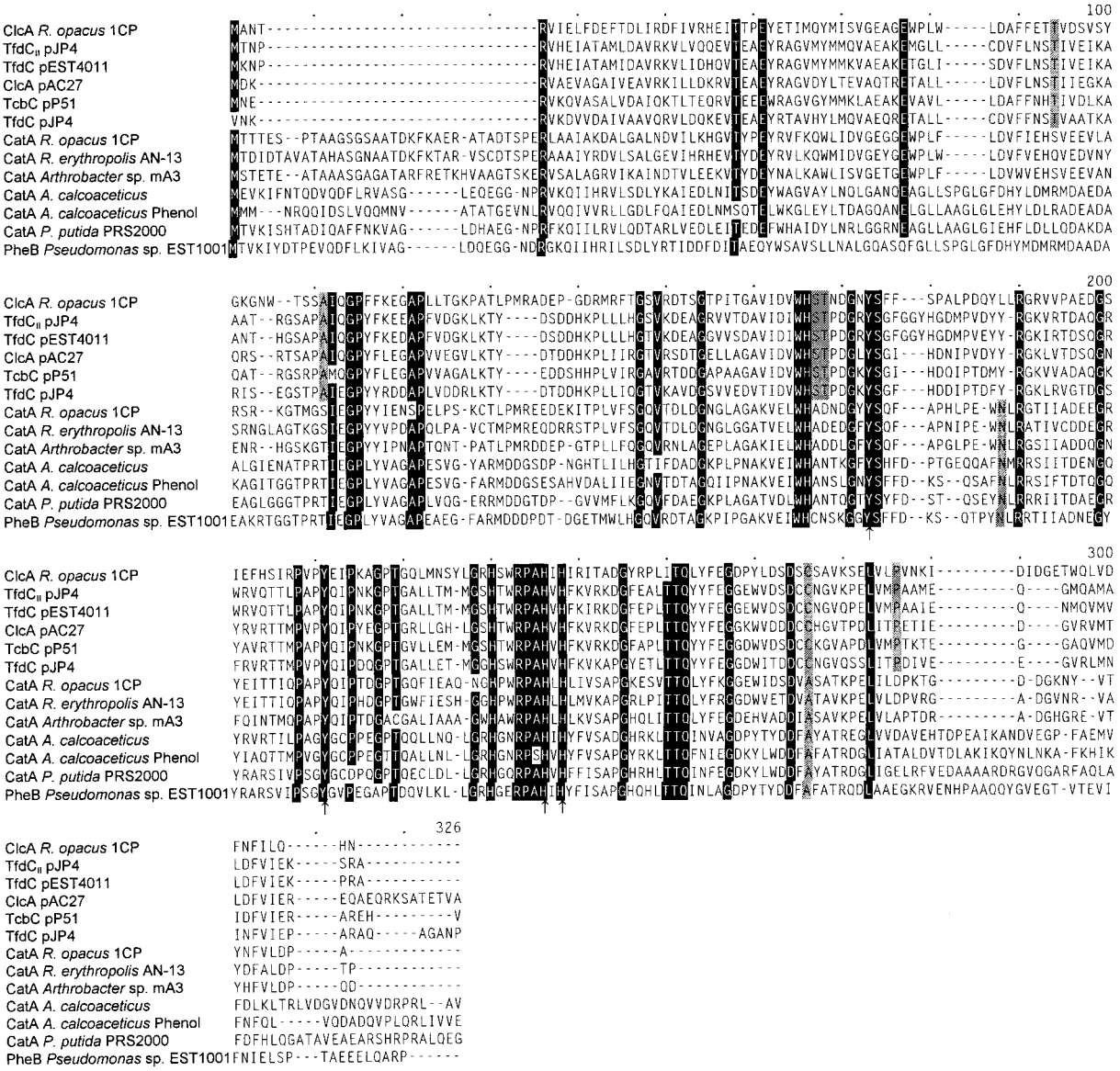

FIG. 6. Sequence alignment of catechol 1,2-dioxygenases (CatA and PheB) and chlorocatechol 1,2-dioxygenases (ClcA, TcbC, TfdC, and TfdC$_{II}$) as calculated by the CLUSTAL program. The numbering above the sequences refers to positions in the alignment, not in single sequences. Amino acids in which at least 12 of the 13 sequences are identical are highlighted. Amino acids that are identical in all catechol but different in all chlorocatechol 1,2-dioxygenases (or vice versa) are shaded. Histidine and tyrosine residues that were previously reported to be involved in iron binding (28, 41) are indicated by arrows. Accession numbers and references for the published sequences: TfdC$_{II}$, U16782 (67); TfdC of pEST4011, U32188 (38); ClcA of pAC27, M16964 (20); TcbC, M57629 (69); TfdC of pJP4, M35097 (52); CatA *R. opacus* 1CP, X99622 (17); CatA *R. erythropolis* AN-13, D83237 (47); CatA *Arthrobacter* sp. mA3, M94318 (14); CatA *A. calcoaceticus* ADP1, M76991 (49); CatA for phenol degradation by *A. calcoaceticus* NCIB8250, Z36909 (15); CatA *P. putida* PRS2000, U12557 (35); PheB *Pseudomonas* sp. EST1001, M57500 (37). Two sequences of catechol 1,2-dioxygenases from *Pseudomonas* strains (48) which are very similar to the sequence of CatA from *P. putida* PRS2000 have been omitted for clarity.

N-terminal sequence of the purified protein (43). ClcA shows high sequence similarities to bacterial catechol and chlorocatechol 1,2-dioxygenases: 38.9 to 40.5% identical positions with *Rhodococcus* or *Arthrobacter* CatA (in the alignment of Fig. 6), 29.2 to 33.3% with catechol 1,2-dioxygenases of γ-proteobacteria, and 34.1 to 39.5% identical positions with chlorocatechol dioxygenases of proteobacterial plasmids. The amino acids that have been reported to be involved in iron binding (28, 41) are also conserved in the sequence predicted for chlorocatechol 1,2-dioxygenase of strain 1CP. The calculated molecular mass of ClcA is 28,952 Da, which is consistent with calculated molecular masses of other chlorocatechol dioxygenases (27.6 to 29.0 kDa). The rhodococcal ClcA thus does not seem to be significantly shorter than its counterparts from gram-negative strains as suggested by Maltseva et al. (43), who reported to have determined a molecular mass of 26.5 kDa for the purified enzyme by SDS-polyacrylamide gel electrophoresis.

The fourth open reading frame, between positions 4429 and 5187, represents *clcD*, the gene of the dienelactone hydrolase. All tryptic peptides obtained from the purified enzyme occur within the predicted sequence (Table 2; Fig. 3). The calculated molecular mass of the *clcD* gene product (27,304 Da) is somewhat lower than the one (30 kDa) determined for the purified protein (44). With the dienelactone hydrolases TfdE of pJP4,

FIG. 7. Sequence alignment of dienelactone hydrolases as calculated by the CLUSTAL program. Numbers above the sequences refer to positions in the alignment, not in individual sequences. Amino acids in those positions in which at least five of the six sequences are identical are highlighted. Presumed catalytic triad residues (50) are indicated by arrows. Accession numbers and references for the published sequences: TfdE$_{II}$, U16782 (67); TfdE of pEST4011, U32188 (38); ClcD of pAC27, M16964 (20); TcbE, M57629 (69); TfdE of pJP4, M35097 (52).

TcbE of pP51, and ClcD of pAC27, ClcD from *Rhodococcus* shares between 17.9 and 22.2% identical positions in the alignment of Fig. 7. This is considerably less than the respective scores obtained for comparisons of chlorocatechol 1,2-dioxygenases or chloromuconate cycloisomerases from the same sources. The protein predicted from *clcD* was found to be most similar to the dienelactone hydrolases TfdE of pEST4011 and TfdE$_{II}$ of module 2 on pJP4, with which it shares 30.0 and 31.1% identical positions, respectively. Thus, while the rhodococcal dienelactone hydrolase in the dendrogram of Fig. 8 resides on the same branch as two proteobacterial enzymes, TfdE of pEST4011 and TfdE$_{II}$ of module 2 on pJP4, the sequence similarity to them is still significantly less than the similarity between the rhodococcal chloromuconate cycloisomerase or chlorocatechol 1,2-dioxygenase and their proteobacterial counterparts. The catalytic triad residues identified for ClcD of pAC27 (50) are also contained in ClcD from *Rhodococcus* (Fig. 7).

A fifth open reading frame (ORF5; bases 5628 to 6356) was detected 441 bp downstream of *clcD*. Its predicted gene product shows significant sequence similarity to putative transposition-related proteins, especially to those of Tn*4811* from *Streptomyces lividans* 66 (7) (42.2% identical positions) and of IS*1031*C from *Acetobacter xylinum* (9) (38.0% identical positions). Comparison of the sequences of the inverted repeats of IS*1031*C to the sequence in Fig. 3 indicated that the left inverted repeat is most similar (10 out of 21 bases identical) to the segment directly following the stop codon of ORF5 (positions 6357 to 6377). This corresponds to the arrangement in IS*1031*C. The right inverted repeat is most similar (8 out of 21 bases identical) to a segment of the intergenic region between *clcB* and *clcR* (positions 2002 to 2022). Comparison of the *R. opacus* 1CP sequences of the two regions to each other revealed that they form an imperfect inverted repeat as indicated in Fig. 3.

A hairpin structure could be detected between *catR* and *catA*, suggesting that the presumed *clcR* gene might be transcribed independently from the *clcAD* genes (Fig. 3).

DISCUSSION

The chlorocatechol degradative pathways which are represented by the *clcR,ABDE* gene cluster of plasmid pAC27 (8, 20), by the *tfdT,CDEF* cluster (module 1) of pJP4 (40, 52), and by the *tcbR,CDEF* cluster of pP51 (69, 70) diverged from a common ancestral pathway already adapted for chlorocatechol catabolism (58). This conclusion could be drawn basically from two sorts of evidence. (i) The proteins encoded by the genes mentioned above are more similar to each other than to their counterparts in catechol degradation or to enzymes of other pathways. Correspondingly, they belong to the same branches in all four dendrograms depicted in Fig. 8. (ii) The three gene clusters referred to above show a marked similarity (Fig. 9). All three consist of a LysR-type regulatory gene divergently transcribed from an operon comprising, in that order, genes for a chlorocatechol 1,2-dioxygenase, a chloromuconate cycloisomerase, a dienelactone hydrolase, and a maleylacetate reductase, with an additional ORF between the cycloisomerase and the hydrolase genes of pAC27 and pP51. Since such a similar genetic structure is unlikely to have evolved two or more times independently, it should have occurred prior to the divergence of the lines giving rise to the gene clusters now present on pAC27, pJP4, and pP51.

In concordance with our hypothesis derived from biochemical evidence (63), the sequences of all four rhodococcal proteins involved in chlorocatechol degradation differed considerably from those of their protobacterial counterparts, but still proved to be homologous to the respective proteins of catechol and other chlorocatechol pathways (Fig. 4 to 8). In fact, in the dendrograms which illustrate the relatedness of the respective gene products, the *Rhodococcus* proteins represent branches separate from those of the corresponding proteins from proteobacteria (Fig. 8). In accordance with the sequence differences, the structure of the chlorocatechol gene cluster of strain 1CP also differs from those of pAC27, pP51, and module 1 of pJP4 (Fig. 9). The cycloisomerase gene *clcB* is transcribed divergently from *clcR,AD*, and a maleylacetate reductase gene could not be detected in this gene cluster and was cloned on a separate fragment (61). Thus, the organization of the genes does not provide any indication that the *Rhodococcus* pathway may have evolved from the same origin as the chlorocatechol pathways of the plasmids pAC27, pP51, and pJP4 (module 1).

The deep branching of the enzymes and regulatory proteins of chlorocatechol pathways in Fig. 8 challenges the general opinion that catechol catabolism evolved first and then gave

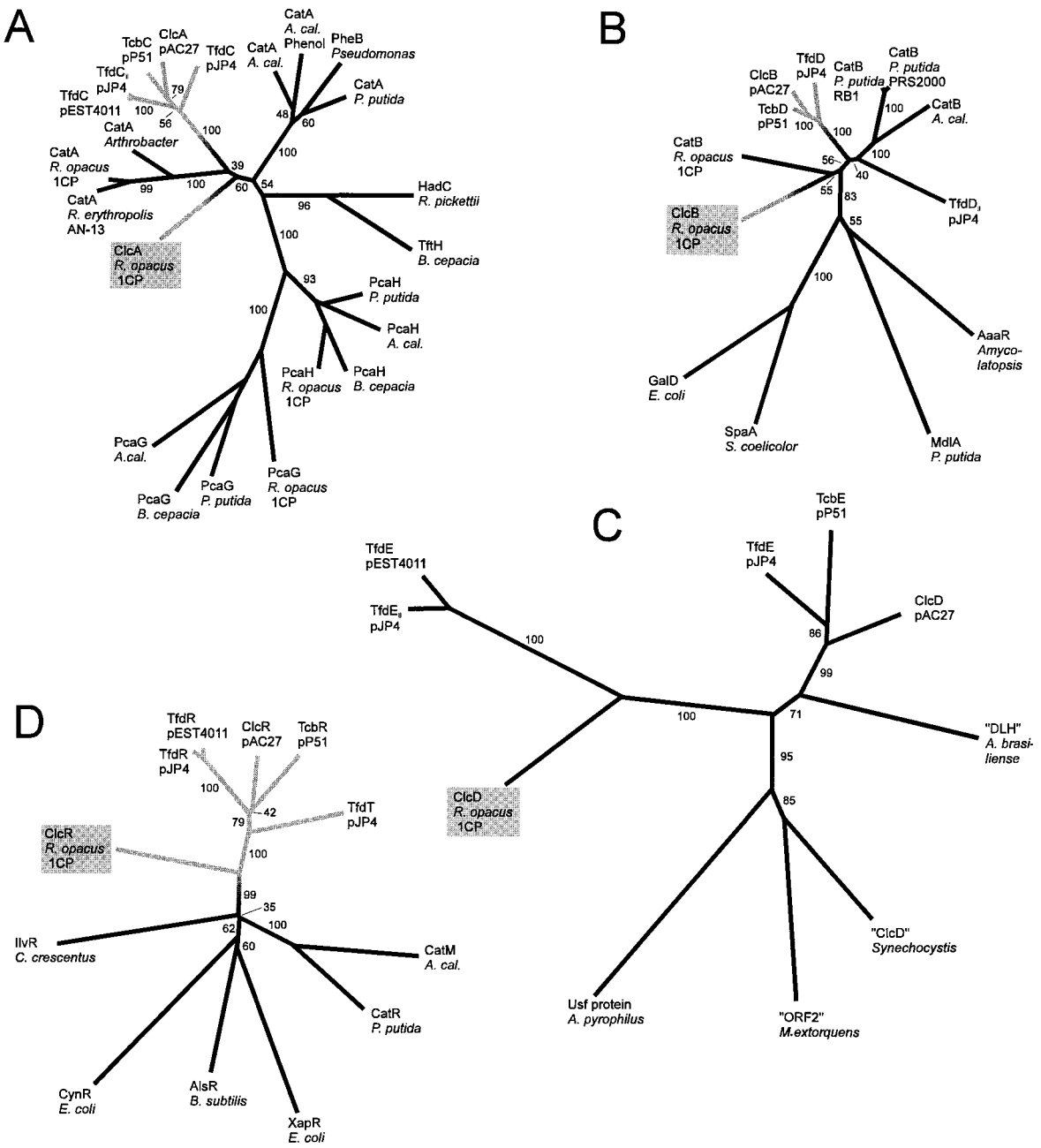

FIG. 8. Dendrograms illustrating sequence similarities between proteins involved in catechol and chlorocatechol catabolism. (A) Catechol and chlorocatechol 1,2-dioxygenases; (B) muconate and chloromuconate cycloisomerases; (C) dienelactone hydrolases; (D) LysR-type regulatory proteins. The dendrograms were calculated by the PHYLIP program package (19) based on CLUSTAL alignments of the sequences in Fig. 4 to 7 and the respective outgroup sequences. For the sequences used in Fig. 4 to 7, the resulting alignments were very similar but, close to some gaps, not quite identical to those shown. Numbers on the branches indicate the frequency (percentage) with which the corresponding cluster occurred during bootstrap analyses. They show that the affiliation of a certain protein to one of the major groups could be deduced with great confidence. The exact branching pattern of these major groups, in contrast, as well as the branching within the groups is not certain, as indicated by low bootstrap numbers. The dendrograms for cycloisomerases and dienelactone hydrolases were calculated from the complete alignments (with the exception that the amino acids in positions 3 to 38 of "ORF2" of *Methylobacterium extorquens* were omitted). For the dioxygenases, only positions 140 to 258 of the alignment in Fig. 6, and for the regulators, only positions 1 to 217 of the alignment in Fig. 5, were considered. All dendrograms are drawn to the same scale with respect to evolutionary distance (in centimeters per unit branch length). Gray branches represent enzymes which are specific for chlorocatechol catabolism; the position of the transition from black to gray was chosen arbitrarily. Functions and accessions numbers of the proteins used as outgroups: PcaG and PcaH, α and β subunits of protocatechuate 3,4-dioxygenases, of *A. calcoaceticus* (*A. cal.*) (M33798), *Burkholderia cepacia* (M30791), *P. putida* (L14836) and *R. opacus* (AF003947) TftH, hydroxyquinol 1,2-dioxygenase, of *B. cepacia* (U19883); HadC, hydroxyquinol 1,2-dioxygenase, of *Ralstonia pickettii* (D86544); GalD, galactonate dehydratase, of *E. coli* (L10328 and U19577); SpaA, protein possibly involved in the metabolism of signalling lactones, of *Streptomyces coelicolor* (X94190); MdlA, mandelate racemase, of *P. putida* (J05293); AaaR, *N*-acyl amino acid racemase, of *Amycolatopsis* sp. (D30738); IlvR, regulator of isoleucine and valine synthesis, of *Caulobacter crescentus* (L24392); CynR, regulator for cyanate detoxification, of *E. coli* (M93053); AlsR, acetoin synthesis regulator, of *Bacillus subtilis* (L04470); XapR, xanthosine catabolic regulator, of *E. coli* (X73828); Usf protein, hypothetical protein, of *Aquifex pyrophilus* (U17575); "ORF2," hypothetical protein, of *M. extorquens* (U72662); "ClcD," hypothetical protein homologous to dienelactone hydrolases, of *Synechocystis* sp. PCC6803 (D90904); "DLH," hypothetical protein homologous to dienelactone hydrolases, of *Azospirillum brasiliense* (X67216).

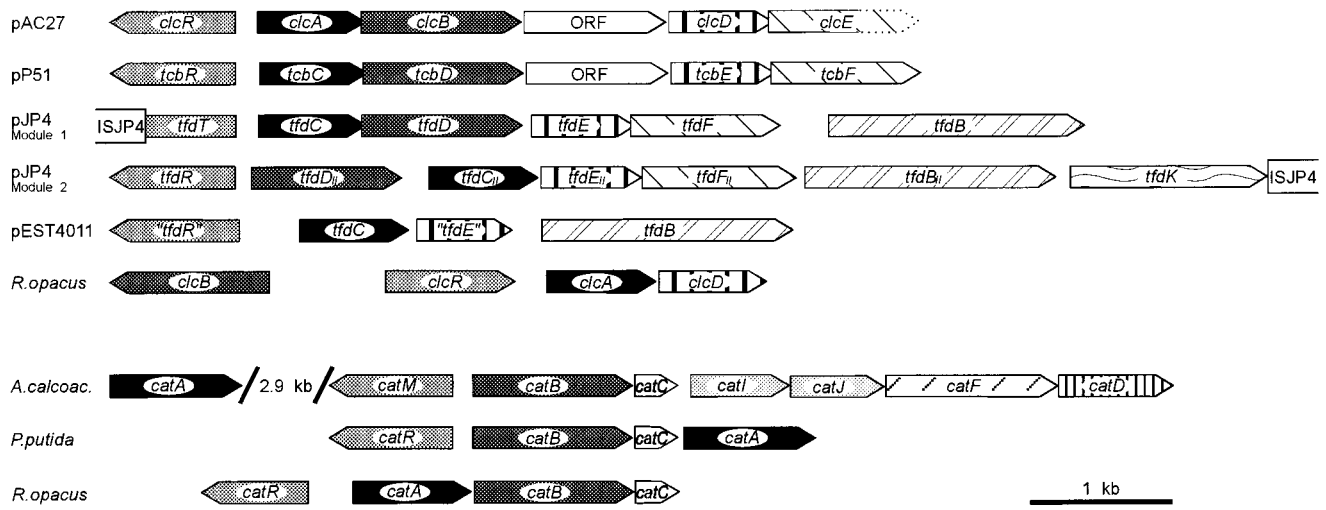

FIG. 9. Structures of gene clusters for catechol and chlorocatechol catabolism. Homologous genes are shaded in the same way. Functions of the proteins encoded by the genes are evident from Fig. 1. For *tfdR* of pEST4011, see the legend to Fig. 5. The information for this figure was compiled from references given in a previous review (58) as well as from references 17, 35, 38, 40, 55, and 67.

rise to chlorocatechol catabolism. However, the two different known branches of catechol degradative pathways, from gram-positive bacteria on the one hand and from γ-proteobacteria on the other, have one feature in common: *catC* directly follows *catB* (Fig. 9). This feature is also conserved in *R. eutropha*, a β-proteobacterium (32). Since the *catC* product, muconolactone isomerase, is an enzyme characteristic of catechol catabolism (Fig. 1), this conserved structure suggests that the catechol pathway has indeed evolved first. This conclusion is consistent with aromatic compounds, which give rise to catechol during their degradation, being more common in nature than chloroaromatic compounds that are degraded via chlorocatechols. Thus, according to present knowledge, chlorocatechol degradation did in fact evolve from catechol degradation. And obviously, it did so independently in gram-positive bacteria and proteobacteria, that is, by functionally convergent evolution.

The sequences of proteobacterial dienelactone hydrolases are known to be so dissimilar to those of their counterparts in catechol pathways, the 3-oxoadipate enol-lactone hydrolases, that recruitment of a dienelactone hydrolase from a preexisting pathway other than for catechol or protocatechuate turnover had to be postulated (58). Given that the *Rhodococcus* chlorocatechol pathway evolved independently from that of pAC27, pP51, and pJP4 (module 1), it was considered possible that *Rhodococcus* found a completely different solution for lactone hydrolysis. It was therefore interesting to find that ClcD of strain 1CP is, in fact, related to other dienelactone hydrolases and not, for example, to 3-oxoadipate enol-lactone hydrolases. However, in contrast to the chlorocatechol 1,2-dioxygenase and chloromuconate cycloisomerase of *R. opacus* 1CP, its dienelactone hydrolase resides on the same branch of the dendrogram as two putative isofunctional enzymes of proteobacterial origin (Fig. 8C). These putative dienelactone hydrolases are encoded by plasmid pEST4011 and by a second chlorocatechol module on pJP4, respectively (38, 67 [U16782]). The respective catabolic genes have only very recently been characterized, and no biochemical evidence has yet been reported for the function of either of the putative dienelactone hydrolases. But even if one accepts that these gene products are dienelactone-hydrolyzing enzymes, their relatively close evolutionary relationship to the dienelactone hydrolase of *R. opacus* 1CP does not call into question the conclusion that the chlorocatechol pathway of this strain evolved independently from that of proteobacterial plasmids. Considering that the absolute evolutionary distance between *Rhodococcus* ClcD and TfdE of pEST4011 or $TfdE_{II}$ of pJP4 is even greater than that between the major groups of dioxygenases and cycloisomerases, a plausible interpretation of the similarity pattern of the dienelactone hydrolases would be that similar enzymes and their genes were independently recruited into the gene clusters of pEST4011 and module 2 of pJP4 on the one hand and *R. opacus* 1CP on the other.

It is interesting that while the putative dienelactone hydrolases of pEST4011 and module 2 of pJP4 are relatively similar to the *Rhodococcus* enzyme, the respective chlorocatechol 1,2-dioxygenases are more closely related to the other proteobacterial dioxygenases (Fig. 8A). The latter is also true for the LysR-type regulator TfdR of pJP4 (46, 74) and for a presumed analogous regulator which may be deduced from the reported sequence of pEST4011 when allowing for one frameshift (Fig. 8D). As argued above, the ancestral chlorocatechol gene cluster from which those of pAC27, pP51, and module 1 of pJP4 diverged must be assumed to have included all necessary genes, among them those coding for chloromuconate cycloisomerase and dienelactone hydrolase. Since the dioxygenases of pEST4011 and of module 2 of pJP4 diverged from that same origin (Fig. 8A), they should, at some time, have been associated with cycloisomerases and hydrolases closely related to those of pAC27, pP51, and module 1 of pJP4. By some sort of gene shuffling, the hydrolase genes have apparently been replaced by those more closely related to the corresponding *Rhodococcus* gene. A cycloisomerase gene has so far not been found in the gene cluster of pEST4011. The presumed chloromuconate cycloisomerase of module 2 of pJP4, $TfdD_{II}$, might even represent a new branch of cycloisomerizing enzymes (Fig. 8B). This, however, must be considered speculation because of the absence of biochemical data on $TfdD_{II}$.

Not only with respect to the evolution of whole pathways but also for the constituting groups of enzymes, the sequence information from *R. opacus* 1CP bears potentially significant new information. Comparison of the dioxygenase sequences revealed that even though they represent different evolutionary groups, the chlorocatechol 1,2-dioxygenases of *R. opacus* and

of proteobacterial plasmids are considerably shorter at the N termini than are catechol 1,2-dioxygenases. It would be interesting to see whether this feature influences the catalytic properties. Especially informative should be those positions in the dioxygenases and cycloisomerases in which the chlorocatechol enzymes of *Rhodococcus* and proteobacteria carry the same amino acids but different ones than the respective catechol enzymes (Fig. 4 and 6). These positions either may have been conserved or may have been gained independently because of their importance for the turnover of chlorinated substrates. Based on structural data of muconate and chloromuconate cycloisomerases (30, 33), we have previously identified some of the amino acids relevant for substrate specificity by using site-directed mutagenesis (73). The occurrence of functionally convergent evolution should allow us to obtain new insight into the adaptations which are necessary at the molecular level to yield an enzyme with altered catalytic properties.

ACKNOWLEDGMENTS

We are indebted to H.-J. Knackmuss for valuable discussions and for making it possible for us to do this work. We thank H. Weber and R. Getzlaff for protein and peptide sequencing, R. Schmid for the opportunity to use an automated DNA sequencer, and S. Bürger as well as S. Lakner for performing the DNA sequencing. We are grateful to W. Reineke and S. R. Kaschabek for a gift of *cis*-dienelactone.

The work was funded by a fellowship of the Landesgraduiertenförderung to D.E. and by grants from the Deutsche Forschungsgemeinschaft.

REFERENCES

1. **Aldrich, T. L., and A. M. Chakrabarty.** 1988. Transcriptional regulation, nucleotide sequence, and localization of the promoter of the *catBC* operon in *Pseudomonas putida*. J. Bacteriol. **170:**1297–1304.
2. **Alting-Mees, M. A., J. A. Sorge, and J. M. Short.** 1992. pBluescriptII: multifunctional cloning and mapping vectors. Methods Enzymol. **216:**483–495.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Amy, P. S., J. W. Schulke, L. M. Frazier, and R. J. Seidler.** 1985. Characterization of aquatic bacteria and cloning of genes specifying partial degradation of 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. **49:** 1237–1245.
5. **Bethesda Research Laboratories.** 1986. BRL pUC host: *E. coli* DH5α competent cells. Bethesda Res. Lab. Focus **8(2):**9.
6. **Chatterjee, D. K., and A. M. Chakrabarty.** 1982. Genetic rearrangements in plasmids specifying total degradation of chlorinated benzoic acids. Mol. Gen. Genet. **188:**279–285.
7. **Chen, C. W., T.-W. Yu, H.-M. Chung, and C.-F. Chou.** 1992. Discovery and characterization of a new transposable element, Tn*4811*, in *Streptomyces lividans* 66. J. Bacteriol. **174:**7762–7769.
8. **Coco, W. M., R. K. Rothmel, S. Henikoff, and A. M. Chakrabarty.** 1993. Nucleotide sequence and initial functional characterization of the *clcR* gene encoding a LysR family activator of the *clcABD* chlorocatechol operon in *Pseudomonas putida*. J. Bacteriol. **175:**417–427.
9. **Coucheron, D. H.** 1993. A family of IS*1031* elements in the genome of *Acetobacter xylinum*: nucleotide sequences and strain distribution. Mol. Microbiol. **9:**211–218.
10. **D'Aquila, R. T., L. J. Bechtel, J. A. Videler, J. J. Eron, P. Gorczyca, and J. C. Kaplan.** 1991. Maximizing sensitivity and specificity of PCR by preamplification heating. Nucleic Acids Res. **19:**3749.
11. **Don, R. H., A. J. Weightman, H.-J. Knackmuss, and K. N. Timmis.** 1985. Transposon mutagenesis and cloning analysis of the pathways for degradation of 2,4-dichlorophenoxyacetic acid and 3-chlorobenzoate in *Alcaligenes eutrophus* JMP134(pJP4). J. Bacteriol. **161:**85–90.
12. **Doolittle, R. F.** 1994. Convergent evolution: the need to be explicit. Trends Biochem. Sci. **19:**15–18.
13. **Dorn, E., and H.-J. Knackmuss.** 1978. Chemical structure and biodegradability of halogenated aromatic compounds. Substituent effects on 1,2-dioxygenation of catechol. Biochem. J. **174:**85–94.
14. **Eck, R., and J. Belter.** 1993. Cloning and characterization of a gene coding for the catechol 1,2-dioxygenase of *Arthrobacter* sp. mA3. Gene **123:**87–92.
15. **Ehrt, S., F. Schirmer, and W. Hillen.** 1995. Genetic organization, nucleotide sequence and regulation of expression of genes encoding phenol hydroxylase and catechol 1,2-dioxygenase in *Acinetobacter calcoaceticus* NCIB8250. Mol. Microbiol. **18:**13–20.
16. **Eulberg, D., and M. Schlömann.** 1997. Convergent evolution of chlorocatechol catabolism in *Rhodococcus* and proteobacteria. Biospektrum (special issue), abstr. VU016, p. 35.
17. **Eulberg, D., L. A. Golovleva, and M. Schlömann.** 1997. Characterization of catechol catabolic genes from *Rhodococcus erythropolis* 1CP. J. Bacteriol. **179:**370–381.
18. **Eulberg, D., S. Lakner, L. A. Golovleva, and M. Schlömann.** 1998. Characterization of a protocatechuate catabolic gene cluster from *Rhodococcus opacus* 1CP: evidence for a merged enzyme with 4-carboxymuconolactone-decarboxylating and 3-oxoadipate enol-lactone-hydrolyzing activity. J. Bacteriol. **180:**1072–1081.
19. **Felsenstein, J.** 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, Wash.
20. **Frantz, B., and A. M. Chakrabarty.** 1987. Organization and nucleotide sequence determination of a gene cluster involved in 3-chlorocatechol degradation. Proc. Natl. Acad. Sci. USA **84:**4460–4464.
21. **Fulthorpe, R. R., C. McGowan, O. V. Maltseva, W. E. Holben, and J. M. Tiedje.** 1995. 2,4-Dichlorophenoxyacetic acid-degrading bacteria contain mosaics of catabolic genes. Appl. Environ. Microbiol. **61:**3274–3281.
22. **Gerlt, J. A., and P. G. Gassman.** 1992. Understanding enzyme-catalyzed proton abstraction from carbon acids: details of stepwise mechanisms for β-elimination reactions. J. Am. Chem. Soc. **114:**5928–5934.
23. **Ghisalba, O.** 1983. Chemical wastes and their biodegradation—an overview. Experientia **39:**1247–1257.
24. **Gish, W., and D. J. States.** 1993. Identification of protein coding regions by database similarity search. Nat. Genet. **3:**266–272.
25. **Gorlatov, S. N., O. V. Maltseva, V. I. Shevchenko, and L. A. Golovleva.** 1989. Degradation of chlorophenols by a culture of *Rhodococcus erythropolis*. Mikrobiologiya **58:**802–806. (Microbiology **58:**647–651.)
26. **Häggblom, M. M.** 1992. Microbial breakdown of halogenated aromatic pesticides and related compounds. FEMS Microbiol. Rev. **103:**29–72.
27. **Hammer, A., T. Hildenbrand, H. Hoier, K.-L. Ngai, M. Schlömann, and J. J. Stezowski.** 1993. Crystallization and preliminary X-ray data of chloromuconate cycloisomerase from *Alcaligenes eutrophus* JMP134 (pJP4). J. Mol. Biol. **232:**305–307.
28. **Hartnett, C., E. L. Neidle, K.-L. Ngai, and L. N. Ornston.** 1990. DNA sequences of genes encoding *Acinetobacter calcoaceticus* protocatechuate 3,4-dioxygenase: evidence indicating shuffling of genes and of DNA sequences within genes during their evolutionary divergence. J. Bacteriol. **172:**956–966.
29. **Hartnett, G. B., and L. N. Ornston.** 1994. Acquisition of apparent DNA slippage structures during extensive evolutionary divergence of *pcaD* and *catD* genes encoding identical catalytic activities in *Acinetobacter calcoaceticus*. Gene **142:**23–29.
30. **Helin, S., P. C. Kahn, B. L. Guha, D. G. Mallows, and A. Goldman.** 1995. The refined X-ray structure of muconate lactonizing enzyme from *Pseudomonas putida* PRS2000 at 1.85 Å resolution. J. Mol. Biol. **254:**918–941.
31. **Higgins, D. G., and P. M. Sharp.** 1989. Fast and sensitive multiple sequence alignments on a microcomputer. CABIOS **5:**151–153.
32. **Hinner, I.-S., M. Lohmann, and M. Schlömann.** 1997. Two catechol catabolic genes in *Ralstonia eutropha* 335. Biospektrum (special issue), abstr. PU074, p. 71.
33. **Hoier, H., M. Schlömann, A. Hammer, J. P. Glusker, H. L. Carrell, A. Goldman, J. J. Stezowski, and U. Heinemann.** 1994. Crystal structure of chloromuconate cycloisomerase from *Alcaligenes eutrophus* JMP134 (pJP4) at 3 Å resolution. Acta Crystallogr. D**50:**75–84.
34. **Holben, W. E., B. M. Schroeter, V. G. M. Calabrese, R. H. Olsen, J. K. Kukor, V. O. Biederbeck, A. E. Smith, and J. M. Tiedje.** 1992. Gene probe analysis of soil microbial populations selected by amendment with 2,4-dichlorophenoxyacetic acid. Appl. Environ. Microbiol. **58:**3941–3948.
35. **Houghton, J. E., T. M. Brown, A. J. Appel, E. J. Hughes, and L. N. Ornston.** 1995. Discontinuities in the evolution of *Pseudomonas putida cat* genes. J. Bacteriol. **177:**401–412.
36. **Ka, J. O., W. E. Holben, and J. M. Tiedje.** 1994. Genetic and phenotypic diversity of 2,4-dichlorophenoxyacetic acid (2,4-D)-degrading bacteria isolated from 2,4-D-treated field soils. Appl. Environ. Microbiol. **60:**1106–1115.
37. **Kivisaar, M., L. Kasak, and A. Nurk.** 1991. Sequence of the plasmid-encoded catechol 1,2-dioxygenase-expressing gene, *pheB*, of phenol-degrading *Pseudomonas* sp. strain EST1001. Gene **98:**15–20.
38. **Kõiv, V., R. Marits, and A. Heinaru.** 1996. Sequence analysis of the 2,4-dichlorophenol hydroxylase gene *tfdB* and 3,5-dichlorocatechol 1,2-dioxygenase gene *tfdC* of 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011. Gene **174:**293–297.
39. **Kuhm, A. E., M. Schlömann, H.-J. Knackmuss, and D. H. Pieper.** 1990. Purification and characterization of dichloromuconate cycloisomerase from *Alcaligenes eutrophus* JMP 134. Biochem. J. **266:**877–883.
40. **Leveau, J. H. J., and J. R. van der Meer.** 1996. The *tfdR* gene product can successfully take over the role of the insertion element-inactivated TfdT protein as a transcriptional activator of the *tfdCDEF* gene cluster, which encodes chlorocatechol degradation in *Ralstonia eutropha* JMP134(pJP4). J. Bacteriol. **178:**6824–6832.
41. **Lipscomb, J. D., and A. M. Orville.** 1992. Mechanistic aspects of dihydroxy-

benzoate dioxygenases, p. 243–298. *In* H. Sigel and A. Sigel (ed.), Metal ions in biological systems, vol. 28. Degradation of environmental pollutants by microorganisms and their metalloenzymes. Marcel Dekker, Inc., New York, N.Y.

42. **Mäe, A. A., R. O. Marits, N. R. Ausmees, V. M. Kõiv, and A. L. Heinaru.** 1993. Characterization of a new 2,4-dichlorophenoxyacetic acid degrading plasmid pEST4011: physical map and localization of catabolic genes. J. Gen. Microbiol. **139:**3165–3170.
43. **Maltseva, O. V., I. P. Solyanikova, and L. A. Golovleva.** 1994. Chlorocatechol 1,2-dioxygenase from *Rhodococcus erythropolis* 1CP. Kinetic and immunochemical comparison with analogous enzymes from Gram-negative strains. Eur. J. Biochem. **226:**1053–1061.
44. **Maltseva, O. V., I. P. Solyanikova, L. A. Golovleva, M. Schlömann, and H.-J. Knackmuss.** 1994. Dienelactone hydrolase from *Rhodococcus erythropolis* 1CP: purification and properties. Arch. Microbiol. **162:**368–374.
45. **Marchuk, D., M. Drumm, A. Saulino, and F. S. Collins.** 1991. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. Nucleic Acids Res. **19:**1154.
46. **Matrubutham, U., and A. R. Harker.** 1994. Analysis of duplicated gene sequences associated with *tfdR* and *tfdS* in *Alcaligenes eutrophus* JMP134. J. Bacteriol. **176:**2348–2353.
47. **Murakami, S., N. Kodama, R. Shinke, and K. Aoki.** 1997. Classification of catechol 1,2-dioxygenase family: sequence analysis of a gene for the catechol 1,2-dioxygenase showing high specificity for methylcatechols from Gram$^+$ aniline-assimilating *Rhodococcus erythropolis* AN-13. Gene **185:**49–54.
48. **Nakai, C., H. Uyeyama, H. Kagamiyama, T. Nakazawa, S. Inouye, F. Kishi, A. Nakazawa, and M. Nozaki.** 1995. Cloning, DNA sequencing, and amino acid sequencing of catechol 1,2-dioxygenases (pyrocatechase) from *Pseudomonas putida* mt-2 and *Pseudomonas arvilla* C-1. Arch. Biochem. Biophys. **321:**353–362.
49. **Neidle, E. L., C. Hartnett, S. Bonitz, and L. N. Ornston.** 1988. DNA sequence of the *Acinetobacter calcoaceticus* catechol 1,2-dioxygenase I structural gene *catA*: evidence for evolutionary divergence of intradiol dioxygenases by acquisition of DNA sequence repetitions. J. Bacteriol. **170:**4874–4880.
50. **Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag, J. L. Sussman, K. H. G. Verschueren, and A. Goldman.** 1992. The α/β hydrolase fold. Protein Eng. **5:**197–211.
51. **Pearson, W. R., and D. J. Lipman.** 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA **85:**2444–2448.
52. **Perkins, E. J., M. P. Gordon, O. Caceres, and P. F. Lurquin.** 1990. Organization and sequence analysis of the 2,4-dichlorophenol hydroxylase and dichlorocatechol oxidative operons of plasmid pJP4. J. Bacteriol. **172:**2351–2359.
53. **Pieper, D. H., W. Reineke, K.-H. Engesser, and H.-J. Knackmuss.** 1988. Metabolism of 2,4-dichlorophenoxyacetic acid, 4-chloro-2-methylphenoxyacetic acid and 2-methylphenoxyacetic acid by *Alcaligenes eutrophus* JMP 134. Arch. Microbiol. **150:**95–102.
54. **Reineke, W.** 1994. Degradation of chlorinated aromatic compounds by bacteria: strain development, p. 416–454. *In* G. R. Chaudhry (ed.), Biological degradation and bioremediation of toxic chemicals. Dioscorides Press, Portland, Oreg.
55. **Romero-Arroyo, C. E., M. A. Schell, G. L. Gaines III, and E. L. Neidle.** 1995. *catM* encodes a LysR-type transcriptional activator regulating catechol degradation in *Acinetobacter calcoaceticus*. J. Bacteriol. **177:**5891–5898.
56. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
57. **Schell, U.** Unpublished results.
58. **Schlömann, M.** 1994. Evolution of chlorocatechol catabolic pathways. Conclusions to be drawn from comparisons of lactone hydrolases. Biodegradation **5:**301–321.
59. **Schlömann, M., E. Schmidt, and H.-J. Knackmuss.** 1990. Different types of dienelactone hydrolase in 4-fluorobenzoate-utilizing bacteria. J. Bacteriol. **172:**5112–5118.
60. **Schmidt, E., and H.-J. Knackmuss.** 1980. Chemical structure and biodegradability of halogenated aromatic compounds. Conversion of chlorinated muconic acids into maleoylacetic acid. Biochem. J. **192:**339–347.
61. **Seibert, V., L. A. Golovleva, and M. Schlömann.** 1995. Cloning of maleylacetate reductase genes from *Alcaligenes eutrophus* JMP289 and *Rhodococcus erythropolis* 1CP. Biospektrum (special issue), abstr. PA013, p. 57.
62. **Shanley, M. S., A. Harrison, R. E. Parales, G. Kowalchuk, D. J. Mitchell, and L. N. Ornston.** 1994. Unusual G+C content and codon usage in *catIJF*, a segment of the *ben-cat* supra-operonic cluster in the *Acinetobacter calcoaceticus* chromosome. Gene **138:**59–65.
63. **Solyanikova, I. P., O. V. Maltseva, M. D. Vollmer, L. A. Golovleva, and M. Schlömann.** 1995. Characterization of muconate and chloromuconate cycloisomerase from *Rhodococcus erythropolis* 1CP: indications for functionally convergent evolution among bacterial cycloisomerases. J. Bacteriol. **177:**2821–2826.
64. **Stone, K. L., M. B. LoPresti, J. M. Crawford, R. DeAngelis, and K. R. Williams.** 1989. Enzymatic digestion of proteins and HPLC peptide isolation, p. 31–47. *In* P. T. Matsudaira (ed.), A practical guide to protein and peptide purification for microsequencing. Academic Press, Inc., San Diego, Calif.
65. **Top, E. M., W. E. Holben, and L. J. Forney.** 1995. Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. Appl. Environ. Microbiol. **61:**1691–1698.
66. **Tsitko, T. V., G. M. Zaitsev, A. G. Lobanok, and M. S. Salkinoja-Salonen.** Tolerance of *Rhodococcus opacus* strains to aromatic and aliphatic solvents and the role of cellular fatty acids in the adaptation. Submitted for publication.
67. **van der Meer, J. R.** 1996. Direct submission to the databases.
68. **van der Meer, J. R., A. R. W. van Neerven, E. J. de Vries, W. M. de Vos, and A. J. B. Zehnder.** 1991. Cloning and characterization of plasmid-encoded genes for the degradation of 1,2-dichloro-, 1,4-dichloro-, and 1,2,4-trichlorobenzene of *Pseudomonas* sp. strain P51. J. Bacteriol. **173:**6–15.
69. **van der Meer, J. R., R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos.** 1991. Sequence analysis of the *Pseudomonas* sp. strain P51 *tcb* gene cluster, which encodes metabolism of chlorinated catechols: evidence for specialization of catechol 1,2-dioxygenases for chlorinated substrates. J. Bacteriol. **173:**2425–2434.
70. **van der Meer, J. R., A. C. J. Frijters, J. H. J. Leveau, R. I. L. Eggen, A. J. B. Zehnder, and W. M. de Vos.** 1991. Characterization of the *Pseudomonas* sp. strain P51 gene *tcbR*, a LysR-type transcriptional activator of the *tcbCDEF* chlorocatechol oxidative operon, and analysis of the regulatory region. J. Bacteriol. **173:**3700–3708.
71. **Varadaraj, K., and D. M. Skinner.** 1994. Denaturants or cosolvents improve the specificity of PCR amplification of a G + C-rich DNA using genetically engineered DNA polymerases. Gene **140:**1–5.
72. **Vollmer, M. D., P. Fischer, H.-J. Knackmuss, and M. Schlömann.** 1994. Inability of muconate cycloisomerases to cause dehalogenation during conversion of 2-chloro-*cis*,*cis*-muconate. J. Bacteriol. **176:**4366–4375.
73. **Vollmer, M. D., H. Hoier, U. Schell, J. Gröning, J. Pleiss, A. Goldman, and M. Schlömann.** 1995. Differences in the catalytic properties of muconate and chloromuconate cycloisomerases analyzed by site-directed mutagenesis. Biospektrum (special issue), abstr. KV002, p. 24.
74. **You, I.-S., and D. Ghosal.** 1995. Genetic and molecular analysis of a regulatory region of the herbicide 2,4-dichlorophenoxyacetate catabolic plasmid pJP4. Mol. Microbiol. **16:**321–331.