Cloning, DNA Sequence Analysis, and Expression in *Escherichia coli* of the Gene for Mandelate Racemase from *Pseudomonas putida*†

Stephen C. Ransom and John A. Gerlt*
Department of Chemistry and Biochemistry, University of Maryland, College Park, Maryland 20742
Vincent M. Powers and George L. Kenyon*
Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143
Received July 22, 1987; Revised Manuscript Received September 21, 1987

ABSTRACT: The gene for mandelate racemase (EC 5.1.2.2) from *Pseudomonas putida* (ATCC 12633) was cloned in *Pseudomonas aeruginosa* (ATCC 15692). The selection for the cloned gene was based upon the inability of *P. aeruginosa* to grow on (*R*)-mandelate as sole carbon source by virtue of the absence of mandelate racemase in its mandelate pathway. Fragments of *P. putida* DNA obtained by digestion of chromosomal DNA with *Sau*3A were ligated into the *Bam*HI site of the Gram-negative vector pKT230 and transformed into the *P. aeruginosa* host. A transformant able to utilize (*R*)-mandelate as sole carbon source was characterized, and the plasmid was found to contain approximately five kilobase pairs of *P. putida* DNA. Subcloning of this DNA revealed the position of the gene for the racemase within the cloned DNA from *P. putida*. The dideoxy-DNA sequencing procedure was used to determine the sequence of the gene and its translated sequence. The amino acid sequence and molecular weight for mandelate racemase deduced from the gene sequence (38 570) are in excellent agreement with amino acid composition and molecular weight data for the polypeptide recently determined with enzyme isolated from *P. putida*; these recent determinations of the polypeptide molecular weight differ significantly from the originally reported value of 69 500 [Fee, Judith A., Hegeman, G. D., & Kenyon, G. L. (1974) *Biochemistry 13*, 2528], which was used to demonstrate that α-phenylglycidate, an active site directed irreversible inhibitor, binds to the enzyme with a stoichiometry of 1:1. The 5′-flanking sequence is homologous to the consensus sequences for *Escherichia coli* promoters as well as portions of promoter sequences reported for some *Pseudomonas* genes, suggesting that the gene for mandelate racemase is the first in the mandelate operon. Placement of the gene behind the *trc* promoter permitted expression of soluble and catalytically active mandelate racemase in *E. coli*, with the levels of expression being approximately 1% of the total cellular protein in induced cells; homogeneous, catalytically active enzyme has been isolated from extracts of these induced cells. Comparison of the mandelate racemase encoding DNA sequence and the deduced amino acid sequence with available DNA and amino acid sequence data bases revealed no significant sequence homologies.

The mandelate pathway in *Pseudomonas putida* is composed of five enzymes that convert the enantiomeric pair of mandelic acids to benzoic acid; catabolism of benzoic acid is accomplished by its oxidative ring opening and degradation to succinate and acetyl coenzyme A (acetyl-CoA) by the enzymes of the β-ketoadipate pathway (Gunsalus et al., 1953). The enzymes of the mandelate pathway include mandelate racemase (Kenyon & Hegeman, 1979), which catalyzes the interconversion of the enantiomers of mandelate, (*S*)-mandelate dehydrogenase, which catalyzes the oxidation of (*S*)-mandelate to form benzoyl formate, benzoylformate decarboxylase, which utilizes thiamine pyrophosphate to catalyze the formation of benzaldehyde, and two benzaldehyde dehydrogenases (one NAD^+ specific and one $NADP^+$ specific), which catalyze the formation of benzoic acid.

The genes encoding these enzymes constitute an operon that is induced by either enantiomer of mandelic acid (Hegeman, 1966) and can be cotransduced (Wheelis & Stanier, 1970). Despite the fact that the mandelate operon was one of the earliest described operons, no further work on its structure has been reported. Because we are interested in mechanistic characterization of enzymes in the mandelate pathway that depends upon the availability of cloned and sequenced genes, we have begun work to characterize the operon on a molecular level by cloning and DNA sequence analysis. We expect this project to allow both enhanced expression of the various enzymes in suitable bacterial species (*Escherichia coli* or pseudomonad species) and generation of mutant enzymes that will allow further description of the mechanisms of the enzymatic reactions.

In this paper we report the cloning and DNA sequence analysis of the gene for mandelate racemase as well as its expression in *E. coli*. The gene encodes a protein with 359 amino acids having a M_r of 38 570. This value differs significantly from the figure of 69 500 daltons reported earlier (Fee et al., 1974a). Since previous studies (Fee et al., 1974b; Whitman et al., 1985) have detailed the irreversible inactivation of mandelate racemase by α-phenylglycidate and the stoichiometry of the alkylation was concluded to be one inhibitor molecule per racemase polypeptide on the basis of the previously reported molecular weight, our revision of the molecular weight necessitates further investigation of the interaction of α-phenylglycidate with the enzyme. Homogeneous mandelate racemase has been isolated from *E. coli* transformed with an expression plasmid in which the gene has been placed downstream of the *trc* promoter (Amann & Brosius, 1985). Finally, the sequence data we have obtained for the 5′-flanking region indicate the presence of a promoter sequence imme-

†This research was supported by Grants GM-34572 to J.A.G. and AR-17323 to G.L.K. from the National Institutes of Health. This is paper 8 in a series on mandelate racemase. Paper 7 is Whitman et al. (1985). V.M.P. is the 1987 George V. Doerr Memorial Fellow of the American Foundation for Pharmaceutical Education.

0006-2960/88/0427-0540$01.50/0

diately upstream of the gene and, therefore, suggest that the gene for the racemase is the first in the operon.

Materials and Methods

The broad host range expression vector pKT230 (Bagdasarian et al., 1981), which is capable of replication within both *P. putida* and *E. coli*, was obtained from Dr. L. Nicholas Ornston, Yale University. This vector and recombinant plasmids derived from it were propagated in *Pseudomonas aeruginosa* (ATCC 15692). pKT230 confers kanamycin and streptomycin resistance to this host; the vector contains a unique *Bam*HI site that can be used for cloning that does not inactivate either resistance marker. The *P. aeruginosa* host was rendered temporarily restriction deficient by overnight growth at 42 °C (L. N. Ornston, private communication) presumably by thermal inactivation of the restriction system.

Bacteriophages M13mp18 and M13mp19 and plasmid pUC19 were obtained from Bethesda Research Laboratories. The bacteriophage were propagated in *E. coli* strain JM101, and the pUC plasmid was propagated in *E. coli* strain TB1. An expression vector, pKK233-2, containing the *trc* promoter followed by the *lac*Z ribosome binding site and unique *Nco*I restriction site for cloning, was obtained from Pharmacia Biochemicals and propagated in *E. coli* strain JM105 (Amann & Brosius, 1985). Expression vectors containing the bacteriophage λ P_L promoter followed by the λ cII gene ribosome binding site and unique restriction sites for cloning (*Nco*I in pOTS-Nco and *Nde*I in pMG27) were obtained from Dr. Martin Rosenberg, Smith Kline & Beckman Laboratories, Swedeland, PA, and propagated in *E. coli* strain M5219; this host strain contains a lytic deficient λ phage lysogenized in the chromosome that directs the synthesis of the cI857 temperature-sensitive λ repressor (Shatzman & Rosenberg, 1986).

Restriction endonucleases were obtained from New England Biolabs. The Klenow fragment of DNA polymerase I from *E. coli*, bacteriophage T_4 DNA ligase, bacteriophage T_4 polynucleotide kinase, and calf intestinal alkaline phosphatase were purchased from Boehringer Mannheim Biochemicals. All enzymes were used according to the supplier's directions.

Chromosomal DNA from *P. putida* (ATCC 12633) was isolated according to the following procedure (L. N. Ornston and K.-L. Ngai, private communication). Cells grown in 50 mL of LB were harvested and resuspended in 10 mL of 50 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl), pH 8.0, containing 20 mM ethylenediaminetetraacetic acid (EDTA). Pronase (1 mL of a 5 mg/mL solution) and sodium dodecyl sulfate (SDS) (0.1 mL of a 10% solution) were added, and the suspension was incubated for 6 h at 37 °C. The lysate was extracted 3 times with 10 mL of buffer-saturated phenol followed by three extractions with ether. The aqueous layer was dialyzed against 10 mM Tris-HCl, pH 7.5, containing 1 mM EDTA and 100 mM NaCl for 24 h with several changes of buffer followed by dialysis against 10 mM Tris-HCl, pH 7.5, containing 1 mM EDTA for 24 h with several changes of buffer. The clear, viscous solution was used directly for restriction digestions.

DNA sequence analyses were performed according to the dideoxy chain-terminating procedure and with [α-^{35}S]dATP as the radioactive nucleotide; regions of severe band compression were resolved with deoxy-7-deazaguanosine 5′-triphosphate in place of dGTP in the DNA polymerase reaction (Mizusawa et al., 1986). The commercially available "−40" M13 sequencing primer (New England Biolabs) was used for most sequencing reactions; in one case an oligonucleotide primer incorporating a sequence within the mandelate racemase gene that was determined on the complementary strand was prepared with a Systec 1450A DNA synthesizer according to the standard phosphoramidite chemistry program. Both strands of DNA comprising the gene were sequenced in their entirety.

Construction of a Library of Chromosomal DNA from P. putida in pKT230. Chromosomal DNA from *P. putida* was partially digested with *Sau*3A, and the fragments were size fractionated by agarose gel electrophoresis. Fragments in the molecular weight range between 2 and 5 kb were isolated by electroelution. The fragments were ligated with pKT230 that had been digested with *Bam*HI and dephosphorylated with calf intestinal alkaline phosphatase. The ligation mixture was used to transform *P. aeruginosa* (Bagdasarian et al., 1983) that had been rendered temporarily restriction deficient, and the transformed cells were plated on LB agar containing streptomycin at a concentration of 1 mg/mL.

Selection for Recombinant Plasmids Containing the Gene for Mandelate Racemase. The library of transformants constructed as described in the previous section was screened for the presence of functional mandelate racemase by replicate plating of the transformants growing on LB agar containing streptomycin to minimal medium plates containing both 2 g/L (*R*)-mandelate as carbon source and streptomycin. The basis for this selection is the ability of *P. aeruginosa* to utilize (*S*)- but not (*R*)-mandelate as sole carbon source due to the absence of mandelate racemase in its mandelate pathway.

Expression of the Gene for Mandelate Racemase in E. coli. The DNA sequence analysis of the cloned gene for mandelate racemase revealed the presence of a unique *Sca*I restriction site (AGTACT, with the product containing blunt ends) in codons three through five. The availability of potential expression vectors with initiation codons contained in unique *Nco*I (CCATGG, *trc* promoter in pKK233-2 and λ P_L promoter in pOTS-Nco) and *Nde*I (CATATG, λ P_L promoter in pMG27) restriction sites suggested a strategy for the convenient placement of the mandelate racemase gene behind these promoters. Palindromic oligonucleotide linkers were synthesized for ligation to the gene that had been restricted with *Sca*I such that either a unique *Nco*I or *Nde*I site could be introduced. In the case of the *Nco*I vectors, the 14-mer ACTTCCATGGAAGT was synthesized, with the initiation codon shown underlined; since the sequences of these vectors require that the second codon have a G in the first base, the linker was synthesized to delete the second amino acid (serine, codon AGT) found in the gene. In the case of the *Nde*I vector, the 22-mer ACTTCACTCATATGAGTGAAGT was synthesized, with the initiation codon again shown underlined; this linker contains all of the codons between the initiation codon and the *Sca*I restriction site. Accordingly, either a *Sca*I–*Hin*dIII restriction fragment (in the case of pKK233-3) or a *Sca*I–*Sac*I restriction fragment (in the case of the λ P_L vectors) containing the remainder of the gene was ligated to the appropriate linker, restricted with either *Nco*I or *Nde*I, and finally ligated separately with the appropriate fragment of each expression vector that had been restricted with either *Nco*I or *Nde*I and *Hin*dIII or *Sac*I (pKK233-3 has a unique *Hin*dIII site and the λ P_L vectors have unique *Sac*I sites downstream of the cloning site).

The plasmid containing the gene cloned in pKK233-3 (designated pMRtrc) is inducible by isopropyl β-thiogalactoside (IPTG). *E. coli* strain JM105 transformed with pMRtrc is grown to midlog phase (OD_{650nm} = 0.8) and induced by the addition of IPTG to a final concentration of 0.2 mM.

The genes cloned in pOTS-Nco and pMG27 are inducible by heat denaturation of the temperature-sensitive λ repressor

encoded by the chromosome of the host strain M5219. Cells grown at 30 °C to late log phase in LB medium were induced by the addition of hot medium to raise the temperature to 42 °C.

Purification of Mandelate Racemase from P. putida. Mandelate racemase was purified from *P. putida* (ATCC 12633) according to modifications of the procedure published by Hegeman (1970). Following lysis of cells in 50 mM Tris buffer, pH 8.0, containing 10 mM $MgCl_2$ and 0.1 mM phenylmethanesulfonyl fluoride (PMSF) and removal of nucleic acids by precipitation with $MnCl_2$, the supernatant obtained by adding solid ammonium sulfate to a saturation of 30% was adjusted to 40% saturation, stirred for 90 min, and centrifuged. The pellet was resuspended in and dialyzed against the Tris buffer. The protein was further fractionated by anion-exchange chromatography on a column of DEAE-Sephacel (2.5 × 31 cm) with a gradient of NaCl (0–0.5 M) in 1 L of the Tris buffer to elute the enzyme. The fractions containing mandelate racemase activity were concentrated by ultrafiltration (Amicon YM-30 membrane), and the proteins were size-fractionated by passage through a column of Sephacryl-200 (2.5 × 90 cm) in the Tris buffer. The DEAE-Sephacel and Sephacryl-200 steps were then repeated. The final fractionation was by fast protein liquid chromatography (FPLC) at room temperature on a column of Mono Q with a linear gradient of ammonium bicarbonate, pH 8.5 (0.39–0.53 M), over 20 min at a flow rate of 1 mL/min as eluent. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS–PAGE) revealed that the fractions containing mandelate racemase activity were homogeneous.

Purification of Mandelate Racemase from E. coli. A single colony of JM105 transformed with pMRtrc grown on M9 plates containing 0.2% glucose as carbon source, 10 μg/mL thiamine, and 35 μg/mL ampicillin was used as the innoculum for 4 mL of LB medium containing 35 μg/mL ampicillin. After growth at 37 °C for 10 h, 1 mL of the liquid culture was used as the innoculum for 50 mL of the LB medium. After growth for 3 h, the 50-mL culture was used to innoculate 1 L of the LB medium. After incubation at 37 °C until the OD_{590} reached 0.7, IPTG was added to a final concentration of 0.2 mM to induce the gene for mandelate racemase. Two hours after induction, the cells were harvested by centrifugation (5000*g* at 4 °C for 15 min) and stored at −20 °C. All further steps were conducted at 0–4 °C except where specified otherwise.

Two-thirds of the frozen cells (2.4 g) were thawed and suspended in 25 mL of 50 mM Tris buffer, pH 8.0, containing 10 mM $MgCl_2$ and 0.1 mM PMSF and disrupted by sonication. Cell debris were removed by centrifugation (27000*g* for 20 min). Nucleic acids were removed from the supernatant by dropwise addition over 15 min of a solution of 25 mg of protamine sulfate in 3 mL of water adjusted to pH 7. After centrifugation (12000*g* for 20 min), the supernatant was adjusted to 25% saturation with ammonium sulfate (4.4 g in 30 mL); after stirring for 75 min, the precipitate was removed by centrifugation. Additional ammonium sulfate was added to the supernatant to 40% saturation (2.8 g in 30 mL); after stirring for 75 min, the precipitate was collected by centrifugation. The pellet was dissolved in and dialyzed overnight against the Tris buffer. The dialyzate was chromatographed on a column of DEAE-Sephacel (2.5 cm × 26 cm) by gradient elution as described for isolation of the enzyme from *P. putida*. Fractions containing 90% of the mandelate racemase activity were combined and concentrated by ultrafiltration (Amicon YM-30 membrane) to a final volume of 3.9 mL. The concentrated solution of enzyme was further fractionated by FPLC on a preparative Mono Q HR 10/10 column by elution with a linear gradient of ammonium bicarbonate, pH 8.5 (0.39–0.55 M), over 40 min at a flow rate of 4 mL/min. SDS–polyacrylamide gel electrophoresis (PAGE) of fractions containing the highest mandelate racemase activity revealed the purity to be greater than 95%. This purification is summarized in Table II. The major difference between this isolation of mandelate racemase from *E. coli* transformed with pMRtrc and that described for the enzyme from *P. putida* is that the gel filtration column was found to be unnecessary for the enzyme from *E. coli* due to the nature of the contaminating proteins.

N-Terminal Sequence Analyses of Mandelate Racemase from P. putida and E. coli. One nanomole of the purified mandelate racemase was dried by vacuum centrifugation (Savant Speedvac) and used for N-terminal sequence analysis with an Applied Biosystems Model 470A gas-phase sequenator. These determinations were performed in the Biomolecular Resource Center located at the University of California, San Francisco.

Assay of Mandelate Racemase. During purification from either *P. putida* or *E. coli*, mandelate racemase activity was determined by use of (*R*)-mandelate as substrate and coupling production of (*S*)-mandelate to the (*S*)-mandelate dehydrogenase as described by Hegeman (1970). Additional assays to measure enzyme activities in extracts of cells (either *P. aeruginosa* or *E. coli*) transformed with plasmids containing the cloned gene were performed by use of (*R*)-mandelate as substrate and measurement of the racemization directly by the circular dichroism assay described by Sharp et al. (1979).

Results and Discussion

Molecular Weight and N-Terminal Sequence Analysis of Authentic Mandelate Racemase from P. putida. In 1974 one of our laboratories reported that homogeneous mandelate racemase purified according to the procedure described by Hegeman (1970) had a subunit molecular weight of 69 500 as assessed by SDS–PAGE and a tetrameric quaternary structure as assessed by cross-linking with dimethyl suberimidate (Fee et al., 1974a). In the course of the present study, highly purified enzyme was obtained for N-terminal sequence analysis, and the molecular weight of the enzyme was redetermined by SDS–PAGE. Homogeneous enzyme obtained by Mono Q anion-exchange chromatography was found to have a molecular weight of ca. 41 000. No reason for this discrepancy is readily apparent. A sample of enzyme isolated in 1974, and kept in buffer in a refrigerator for 13 years, was observed to comigrate with a recently isolated, homogeneous enzyme preparation. Apparently, the molecular weight determination made in 1974, perhaps based on technology that was less sophisticated, was not reliable. Efforts are underway to reexamine the subunit composition of the enzyme in light of the revised molecular weight.

The newly isolated, homogeneous enzyme was subjected to N-terminal sequence analysis by an automated Edman degradation procedure. Sixteen cycles of degradation were performed, with the identity of the first residue being uncertain. The sequence of the next 15 residues was established as Glu-Val-Leu-Ile-Thr-Gly-Leu-Arg-Thr-Arg-Ala-Val-Asn-Val-Pro.

Cloning of the Gene for Mandelate Racemase. The gene for mandelate racemase was expeditiously cloned with a powerful metabolic selection procedure based upon the differences in the mandelate pathways in *P. putida* and in *P. aeruginosa*; the former includes mandelate racemase, but the

latter does not. This difference has the practical consequence that wild type *P. putida* can utilize either (*R*)- or (*S*)-mandelate as carbon source but wild-type *P. aeruginosa* can utilize only (*S*)-mandelate as carbon source. Thus, the strategy for the selection of the cloned gene for mandelate racemase was to prepare a library of *P. putida* DNA partially restricted with *Sau*3A in the *Bam*HI site of pKT230, a vector compatible with a range of Gram-negative organisms, and to transform this library into a wild-type *P. aeruginosa* host; plating of the transformed host on minimal medium plates containing (*R*)-mandelate as the carbon source would yield as transformants only those colonies containing the cloned gene. Our experience indicated that this screen could be best applied after an initial plating of the transformed host on LB agar containing streptomycin, thereby selecting for transformants, followed by replicate plating of the transformants on minimal medium plates containing (*R*)-mandelate as carbon source (as well as streptomycin). This selection afforded two colonies able to utilize (*R*)-mandelate as carbon source from a total of approximately 3000 transformants. One of these grew more rapidly than the other on minimal medium plates containing (*R*)-mandelate as carbon source and was subjected to further characterization; the plasmid contained in these cells was designated pSCR1.

As previously detailed, the library was constructed by the ligation of DNA fragments ranging from 2 to 5 kb in length obtained by digestion with *Sau*3A in the *Bam*HI site of pKT230. Restriction digestion of pSCR1 revealed that neither of the *Bam*HI sites that could have been generated by fragment insertion were present and that the fragment of *P. putida* that had been cloned had a length of approximately 5 kb. pSCR1 was found to possess two *Hin*dIII sites (one in vector DNA and one in the cloned DNA); subcloning of the smaller *Hin*dIII fragment (approximately 3.6 kb in length, with 3 kb of this fragment being DNA derived from *P. putida*) into the unique *Hin*dIII of pKT230 resulted in the generation of a plasmid that allowed the wild-type *P. aeruginosa* host to grow on (*R*)-mandelate as carbon source as well as imparting resistance to kanamycin (implying the same orientation with respect to the vector DNA as in pSCR1). The restriction analysis of pSCR1 also revealed the presence of a unique *Bam*HI site approximately in the middle of the smaller *Hin*dIII fragment that contains the gene for the racemase. Restriction of pSCR1 at the unique *Bam*HI followed by treatment with the Klenow fragment of DNA polymerase in the presence of the four deoxynucleoside triphosphates and finally DNA ligase yielded a plasmid that did not permit the growth of the *P. aeruginosa* host on minimal medium plates containing (*R*)-mandelate as carbon source. These observations suggested that the smaller *Hin*dIII fragment from pSCR1 contained the gene for catalytically active mandelate racemase and that the unique *Bam*HI site present in pSCR1 was located within the gene for the racemase.

DNA Sequence Analysis of the Gene and Its 5′-Flanking Sequence and the Deduced Primary Structure of Mandelate Racemase. Both *Hin*dIII–*Bam*HI fragments from pSCR1 were subcloned into pUC19 and subjected to restriction analysis to ascertain the presence of useful restriction sites for further subcloning into M13mp18 and M13mp19. A limited restriction map of the *Hin*dIII fragment showing the location of the gene for mandelate racemase is shown in the top line of Figure 1. The bottom lines (a–d) summarize the subclones subjected to DNA sequence analysis. The sequence for the gene was determined from the subclones as shown in the figure. The sequence information determined with the "–40" M13

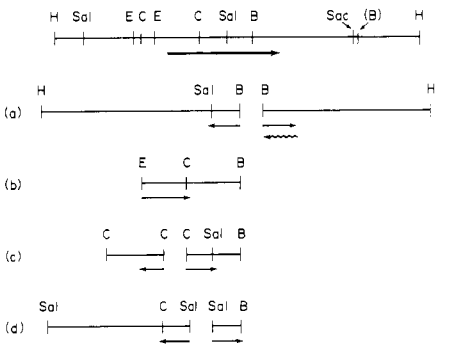

FIGURE 1: (Top line) Limited restriction map of the *Hin*dIII fragment containing the gene for mandelate racemase (indicated by the heavy arrow); (lines a–d) subclones used for DNA sequence determination of the gene. Details of the sequencing strategy are given in the text. Restriction sites: B, *Bam*HI; C, *Cla*I; E, *Eco*RV; H, *Hin*dIII; Sac, *Sac*I; Sal, *Sal*I. The site marked (B) shows the position of the *Bam*HI site used in cloning that was not regenerated by insertion of the *Sau*3A fragment and represents the border between the insert and vector sequences. The total length of the *Hin*dIII fragment is approximately 3.6 kb, and the positions of the restriction sites are drawn approximately to scale.

Table I: Observed[a] and Predicted Amino Acid Compositions for Mandelate Racemase

amino acid	observed	predicted	amino acid	observed	predicted
Ala	44	47	Leu	38	42
Arg	14	15	Lys	14	14
Asx	28		Met	9	12
Asn		7	Phe	9	8
Asp		17	Pro	17	19
Cys	9	2	Ser	22	17
Glx	35		Thr	21	19
Gln		15	Trp	5	6
Glu		19	Tyr	8	9
Gly	30	30	Val	30	31
His	8	9	total	359	359
Ile	18	21			

[a] Kenyon and Hegeman, 1979; Fee, 1974.

sequencing primer is denoted with the straight arrows; the sequence from the 3′-end of the gene contained in the right-hand *Bam*HI–*Hin*dIII subclone was determined with a primer synthesized with data obtained from the complementary strand and is denoted with the wavy arrow.

The coding sequence for the gene is shown in Figure 2 along with the translated amino acid sequence; also shown is the 5′-flanking sequence for the 240 bp immediately upstream of the gene. The N-terminal sequence of the open reading frame following the initiating methionine agrees with the 15 residues at the N-terminal sequence of the protein established by automated Edman degradation; the results of the automated sequence analysis and the deduced amino acid sequence indicate that the mandelate racemase produced by *P. putida* is processed to remove the initiating *N*-formylmethionine residue and that the N-terminus of the protein is a serine residue. The experimentally determined (on enzyme greater than 95% pure as assessed by SDS–PAGE) and deduced amino acid compositions of the racemase are compared in Table I; these are in good agreement with the exception of the number of cysteine residues. Finally, the molecular weight of the polypeptide predicted on the basis of the amino acid sequence is 38 570, a value in good agreement with the subunit molecular weight of ca. 41 000 determined by SDS–PAGE.

Thus, the molecular weight of the polypeptide in mandelate racemase must be revealed from the value of 69 500 originally reported (Fee et al., 1974a). This revision may be mechanistically significant since the larger value was used to calculate

```
                -220                -200                -180
TCGTTAAAGCACTTGCGCTAGTGCTAGGCTGTACTCCTGGGTCTGCAACTTTGTATGGCCTTGCAGCACGAGGTGAAACG

                -140                -120                -100        -90
GGTGTTGACGAGGTGCTAACCCTCCTAAAAGCGGATATCGACGCACCCTTGCCCAGATTGGATGCCCTGACATCACCTCC

        -70         -60         -50         -40         -30         -20         -10
CTTTCTCCTGATTACCTCCAAAACGAGGGAGTGACTAACACCGCTCCAGTCGATCACCTCATTGGTAAAGGAACACACGC

                                    30                                          60
ATG AGT GAA GTA CTG ATT ACC GGC CTG AGG ACG CGG GCC GTC AAT GTC CCA TTG GCC TAC
Met Ser Glu Val Leu Ile Thr Gly Leu Arg Thr Arg Ala Val Asn Val Pro Leu Ala Tyr

                                    90                                         120
CCC GTT CAC ACC GCT GTT GGA ACT GTT GGC ACA GCG CCT CTT GTT CTG ATC GAC CTG GCC
Pro Val His Thr Ala Val Gly Thr Val Gly Thr Ala Pro Leu Val Leu Ile Asp Leu Ala

                                   150                                         180
ACC AGT GCC GGC GTG GTA GGC CAT TCC TAC CTG TTC GCA TAC ACC CCC GTT GCG TTG AAG
Thr Ser Ala Gly Val Val Gly His Ser Tyr Leu Phe Ala Tyr Thr Pro Val Ala Leu Lys

                                   210                                         240
TCG CTG AAG CAG CTC CTG GAC GAC ATG GCA GCC ATG ATC GTC AAC GAA CCA TTG GCA CCG
Ser Leu Lys Gln Leu Leu Asp Asp Met Ala Ala Met Ile Val Asn Glu Pro Leu Ala Pro

                                   270                                         300
GTT AGC TTG GAA GCC ATG CTC GCA AAA CGC TTC TGC CTG GCA GGT TAT ACG GGA TTG ATC
Val Ser Leu Glu Ala Met Leu Ala Lys Arg Phe Cys Leu Ala Gly Tyr Thr Gly Leu Ile

                                   330                                         360
CGA ATG GCC GCT GCT GGC ATC GAT ATG GCA GCT TGG GAT GCG CTG GGC AAG GTT CAC GAA
Arg Met Ala Ala Ala Gly Ile Asp Met Ala Ala Trp Asp Ala Leu Gly Lys Val His Glu

                                   390                                         420
ACG CCA CTA GTA AAA CTC CTC GGC GCG AAC GCT CGA CCA GTT CAG GCT TAT GAC AGC CAC
Thr Pro Leu Val Lys Leu Leu Gly Ala Asn Ala Arg Pro Val Gln Ala Tyr Asp Ser His

                                   450                                         480
AGC TTG GAT GGA GTC AAA CTC GCG ACT GAG CGC GCT GTG ACC GCA GCA GAA CTC GGA TTC
Ser Leu Asp Gly Val Lys Leu Ala Thr Glu Arg Ala Val Thr Ala Ala Glu Leu Gly Phe

                                   510                                         540
CGG GCG GTT AAG ACC AAG ATC GGC TAT CCG GCA TTG GAT CAA GAT CTG GCA GTC GTG CGC
Arg Ala Val Lys Thr Lys Ile Gly Tyr Pro Ala Leu Asp Gln Asp Leu Ala Val Val Arg

                                   570                                         600
AGC ATA CGC CAA GCA GTA GGT GAC GAC TTC GGC ATC ATG GTC GAC TAC AAC CAG AGT TTG
Ser Ile Arg Gln Ala Val Gly Asp Asp Phe Gly Ile Met Val Asp Tyr Asn Gln Ser Leu

                                   630                                         660
GAT GTA CCA GCC GCA ATC AAA CGC AGC CAG GCG CTG CAG CAA GAG GGC GTC ACC TGG ATT
Asp Val Pro Ala Ala Ile Lys Arg Ser Gln Ala Leu Gln Gln Glu Gly Val Thr Trp Ile

                                   690                                         720
GAA GAG CCG ACG CTT CAA CAC GAT TAC GAA GGC CAT CAG CGC ATT CAA AGC AAG CTC AAT
Glu Glu Pro Thr Leu Gln His Asp Tyr Glu Gly His Gln Arg Ile Gln Ser Lys Leu Asn

                                   750                                         780
GTG CCC GTC CAG ATG GGT GAG AAC TGG CTC GGC CCT GAG GAG ATG TTC AAA GCA TTG AGC
Val Pro Val Gln Met Gly Glu Asn Trp Leu Gly Pro Glu Glu Met Phe Lys Ala Leu Ser

                                   810                                         840
ATC GGT GCA TGC CGG TTG GCT ATG CCA GAT GCA ATG AAG ATC GGT GGC GTG ACG GGC TGG
Ile Gly Ala Cys Arg Leu Ala Met Pro Asp Ala Met Lys Ile Gly Gly Val Thr Gly Trp

                                   870                                         900
ATC CGA GCC AGT GCA CTC GCA CAA CAA TTC GGT ATT CCA ATG TCC AGC CAC CTG TTC CAA
Ile Arg Ala Ser Ala Leu Ala Gln Gln Phe Gly Ile Pro Met Ser Ser His Leu Phe Gln

                                   930                                         960
GAA ATC AGC GCT CAC CTG CTA GCC GCC ACC CCA ACT GCG CAT TGG CTG GAG CGT TTG GAT
Glu Ile Ser Ala His Leu Leu Ala Ala Thr Pro Thr Ala His Trp Leu Glu Arg Leu Asp

                                   990                                        1020
CTC GCC GGC AGC GTC ATC GAG CCT ACT CTT ACA TTT GAG GGC GGA AAT GCT GTG ATT CCG
Leu Ala Gly Ser Val Ile Glu Pro Thr Leu Thr Phe Glu Gly Gly Asn Ala Val Ile Pro

                                  1050                                        1080
GAT CTC CCT GGC GTT GGG ATC ATC TGG CGC GAG AAA GAA ATC GGG AAA TAT CTG GTG TAA
Asp Leu Pro Gly Val Gly Ile Ile Trp Arg Glu Lys Glu Ile Gly Lys Tyr Leu Val End
```

FIGURE 2: DNA sequence of the gene and deduced primary sequence of the polypeptide for mandelate racemase from *P. putida*.

the 1:1 stoichiometry of covalent binding of the active site directed irreversible inhibitor α-phenylglycidate (Fee et al., 1974b) and also of the divalent cation required for catalysis (Maggio et al., 1975). The new value suggests that these stoichiometries may actually be approximately 0.5 molecule of inhibitor and divalent metal ion per active site. Since the inhibition and alkylation by α-phenylglycidate are protected by substrate, the less than stoichiometric alkylation may provide evidence that the inhibition displays half of sites reactivity such that alkylation and inactivation of an active site inhibit the reaction at a remote active site. These stoichiometries are being reinvestigated.

We have sequenced approximately 1100 bp of the 5′-flanking sequence, and the 240 bp immediately upstream of the initiation codon are also shown in Figure 2. Although relatively few *Pseudomonas* promoters have been sequenced, the underlined sequence approximately 90 bp upstream of the gene (CTGACATC) is homologous to the −26 region of the consensus sequence proposed for the promoters of *ntr*A-dependent genes in *Klebsiella* species (CTGGPyAPyPu); variations of this sequence have been found in additional putative pseudomonad promoters. However, the sequence we have determined does not appear to contain a sequence homologous the −12 region of these previously characterized promoters. Alternatively, the underlined sequence approximately 90 bp upstream of the gene and the underlined sequence approximately 65 bp upstream of the gene (GATTAC) are homologous to the consensus sequences for *E. coli* promoters (TTGACA and TATAAT, respectively) (Rosenberg & Court, 1979). Thus, we conclude that a competent promoter for the mandelate operon is immediately upstream of the gene for the racemase so that the first gene in the operon codes for the first enzyme in the pathway. The upstream sequence also reveals the presence of a putative ribosome binding site (AGGA) (also underlined) appropriately located 8 bp upstream of the initiation codon.

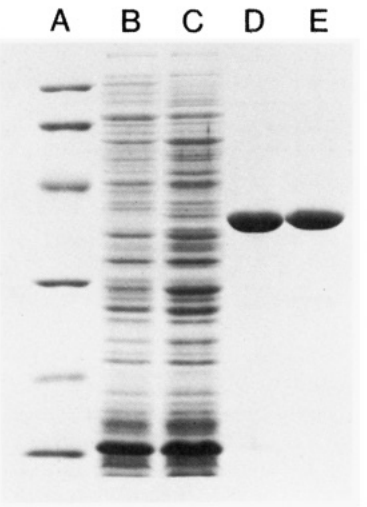

FIGURE 3: SDS–PAGE (12.5% acrylamide) showing (lane A) molecular weight standards [(top to bottom) 94K, 67K, 43K, 30K, 20.1K, and 14.4K], (lane B) JM105 transformed with pMRtrc before induction by IPTG, (lane C) JM105 transformed with pMRtrc after induction by IPTG, (lane D) mandelate racemase purified from induced JM105 transformed with pMRtrc, and (lane E) mandelate racemase purified from *P. putida*.

Expression of the Gene in E. coli. Mechanistic and structural studies of mandelate racemase would be facilitated by a convenient source of large amounts of the enzyme. As an initial step toward achieving this goal, we have ascertained whether the racemase can be produced in *E. coli* by fusing the coding sequence for the enzyme to efficient promoters and ribosome binding sites. Two systems were selected: the hybrid *trc* promoter with the *lacZ* ribosome binding site as contained in the commercially available pKK233-2 vector (Amann & Brosius, 1985) and the bacteriophage λ P_L promoter with the λ *cII* gene ribosome binding site as contained in vectors constructed by Dr. Martin Rosenberg and co-workers (Shatzman & Rosenberg, 1986); each type of vector contains a unique restriction size incorporating the initiation codon.

As described under Materials and Methods, we exploited the presence of a unique *Sca*I site (AGTACT) at the fourth codon in the racemase gene to fuse the coding sequence to the initiation codon found in each of these vectors.

Induction of the racemase gene under the control of the hybrid *trc* promoter (in the plasmid designated pMRtrc) by addition of IPTG to cells in the midlog phase resulted in the production of soluble and catalytically active mandelate racemase in cell extracts; examination of the total cellular protein before and after induction by SDS–PAGE revealed that approximately 1% of the total cellular protein was present

Table II: Summary of Purification of Mandelate Racemase from *E. coli* JM105 Transformed with pMRtrc

purification step	total vol (mL)	total protein (mg)	total act. (units)	sp act. (units/mg)
crude extract	27.0	370	580	1.6
protamine sulfate	30.5	320	550	1.7
$(NH_4)_2SO_4$ (0–25% sat.) supernatant	30	190	520	2.8
$(NH_4)_2SO_4$ (25–40% sat.) precipitate	10	22	390	18
DEAE-Sephacel	50	2.6	230	90
Mono Q	0.15	0.42	66	160

as mandelate racemase after induction (Figure 3, lanes B and C). Induction of the racemase gene under the control of the bacteriophage λ P_L promoter in either of two analogous expression plasmids by heat denaturation of the temperature-sensitive λ repressor encoded by the host chromosome resulted in the production of an insoluble polypeptide having the same molecular weight as authentic mandelate racemase; examination of the total cellular protein before and after induction revealed that approximately 10% of the total cellular protein was present as this insoluble protein after induction (data not shown). We assume that the insolubility of the protein produced by induction of the bacteriophage λ P_L promoter is the result of incorrect folding of the rapidly synthesized polypeptide.

Purification and Characterization of Mandelate Racemase Isolated from E. coli Transformed with PMRtrc. Homogeneous mandelate racemase was isolated from *E. coli* strain JM105 transformed with pMRtrc according to the procedure described under Materials and Methods and summarized in Table II. The SDS–PAGE gel shown in Figure 3 also compares the enzymes isolated from *E. coli* and *P. putida* (lanes D and E).

The specific activity of the enzyme isolated from *E. coli* transformed with pMRtrc (160 units/mg) is the same within the error of the coupled mandelate dehydrogenase assay as that of enzyme isolated from *P. putida* (180 units/mg).

The enzyme isolated from *E. coli* was subjected to N-terminal sequence analysis to determine whether any amino acid residues were removed subsequent to translation. Six cycles of automated Edman degradation were performed, and the sequence Met-Glu-Val-Leu-Ile-Thr was revealed. This sequence is that predicted by the procedure used to construct the expression plasmid (which deletes the second codon of the cloned gene) and indicates that no N-terminal processing of the protein occurs in *E. coli* (recall that the enzyme isolated from *P. putida* is processed to remove the N-terminal methionine residue).

Comparison of the DNA and Amino Acid Sequences with Data Bases. The mandelate racemase encoding DNA sequence and deduced amino acid sequence were compared with available data bases [nucleic acid: Genbank (May 20, 1987, release) and Dayhoff NBRF (January 15, 1987, release); protein: Dayhoff NBRF (March 17, 1987, release)] using the programs DFASTN and DFASTP for nucleic and protein sequences, respectively (Lipman & Pearson, 1985), which were available at the UCSF Biomathematical Computational Laboratory. No significant similarities were observed with any of the sequences searched.

Summary. The cloning and DNA sequence determination of the gene for mandelate racemase from *P. putida* has been accomplished. The expression of the gene in *E. coli* to moderate levels of total cellular protein is facilitating isolation of the racemase for further detailed mechanistic studies. Finally, the availability of the cloned and sequenced gene is likely to allow isolation and characterization of interesting random mutants of the gene and its regulatory regions.

ACKNOWLEDGMENTS

We thank Professor L. Nicholas Ornston and Dr. Ka-Leung Ngai for valuable discussions and plasmids, Professor George D. Hegeman for helpful advice and support, and Amy Tsou for experimental assistance.

Registry No. DNA (*Pseudomonas putida* mandelate racemase gene), 111793-51-2; madelate racemase (*Pseudomonas putida* reduced), 111793-52-3; mandelate racemase, 9024-04-8.

REFERENCES

Amann, E., & Brosius, J. (1985) *Gene 40*, 183.
Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F. C. H., Bagdasarian, M. M., Frey, J., & Timmis, K. N. (1981) *Gene 16*, 237.
Bagdasarian, M. M., Amann, E., Lurz, R., Rückert, B., & Bagdasarian, M. (1983) *Gene 26*, 273.
Fee, Judith A. (1974) Ph.D. Dissertation, University of California.
Fee, Judith A., Hegeman, G. D., & Kenyon, G. L. (1974a) *Biochemistry 13*, 2528.
Fee, Judith A., Hegeman, G. D., & Kenyon, G. L. (1974b) *Biochemistry 13*, 2533.
Gunsalus, I. C., Gunsalus, F. C., & Stanier, R. Y. (1953) *J. Bacteriol. 66*, 538.
Hegeman, G. D. (1966) *J. Bacteriol. 91*, 1140.
Hegeman, G. D. (1970) *Methods Enzymol. 17*, 670.
Kenyon, G. L., & Hegeman, G. D. (1979) *Adv. Enzymol. Relat. Areas Mol. Biol. 50*, 325.
Lipman, D. J., & Pearson, W. R. (1985) *Science (Washington, D.C.) 22*, 1435.
Maggio, E. T., Kenyon, G. L., Mildvan, A. S., & Hegeman, G. D. (1975) *Biochemistry 14*, 1131.
Mizusawa, S., Nishimura, S., & Seela, F. (1986) *Nucleic Acids Res. 14*, 1319.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319.
Sharp, T. R., Hegeman, G. D., & Kenyon, G. L. (1979) *Anal. Biochem. 94*, 329.
Shatzman, A. R., & Rosenberg, M. (1986) *Ann. N.Y. Acad. Sci. 478*, 233.
Whitman, C. P., Hegeman, G. D., Cleland, W. W., & Kenyon, G. L. (1985) *Biochemistry 24*, 3936.