3-Carboxy-*cis,cis*-muconate Lactonizing Enzyme from *Pseudomonas putida* Is Homologous to the Class II Fumarase Family: A New Reaction in the Evolution of a Mechanistic Motif†

Suzanne E. Williams,‡,§ Elisa M. Woolridge,‡ Stephen C. Ransom,‡,∥ James A. Landro,‡,⊥ Patricia C. Babbitt,# and John W. Kozarich*,‡

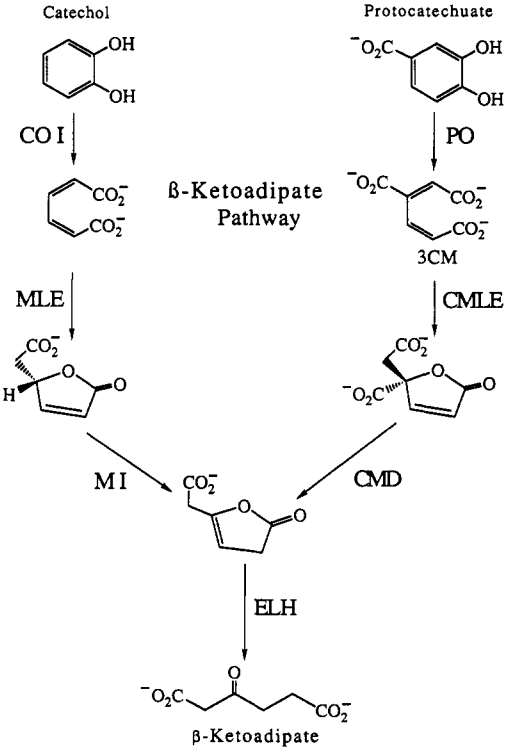
Department of Chemistry and Biochemistry and Center for Agricultural Biotechnology, University of Maryland, College Park, Maryland 20742, and Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143

Received May 11, 1992; Revised Manuscript Received July 16, 1992

ABSTRACT: The gene (*pca*B) for 3-carboxymuconate lactonizing enzyme (CMLE; 3-carboxymuconate cycloisomerase; EC 5.5.1.2) from *Pseudomonas putida* has been cloned into pMG27NS, a temperature-sensitive expression vector, and expressed in *Escherichia coli* N4830. The specific activity and kinetic parameters of the recombinant CMLE were comparable to those previously reported. A comparison of the deduced amino acid sequence of CMLE with sequences available in the PIR and Genbank databases revealed that CMLE has highly significant sequence homology to the class II fumarase family, particularly to adenylosuccinate lyase from *Bacillus subtilis*. CMLE has no significant homology to muconate lactonizing enzyme (MLE) from *P. putida*, its sister enzyme in the β-ketoadipate pathway. These findings fully corroborate a prediction made by us on the basis of mechanistic and stereochemical analyses of CMLE and MLE [Chari, R. V. J., Whitman, C. P., Kozarich, J. W., Ngai, K.-L., & Ornston, L. N. (1987) *J. Am. Chem. Soc. 109*, 5514–5519] and suggest that CMLE and MLE were recruited into this specialized pathway from two different enzyme families.

The β-ketoadipate pathway (Scheme I) consists of a battery of chromosomally encoded enzymes arrayed in two branches that are responsible for the catabolism of a variety of aromatic compounds by prokaryotes, such as *Pseudomonas putida* and *Acinetobacter calcoaceticus* (Stanier & Ornston, 1973; Ornston & Yeh, 1982). Two specific non-heme iron-dependent dioxygenases (CO I[1] and PO) effect ortho cleavage of the aromatic nucleus affording *cis,cis*-muconate from catechol and 3-carboxy-*cis,cis*-muconate (3CM) from protocatechuate. Muconate lactonizing enzyme (MLE; *cis,cis*-muconate cycloisomerase; EC 5.5.1.1) catalyzes the lactonization of muconate to muconolactone, while 3-carboxymuconate lactonizing enzyme (CMLE; 3-carboxymuconate cycloisomerase; EC 5.5.1.2) catalyzes the cyclization of 3CM to 4-carboxymuconolactone. Both branches converge via the formation of the enol lactone from muconolactone and from 4-carboxymuconolactone by the action of muconolactone Δ-isomerase

Scheme I



(MI) and 4-carboxymuconolactone decarboxylase (CMD), respectively. Hydrolytic ring-opening of the enol lactone by an enol-lactone hydrolase (ELH) affords β-ketoadipate. Complete biocombustion is accomplished by conversion of β-ketoadipate into acetyl-CoA and succinate by the action of thiolase and acetyl-CoA transferase.

The obvious similarities in the analogous reactions found in the two branches have suggested that common ancestral

‡ This research was supported by NIH Grant GM-37210 (J.W.K.) and NRSA Fellowship GM-14602-01 (E.M.W.). P.C.B. is supported by NIH Grants AR 17323 and GM 40570 to Dr. George L. Kenyon, UCSF.

* Author to whom correspondence should be addressed.
‡ University of Maryland.
§ Present address: American Red Cross, Rockville, MD 20855.
∥ Present address: NIDDK, Bethesda, MD 20892.
⊥ Present address: Department of Biology, MIT, Cambridge, MA 02139.
University of California.

[1] Abbreviations: ADSL, adenylosuccinate lyase; BSA, bovine serum albumin; 3-CM, 3-carboxy-*cis,cis*-muconate; CMD, 4-carboxymuconolactone decarboxylase; CMLE, 3-carboxy-*cis,cis*-muconate lactonizing enzyme (also known as 3-carboxy-*cis,cis*-muconate cycloisomerase); CO I, catechol 1,2-dioxygenase; DTT, dithiothreitol; ELH, enol-lactone hydrolase; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; IPTG, isopropyl β-D-thiogalactopyranoside; LB, Luria–Bertani medium; MI, muconolactone isomerase; MLE, *cis,cis*-muconate lactonizing enzyme (also known as *cis,cis*-muconate cycloisomerase); PIR, Protein Identification Resource; PO, protocatechuate 3,4-dioxygenase; RF, replicative form; Tris, tris(hydroxymethyl)aminomethane.

0006-2960/92/0431-9768$03.00/0

enzymes might serve as progenitors for the substrate-specific enzymes found in the converging branches. In some cases, evidence supports this contention; in others, it does not. Recent studies have shown that catechol 1,2-dioxygenase (CO I) and protocatechuate 3,4-dioxygenase (PO) have a significant degree of amino acid sequence homology. They differ, however, in oligomeric structure (Neidle et al., 1988). MI and CMD could also easily be envisioned as proceeding via a common intermediate since both enzymes afford the enol lactone. However, they have also been shown to have an opposing stereofacial specificity for substrate and for the enol lactone (Chari et al., 1987b). Little sequence homology appears to exist between this pair.[2]

The case for an evolutionary relationship between MLE and CMLE has been mixed. The cycloisomerases seem to differ only in specificity for the 3-substituent of the muconic acid, with both enzymes catalyzing a similar intramolecular 1,4-conjugate addition. However, several catalytic and mechanistic differences have argued for a more distant relationship. MLE binds one Mn^{2+} per subunit which is required for activity; CMLE has no apparent metal requirement. We have previously reported the determination of the absolute stereochemical course of the CMLE and of the MLE reactions (Chari et al., 1987a). The results established that MLE proceeds by a syn mechanism, while CMLE catalyzes an anti cycloisomerization. Moreover, MLE and CMLE recognize opposing mirror-image and out-of-plane conformers of their respective substrates, suggesting little active site homology with respect to either binding or catalysis.

We describe herein the cloning, sequencing, expression in *Escherichia coli*, and kinetic parameters of CMLE from *P. putida*. We also report on the results of an extensive amino acid sequence homology search for proteins that may be related to CMLE. The findings indicate that, while CMLE is not homologous to MLE, it has significant homology to the class II fumarate-dependent enzymes with remarkably high homology to adenylosuccinate lyase. CMLE is thus appropriately viewed as a new member of a broad class of anti addition enzymes.

MATERIALS AND METHODS

General. All chemicals were of the highest quality commercially available and used without further purification. 3-Carboxy-*cis*,*cis*-muconate was prepared according to the procedure of Ainsworth and Kirby (1968). The plasmid pPX81 that contains *pca*B, the gene for CMLE, and a portion of *pca*D, the gene for enol-lactone hydrolase (ELH) from *P. putida* strain PRS2000, was obtained from Dr. L. Nicholas Ornston, Yale University (Hughes et al., 1988). Bacteriophages M13mp18 and mp19 and *E. coli* JM 101 were obtained from New England Biolabs, Inc. *E. coli* DH5α was obtained from Bethesda Research Laboratories. The expression vector pMG27NS (pAS1 derivative, λ P_L, *CI857*, N, Nut, C^{II}, *cro*, gal K, t_o, amp^+) (Gross et al., 1985) was obtained from the laboratory of Dr. John Gerlt, University of Maryland. The gene product of *cI857* is a temperature-sensitive λ repressor, thus, the transcription of any gene placed downstream of the promoter is initiated via an increase in the temperature of the growth medium. pMG27NS and its derivatives were propagated in the *E. coli* strain N4830 [F^-, *su*$^-$, *his*$^-$, *ilv*$^-$, *gal*$^+$, Δ8(λ *cI857* Δ*Bam*HI)] (Gottesman et al., 1980), also obtained from Dr. John Gerlt. All other molecular biology reagents were obtained from commercial vendors. Large-scale (10 L) growth of *E. coli* N4830 (pMCW1) was conducted in a New Brunswick Scientific Co. Microferm fermentor. The Bio-Rad protein assay was used for protein concentration determination; BSA served as the standard protein. Kinetic assays were performed with a Gilford Response II UV–vis spectrophotometer.

*Sequencing of pca*B. The plasmid pPX81 was digested with *Bam*HI and *Sal*I to yield a fragment containing *pca*B and part of *pca*D. This fragment was cloned into *Bam*HI/*Sal*I restricted M13mp18 and mp19 for DNA sequence analysis, and the resulting replicative forms of M13 were designated pA151 and pB151, respectively. To facilitate complete sequencing of pA151, this plasmid was digested with *Pst*I and then religated, which resulted in a loss of approximately 600 bp from the C terminus of *pca*B and all of *pca*D. The shorter *pca*B RF was designated pA151-*Pst*. Sequencing was done by the Sanger dideoxy method using a United States Biochemical Corp. Sequenase kit. Regions of compression were resolved by incorporating deoxy-7-deazaguanosine 5′-triphosphate or deoxyinosine 5′-triphosphate in place of dGTP. Sequencing primers were synthesized by the University of Maryland Protein and Nucleic Acid Laboratory.

Expression of CMLE. The expression vector pMG27NS has an *Nde*I restriction site directly downstream of the P_L promoter. This site is convenient for protein expression as it contains the initiation sequence ATG. Site-specific mutagenesis was used to create an *Nde*I site at the beginning of *pca*B in pA151 with the oligonucleotide-directed in vitro mutagenesis system, version 2 (Amersham). The synthetic oligonucleotide used to introduce the *Nde*I site was 5′-CAG TTG GTT GGT CAT ATG ACG TCC TTA G-3′; the mutated nucleotides are underlined. The mutated plasmid (pHHH) was then digested with *Nde*I and *Hin*dIII, the fragment cloned into pMG27NS, and the resulting plasmid (pMCW1) transformed into *E. coli* N4830. The cultures were grown at 30 °C in LB at 35 μg/mL ampicillin to an absorbance of 1.0 OD at 600 nm, and then expression was induced by rapidly raising the culture temperature to 42 °C.

Assay for CMLE Activity. A modification of the procedure described by Ornston (1966) was used to measure CMLE activity. A stock solution of enzyme (4.7 μg/mL) was prepared fresh daily in either 0.1 M Tris-HCl, pH 7.5, or 0.1 M HEPES (K^+ salt), pH 7.5, containing 3 mM DTT and 0.5 mg/mL BSA. The assays (1.0 mL total volume) contained 0.1 M HEPES (K^+ salt), pH 7.5, 1 mM DTT, 0.005–0.77 mM 3CM, and 5 μL of the CMLE dilution (0.022 unit, 0.54 nM active site); assay was initiated with the addition of enzyme. Cuvettes with a path length of 0.2 cm were used for 3CM concentrations greater than 0.16 mM; otherwise, 1 cm path length cells were used. Activity assays for the purpose of monitoring CMLE activity during purification contained 0.16 mM 3CM. Substrate consumption was followed as the decrease in absorbance at 270 nm (ϵ_{270} = 6390 M^{-1} cm^{-1}); the absorbance of the product 4-carboxymuconolactone is negligible at this wavelength. The initial velocity was calculated from the absorbance decrease during the first minute after enzyme addition. A subunit M_r of 42 446 was used to calculate k_{cat}. One unit of CMLE activity is defined as that amount of enzyme required to convert 1.0 μmol of 3CM to product in 1 min at 25 °C.

Purification of Expressed CMLE. A 4-mL culture of *E. coli* N4830 (pMCW1) that had been grown for 12 h at 30 °C in LB with 35 μg/mL ampicillin was used to inoculate 200 mL of the same media. After growth at 30 °C for 8 h, the 200-mL culture was used to inoculate 10 L of LB/ampicillin media. The culture was stirred at 450 rpm and aerated at 8

[2] J. W. Kozarich, unpublished observations.

L/min. Growth proceeded for 5 h at 30 °C, and then the temperature was increased to 42 °C for an additional 2 h. Cells were harvested by centrifugation at 5000*g* for 15 min, and the pellet (69 g) was stored at −20 °C.

Frozen cells (30 g) were thawed and resuspended in 430 mL of cold 10 mM ethylenediamine dihydrochloride, pH 7.3, containing 1 μM $MnCl_2$ (buffer A). The cell suspension was placed on ice and disrupted by sonication, and the debris was removed by centrifugation (10000*g* for 20 min, 4 °C). All subsequent steps were performed at 4 °C.

A solution of protamine sulfate (510 mg in 30 mL of H_2O) was added dropwise to the stirring cell extract (450 mL) over a period of 10 min to remove nucleic acids. The mixture was stirred for an additional 15 min and centrifuged (10000*g*) for 20 min. The supernatant was then 60% saturated with ammonium sulfate (163 g in 450 mL) and, after stirring for 75 min, the precipitate collected by centrifugation. The pellet was resuspended in buffer A (140 mL) and dialyzed against 2 L of buffer A with a single buffer change after 4 h.

The dialysate was chromatographed on a DEAE-Sephacel column (2.5 cm × 30 cm) equilibrated with buffer A at a flow rate of 34 mL/h. After washing with buffer A until the eluate was free of absorbance at 280 nm, the protein was eluted with a linear gradient from 0.0 to 0.5 M NaCl in buffer A (1 L, total volume). Fractions (8 mL/tube) containing 75% of the CMLE activity were pooled (56 mL) and concentrated by ultrafiltration (Amicon PM-30 membrane) to 9.3 mL.

This CMLE solution was gel filtered on an Sephacryl S-200 column (2.5 cm × 95 cm) equilibrated in buffer A at a flow rate of 14 mL/h. Fractions (6.7 mL) containing 75% of the CMLE activity were pooled (47 mL) and concentrated by ultrafiltration to 4.25 mL. The enzyme solution was then divided into 0.25-mL aliquots and stored at −70 °C. Purity of CMLE was >95% as determined by SDS–PAGE.

N-Terminal Sequence Analysis of CMLE. Purified CMLE was concentrated and exchanged into water to a final concentration of 30 mg/mL. This preparation was used for N-terminal sequence analysis with an Applied Biosystems Model 477A protein sequencer. The sequence was determined by the University of Maryland Protein and Nucleic Acid Laboratory.

Homology Analysis and Alignment of CMLE. Sequence analysis was performed at the Computer Graphics Laboratory, UCSF. The deduced amino acid sequence of CMLE was compared to those in the Protein Identification Resource (PIR) data base (George et al., 1986; release 31) and to the coding regions of the Genbank database (Bilofsky et al., 1986; release 70) using the algorithms FASTA (Pearson & Lipman, 1988) and BLAST (Altschul et al., 1990). The highest-scoring sequences from these searches were subjected to further analysis to determine the statistical significance of their relationships to CMLE. This was done by pairwise comparisons with CMLE using the algorithms Lawrence (Lawrence & Goldman, 1988), Monte (Altschul & Erickson, 1986), and RSS (Pearson, 1990) based on the RDF2 algorithm. The Lawrence and Monte algorithms were used as presented in the Eugene sequence analysis package (Molecular Biology Information Resource, Baylor College of Medicine, Houston, TX). The statistical significance of the similarities between CMLE and each sequence was tested on the Monte algorithm using 50 randomizations of the CMLE sequence and on the RSS algorithm using 100 randomizations of the CMLE sequence. On the basis of the results of these analyses, the sequences considered to be the most like CMLE were multiply aligned in a master alignment using the PIMA algorithm (Smith & Smith, 1990, 1992).

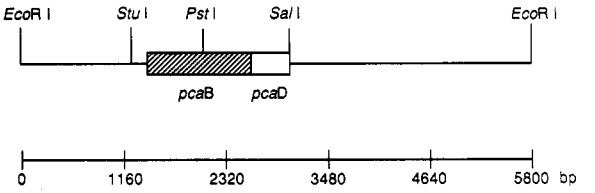


FIGURE 1: Linear map of pPX81 with relevant restriction sites labeled.

RESULTS

DNA Sequence Analysis of the CMLE Gene. The plasmid pPX81 was previously constructed by Tn5 mutagenesis of *P. putida* strain PRS2000 and subsequent cloning into *E. coli* (Hughes et al., 1988). It contains the full-length gene for CMLE, *pca*B, plus a portion of *pca*D, the gene for enol-lactone hydrolase. Figure 1 shows a linear map of pPX81; the significant restriction sites are labeled. A series of DNA oligomers were designed and synthesized so that both strands of *pca*B could be sequenced. The cloning of *pca*B into M13 allowed for single-stranded sequencing using the dideoxy method. The nucleotide and the deduced amino acid sequences are shown in Figure 2. A potential Shine–Dalgarno ribosome binding site (underlined) is located 6 bp upstream of the ATG initiation codon (Shine & Dalgarno, 1974). The CMLE gene encodes 407 amino acids and has a predicted monomer weight of 42 446. This agrees well with the Coomassie-stained band of purified CMLE at 43 000 (see Figure 4).

Expression of CMLE in E. coli N4830. Initial attempts to produce a viable expression system involved cloning *pca*B into a commercially available vector, pTZ19U (United States Biochemical; Yanisch-Perron et al., 1985). The resulting clone, pSWE1, was constructed using the *Stu*I/*Eco*RI fragment of pPX81 and contained all of *pca*B, a portion of *pca*D, and 200 bp upstream of *pca*B. This recombinant vector failed to produce viable transformed colonies in *E. coli* BL21-(DE3) (lac UV5 promoter; Studier & Moffatt, 1986). To exclude the possibility that the basal production of T7 polymerase in BL21(DE3) (without induction by IPTG) was affording potentially toxic levels of CMLE, pSWE1 was transformed into *E. coli* BL21(DE3)LysS and BL21(DE3)-LysE (Studier et al., 1990). T7 lysozyme effectively eliminates the uninduced basal level of T7 polymerase. This system also failed to produce viable colonies. Subsequent transformation of pSWE1 into *E. coli* ACT7, another host containing T7 polymerase, yielded viable colonies; however, no production of CMLE protein as evidenced by SDS–PAGE and activity assays with 3CM was detected. We concluded from these experiments that the 200 bp upstream of *pca*B in pSWE1 was affecting efficient transcription.

To eliminate this problem, we chose a new vector, pMG27NS, which contains a temperature-sensitive promoter (Gross et al., 1985). This vector contains two restriction sites that have ATG in their recognition sequences. These sites are located downstream of the strong λ promoter and allow one to clone a gene directly at its initiation methionine. Since *pca*B does not contain an *Nde*I site, we used site-directed mutagenesis to create one. The end product of the mutagenesis, pHHH, was digested with *Nde*I and *Hin*dIII, and the fragment containing *pca*B was cloned into pMG27NS to generate pMCW1 (Figure 3).

Transformation of pMCW1 into *E. coli* N4830 (Gottesman et al., 1980) provided a means to regulate CMLE production: a temperature increase inactivates the *cI857* λ repressor

```
CGATTGCCTTCCTGTTCAGCCTGCGCCTGCCGAAGCAGGCCGCGTACCTGCACCACGATGACTAAGGACGTGGC ATG ACC AAC CAA CTG TTC GAC GCC TAC TTC ACC GCG CCG GCC
                                                                              Met Thr Asn Gln Leu Phe Asp Ala Tyr Phe Thr Ala Pro Ala   14

ATG CGC GAG ATT TTC TCC GAC CGT GGC CGC TTG CAG GGC ATG CTC GAT TTC GAA GCC GCG CTG GCC CGT GCC GAA GCT GCT GCG GGG CTG GTC CCG CAC
Met Arg Glu Ile Phe Ser Asp Arg Gly Arg Leu Gln Gly Met Leu Asp Phe Glu Ala Ala Leu Ala Arg Ala Glu Ala Ala Ala Gly Leu Val Pro His   47

AGC GCC GTG GCC GCC ATC GAG GCG GCA TGC AAC GCC GAG CGC TAT GAC GTG GGG GCG CTG GCC AAT GCC ATC GCA ACC GCT GGC AAC TCG GCA ATC CCG
Ser Ala Val Ala Ala Ile Glu Ala Ala Cys Lys Ala Glu Arg Tyr Asp Val Gly Ala Leu Ala Asn Ala Ile Ala Thr Ala Gly Asn Ser Ala Ile Pro   80

CTG GTG AAC GCG TTG GGC AAG GTG ATT GCC AGT GGC GTG CCC GAG GCC GAG CGC TAT GTG CAC TTG GGC GCC ACC AGC CAG GAC GCG ATG GAC ACC GGT
Leu Val Lys Ala Leu Gly Lys Val Ile Ala Ser Gly Val Pro Glu Ala Glu Arg Tyr Val His Leu Gly Ala Thr Ser Gln Asp Ala Met Asp Thr Gly  113

CTG GTG TTG CAG TTG CGC GAT GCC CTC GAC CTG ATC GAA GCG GAC CTG GGC AAA CTG GCC GAT ACC CTG TCG CAG CAG GCA TTG AAG CAC GCC GAT ACG
Leu Val Leu Gln Leu Arg Asp Ala Leu Asp Leu Ile Glu Ala Asp Leu Gly Lys Leu Ala Asp Thr Leu Ser Gln Gln Ala Leu Lys His Ala Asp Thr  146

CCA ATG GTG GGC CGC ACC TGG CTG CAA CAC GCT ACC CCG GTG ACC CTG GGC ATG AAA CTG GCT GGC GTG CTG GGG GCA TTG ACC CGC CAC CGT CAG CGC
Pro Met Val Gly Arg Thr Trp Leu Gln His Ala Thr Pro Val Thr Leu Gly Met Lys Leu Ala Gly Val Leu Gly Ala Leu Thr Arg His Arg Gln Arg  179

CTG CAA GAG CTC GGC CCG CCC TGT TGG TGC TGC AGT TCC GGC GGC GCC TCG GGC AGC CTG GCC GCT CTG GGC AGC AAG GCG ATG CCG GTG GCC GAG GCC
Leu Gln Glu Leu Gly Pro Pro Cys Trp Cys Cys Ser Ser Gly Gly Ala Ser Gly Ser Leu Ala Ala Leu Gly Ser Lys Ala Met Pro Val Ala Glu Ala  212

CTG GCC GAG CAA CTG AAG CTG AGC CTG CCC GAG CAA CCC TGG CAC ACC CAG CGT GAT CGC CTG GTG GAG TTT GCC TCG GTG CTG GGC CTT GTG GCC GGC
Leu Ala Glu Gln Leu Lys Leu Ser Leu Pro Glu Gln Pro Trp His Thr Gln Arg Asp Arg Leu Val Glu Phe Ala Ser Val Leu Gly Leu Val Ala Gly  245

AGC CTG GGC AAG TTC GGC CGC GAT GTC AGC CTG CTG ATG CAA ACC GAG GCG GGG GAG GTG TTC GAG CCT TCT GCA CCA GGC AAG GGC GGC TCC TCG ACC
Ser Leu Gly Lys Phe Gly Arg Asp Val Ser Leu Leu Met Gln Thr Glu Ala Gly Glu Val Phe Glu Pro Ser Ala Pro Gly Lys Gly Gly Ser Ser Thr  278

ATG CCA CAC AAG CGC AAC CCG GTG GGC GCT GCG GTG CTG ATC GGT GCC GCG ACT CGC GTG CCG GGG CTG GTA TCG ACG CTG TTC GCC GCC ATG CCC CAG
Met Pro His Lys Arg Asn Pro Val Gly Ala Ala Val Leu Ile Gly Ala Ala Thr Arg Val Pro Gly Leu Val Ser Thr Leu Phe Ala Ala Met Pro Gln  311

GAG CAC GAG CGC AGC CTG GGC CTG TGG CAT GCC GAA TGG GAA ACC CTC CCG GAC ATC TGC TGC CTG GTC TCC GGC GCG CTG CGC CAG GCC CAG GTC ATT
Glu His Glu Arg Ser Leu Gly Leu Trp His Ala Glu Trp Glu Thr Leu Pro Asp Ile Cys Cys Leu Val Ser Gly Ala Leu Arg Gln Ala Gln Val Ile  344

GCC GAG GGC ATT GAA GTA GAC GCA GCG CGC ATG CGC CGT AAC CTC GAC CTG ACC CAA GGG CTG GTG CTG GCA GAA GCG GTG AGT ATC GTC CTG GCC CGA
Ala Glu Gly Ile Glu Val Asp Ala Ala Arg Met Arg Arg Asn Leu Asp Leu Thr Gln Gly Leu Val Leu Ala Glu Ala Val Ser Ile Val Leu Ala Arg  377

ACG CCT GGG CCG CGA CCG TGC CCA CCA CTT GCT GGA GCA ATG CTG CCA GCG AGC CGT GGC CGA ACA GCG GCA CCT GCG TGC GGT GCT GGG TGA TGA CCC
Thr Pro Gly Pro Arg Pro Cys Pro Pro Leu Ala Gly Ala Met Leu Pro Ala Ser Arg Gly Arg Thr Ala Ala Pro Ala Cys Gly Ala Gly End End      407

GCAGGTCAGTGCCGAGCTGTCTGCCGAAGAACTTGATCG
```

FIGURE 2: Nucleotide and translated amino acid sequence of *pca*B. A potential ribosomal binding site sequence is underlined. Numbers refer to amino acid positions.

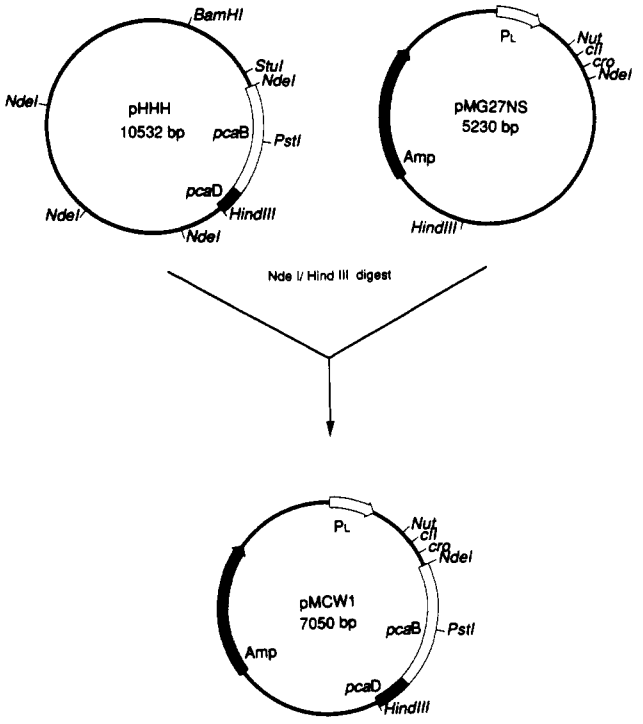


FIGURE 3: Cloning strategy for expression of *pca*B in pMG27NS. pHHH was digested with *Nde*I and *Hin*dIII to generate a 1820-bp fragment. This fragment was ligated into pMG27NS that had been digested with the same restriction enzymes.

protein, allowing transcription to occur. Cultures of *E. coli* N4830 (pMCW1) were grown under a variety of conditions and, using a heat induction step of 42 °C, an abundance of CMLE protein was expressed as evidenced by SDS–PAGE (data not shown).

Purification of CMLE. An earlier report of the purification of CMLE described a yield of 40 mg of CMLE (800 units/mg) from a 100-L culture of *P. putida* strain PRS2105 (Patel et al., 1973). Our expression system provided 525 mg of CMLE (1054 units/mg) from 30 g of cells from a 4-L culture. The isolation of CMLE from *E. coli* N4830 (pMCW1) was accomplished via standard protein separation techniques; Table I outlines the progress of the purification. SDS–PAGE was performed on an aliquot from each purification step (Figure 4); visualization of protein with Coomassie blue indicated that after Sephacryl S-200 gel filtration, the enzyme was at least 95% pure (lane 5). In addition to the described purification steps, CMLE was also chromatographed through a Pharmacia Superose 12 column (HR 16/50). This last step, however, did not yield a significant increase in the specific activity (1054 to 1097 units/mg) or purity of CMLE (data not shown). Efforts to obtain crystals of CMLE suitable for structural analysis are currently underway.

Table I: CMLE Purification[a]

step	total protein (mg)	total units (μmol/min)	specific activity (units/mg)
crude	3220	6.8×10^5	210
protamine sulfate	2590	7.5×10^5	288
ammonium sulfate (0–60%)	1640	7.6×10^5	461
DEAE-Sephacel	650	6.6×10^5	1019
Sephacryl S-200	595	6.3×10^5	1054

[a] CMLE activity was measured as the decrease in absorbance at 270 nm of 0.16 mM 3CM in 33 mM Tris-HCl, pH 8.0. One unit of CMLE activity is defined as that amount of enzyme required to cycloisomerize 1 μmol of 3CM in 1 min at 25 °C. The Bio-Rad assay was used for protein concentration determination; BSA served as the standard protein.

N-Terminal Sequence Analysis of CMLE. The expressed and purified CMLE was subjected to automated protein sequencing, and the first 15 amino acids were identified. The experimentally obtained residues matched the predicted amino acid sequence (Figure 2); however, the results revealed that two forms of CMLE were present. Thirty percent of the

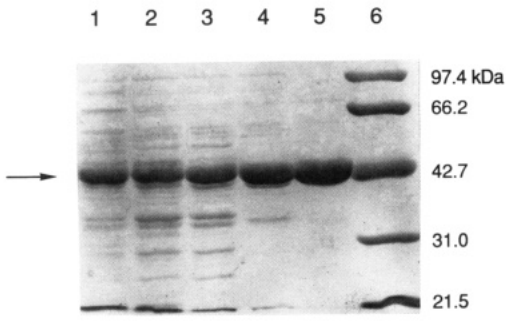


FIGURE 4: SDS–PAGE analysis of the purification of CMLE. Approximately 10 μg of each of the following samples from a CMLE purification was loaded onto a 10% polyacrylamide gel: lane 1, crude supernatant; lane 2, protamine sulfate supernatant; lane 3, 0–60% ammonium sulfate precipitate; lane 4, DEAE-Sephacel chromatography; lane 5, Sephacryl S-200 chromatography; lane 6, low molecular weight markers. Coomassie blue was used to visualize protein.

CMLE was missing the initiation methionine, and the first amino acid was threonine. This is consistent with some degree of posttranslational cleavage of methionine occurring in the host cell, *E. coli* N4830. The N-terminal sequence of the first five amino acids of CMLE had been reported previously as Ile-Leu-Val-Met-Ala (Patel et al., 1973). This is obviously quite different from our N-terminal sequence. Moreover, we could not assign this sequence to any within our deduced sequence. On the basis of the DNA sequence of *pca*B and the high specific activity and purity of our expressed CMLE, we conclude that the deduced amino acid sequence reported here is the correct one.

Characteristics of Purified CMLE. Kinetic parameters for the cycloisomerization of 3CM by CMLE are summarized in Table II. The K_M value (32 μM) agrees well with that described previously of 75 μM in 33 mM Tris-HCl, pH 8.0 (Ornston, 1966). Curiously, as the pH of the Tris buffer is lowered to 7.5, the K_M increases by a factor of 8. This appears to be due to inhibition of CMLE by chloride. Potassium chloride competitively inhibits 3CM cycloisomerization with a K_I of 15 ± 3 mM (data not shown). When Tris-HCl (pH 7.5) is substituted with K-HEPES, the K_M value is comparable to that observed in Tris buffer at pH 8.0. Purified CMLE was found to be relatively unstable in the dilutions necessary to obtain accurate kinetic data. The inclusion of BSA (0.5 mg/mL) in dilute enzyme assays enhanced CMLE stability. In addition, DTT (1 mM) had a modest, positive effect on CMLE activity (~23% enhancement; data not shown). CMLE contains eight cysteine residues per monomer.

Protein Sequence Analysis of CMLE. The initial screening of sequence databases allowed comparison of CMLE with over 36 000 known sequences. CMLE was found to bear probable homology to the adenylosuccinate lyases (Ebbole & Zalkin, 1987; Aimi et al., 1990), the argininosuccinate lyases (Beacham et al., 1984; Matsubasa et al., 1989; O'Brien et al., 1986; Matuo et al., 1988; Debuchy et al., 1989), the δ-crystallins (Nickerson et al., 1985, 1986), the aspartases (Takagi et al., 1985, 1986), and the class II fumarases (Miles & Guest, 1985; Woods et al., 1986; Wu & Tzagoloff, 1987; Kinsella & Donnan, 1986; Sacchettini et al., 1988). It is important to note that the above enzymes are all members of a family of structurally related proteins (vide infra).

On the basis of the original database searches, these sequences represented all those most similar to CMLE, with optimized scores ranging from 139 to 404. The significance of the homology of CMLE to this family is highlighted by noting that the next highest optimized score for the PIR database, 73, was that of the F2 protein of avian infectious bronchitis virus, which is not a member of this family. For the Genbank database, a similar result was obtained. Further FASTA searches were also performed using the member of each enzyme class in this family that was most related to CMLE as a probe. Results of these searches revealed a similar pattern as well; that is, whether CMLE, adenylosuccinate lyase, δ-crystallin, argininosuccinate lyase, aspartase, or fumarase was used as the sequence probe, all of the highest optimized scores for each search represented only members of the class II fumarase family. More rigorous analysis of the statistical significance of the relationships between CMLE and the highest scoring family member sequences was performed by comparing an alignment score between CMLE and each sequence with the distribution of scores generated by alignment of 50 or 100 randomizations of the CMLE sequence with each sequence. Scores from one such set of statistical analyses are shown in Table III.

Comparison of CMLE with sequences in the PIR database using the BLAST algorithm (Altschul et al., 1990) also produced similar results, although there were several sequences not in this family that scored better than a few of the least related family members. Further statistical analysis of these non-family-member sequences and the highest-scoring non-family-member sequences from the FASTA search suggested that these sequences were less likely to be related to CMLE than were the class II fumarase family (Table III).

The highest-scoring sequence from either the FASTA or BLAST search of the PIR was adenylosuccinate lyase from *B. subtilis* (ADSL; 431 aa; Ebbole & Zalkin, 1987). This sequence also showed the most statistically significant relationship to CMLE of any of the sequences reported in Table III. The RSS alignments score for this alignment is 56 (score >6 indicates probable homology). An alignment of the two sequences is shown in Figure 5. The homology between the two sequences extends over 374 amino acids in CMLE with an overall percent identity of 28.5%. Although there is significant sequence similarity throughout the entire alignment, a clustering of identical residues between the sequences occurs in two short regions—11 identities in CMLE residues 150–167 and 10 identities in CMLE residues 275–285. A portion of the first region and all of the second is shown aligned with several other members of the class II fumarase family in Figure 6. These regions show striking sequence conservation among the members of the class II fumarase family and CMLE. The overall alignment of these sequences (not shown) revealed that CMLE was most closely related to ADSL and argininosuccinate lyase followed by the δ-crystallins. It was least related to the aspartases and the class II fumarases (typified by FUMC of *E. coli* and the mammalian fumarases). For one alignment generated by the PIMA algorithm, the highest score (141) was obtained for CMLE with ADSL from *B. subtilis* and the lowest for fumarase from pig heart (91). These scores are different from the RSS statistical significance analysis shown in Table III because the alignments were generated differently. The two algorithms were in general overall agreement, however, in the identification of the sequences most similar to CMLE.

The relationship between argininosuccinate lyase and the δ-crystallins has been the subject of much attention; the crystallin genes apparently evolved by recruitment and duplication of preexisting argininosuccinate lyase genes (Matsubasa et al., 1989; Yeh et al., 1988; Piatigorsky et al., 1988; Takiguchi et al., 1989; Mori et al., 1990). Homology between the aspartases and fumarases has also been noted (Woods et al., 1986; Takagi et al., 1986) as has that of

Table II: Kinetic Parameters for the Cycloisomerization of 3CM by CMLE[a]

buffer	K_M (mM)	k_{cat} (s^{-1})	$10^{-7} \times k_{cat}/K_M$ (M^{-1} s^{-1})
K-HEPES, pH 7.5 (0.1 M)	0.030 ± 0.006	460 ± 70	1.6 ● 0.4
Tris-HCL, pH 7.5 (0.1 M)	0.25 ● 0.01	600 ± 10	0.24 ± 0.09
Tris-HCl, pH 8.0 (0.033 M)	0.032 ± 0.004 (0.075)[b]	370 ± 10 (570)[c]	1.2 ● 0.1

[a] Assays were conducted as described under Materials and Methods. [b] Ornston (1966). [c] Calculated from the specific activity of 800 units/mg reported by Patel et al. (1973).

Table III: Statistical Analysis of Alignments between CMLE and Selected Proteins[a]

identifier	protein	score
	Sequences Most Likely To Be Related to CMLE	
WXBSDS	adenylosuccinate lyase (*B. subtilis*)	56
A35291	adenylosuccinate lyase (chicken)	16
B31658	argininosuccinate lyase (rat)	16
S01010	argininosuccinate lyase (human)	14
B31658	argininosuccinate lyase (yeast)	12
JU0453	δ-crystallin 2 (duck)	21
A25622	δ-crystallin 1 (chicken)	19
UFPSDF	aspartase (*P. fluorescens*)	10
UFECDW	aspartase (*E. coli*)	8
A29804	fumarase (yeast)	17
UFBSC8	fumarase (*B. subtilis*)	12
S06213	fumarase (human)	11
S07138	fumarase (*E. coli*)	10
	Sequences Less Likely To Be Related to CMLE	
A27316	MLE I (*P. putida*)	−0.2
B27316	MLE II (*P. putida*)	−0.4
B33094	F2 protein (avian infectious bronchitis virus)	5
A34796	kinesin-related protein (bovine herpes virus)	3
S04793	mobilization protein (*E. coli*)	4
A32001	thrombomodulin (mouse)	4

[a] The RSS analysis was performed by comparing a similarity score generated between CMLE and each sequence listed to a distribution of scores obtained for the same comparison but using 100 randomizations of the CMLE sequence to obtain the distribution. What is expressed as the RSS score is the number of standard deviations from the mean of that distribution that the score for the authentic CMLE represents. A rough guideline for interpreting these scores is that scores greater than 6 are suggestive of a relationship between sequences while scores less than 3 are doubtful (Pearson, 1990). The class II fumarate-dependent enzymes with the greatest similarities to CMLE based on FASTA, BLAST, Monte Carlo-type, and RSS algorithm analysis are listed in the first part of the table. Not included are the class I fumarases which scored lower in these analyses. For comparison, the lowest-scoring of these sequences was the iron-dependent fumarase from *E. coli* (PIR UFECAQ). For the above analysis conditions, this sequence had an RSS score of 5, similar to the highest of the scores listed in the lower section of the table.

aspartase, ADSL, and argininosuccinate lyase (Aimi et al., 1990). Like CMLE, all of these enzymes (δ-crystallins excluded) have been characterized as tetrameric enzymes. Thus, the interrelationship between CMLE and these proteins suggests that CMLE is also a member of this family.

The relationship is particularly noteworthy in the two regions that are particularly well-conserved among sequences of the class II fumarases and related enzymes (Figure 6). It has been suggested that the most highly conserved region (Figure 6B) represents a "signature sequence" for enzymes which catalyze β-elimination reactions and generate fumarate as product (Aimi et al., 1990). Woods et al. (1988) have also proposed that this region constitutes a consensus sequence, GSxxMxxKxN, that relates the class I and class II fumarases. The class I fumarases (such as FUMA and FUMB of *E. coli*) are thermolabile dimeric enzymes (M_r 60 000) and represent a distinct family of enzymes. From a kinetic standpoint, the large k_{cat}/K_M of 10^7 M^{-1} s^{-1} for CMLE is similar to that reported for fumarase (Teipel et al., 1968).

DISCUSSION

CMLE and MLE from *P. putida* catalyze what superficially appear to be identical reactions in a highly specialized catabolic pathway. This simple recognition of a common reaction type, an intramolecular conjugate addition, had led to a general acceptance of this enzyme pair, as well as other pairs in the parallel branches of the pathway, as a paradigm for the evolution of metabolic diversity (Clarke, 1984) from ancestral enzymes. The underlying assumption of this hypothesis is that the array of active site functional groups required for catalysis has been conserved for both enzymes and that the major distingushing feature between the pair has been the acquisition of substrate specificity. This specificity would be achieved, then, by the enhanced binding of substrate functional groups that are peripheral to the structure and energetics of the transition state of the reaction. By this reasoning the 3-carboxyl group of 3CM is relegated to a less important role in substituent recognition by CMLE.

We have reported a detailed analysis of the absolute stereochemical course of CMLE and MLE (Chari et al., 1987a). A conservation of active site functional groups and geometry would necessarily lead to the prediction that the stereochemical and mechanistic features of these reactions are conserved and are identical. Our findings clearly demonstrated that they are not identical. In fact, the results were consistent with two distinct classes of cycloisomerases: the syn class represented by MLE and the anti class typified by CMLE (Figure 7). It is particularly significant that the geometric difference was not the result of an inversion of the configuration of proton delivery by a protonated active site base but was due to the opposing relative configurations of intramolecular carboxylate attack. Thus, the enzymes must recognize opposite mirror-image, out-of-plane conformers of their respective substrates. This marked difference in molecular recognition suggested a remote evolutionary relationship for MLE and CMLE.

While the geneology of MLE seemed obscure to us at the time, we proposed that CMLE bore a striking resemblance to the fumarate-dependent enzymes (Figure 8; Chari et al., 1987a). Hanson and Rose (1975) had suggested that enzyme-catalyzed 1,2-eliminations that form a trans or cis double bond between two carboxylate groups (i.e., the fumarate- and maleate-dependent enzymes) all occur by anti additions. In addition, they all follow the same absolute stereochemical course, which has been fixed by the common L stereochemistry of the substrates (Figure 8). We suggested, then, that the CMLE reaction might be viewed as an intramolecular example of a fumarate-dependent reaction. This hypothesis was further supported by the observation that CMLE from *Neurospora crassa* catalyzes the conversion of 3CM to the lactone with the opposite regiochemistry. This reaction occurs by a syn mechanism and with the same absolute stereochemistry as MLE (Kirby et al., 1975; Figure 7). This suggests that

```
CMLE    MTNQLFDAYFTAPAMREIFSDRGRLQGMLDFEAALARAEAAAGLVPHSAVAAIEAACKAERYDVGALANAIATAG   75
ADSL    -----MIERYSRPEMSAIWTDENRFQAWLEVEILACEAWAELGVIPKEDV-------KVMRENASFDINRILEIE   63

CMLE    NSAIPLVKALGKVIASGVPEAERYVHLGATSQDAMDTGLVLQLRDALDLIEADLGKLADTLSQQALKHADTPMVG  150
ADSL    KDTRHDVVAFTRAVSESLGEERKWVHYGLTSTDVVDTALSYLLKQANDILLKDLERFVDIIKEKAKEHKYTVMMG  138

CMLE    RTWLQHATPVTLGMKLAGVLGALTRHRQRLQELGPPCWCCSSGGASGSLAALGSKAMPVAEALAEQLKL-SLPEQ  224
ADSL    RTHGVHAEPTTFGLKLALWHEEMKRNLERFKQAKAGIEVGKISGAVGTYANIDP---FVEQYVCEKLGLKAAPIS  210

CMLE    PWHTQRDRLVEFASVLGLVAGSLGKFGRDVSLLMQTEAGEVFEPSAPGKGGSSTMPHKRNPVGAAVLIGAATRVP  299
ADSL    TQTLQRDRHADYMATLALIATSIEKFAVEIRGLQKSETREVEEFFAKGQKGSSAMPHKRNPIGSENMTGMARVIR  285

CMLE    GLVSTLFAAMPQEHERSLGLWHAEWETLPDICCLVSGALRQAQVIAEGIEVDAARMRRNLDLTQGLVLAEAVSIV  374
ADSL    GYMMTAYENVPLWHERDISHSSAERIILPDATIALNYMLNRFSNIVKNLTVFPENMKRNMDRTLGLIYSQRVLLA  360

CMLE    LARTPGPR--PCPPLAGAMLPASRGRTAAPACGAG----------------------------------       407
ADSL    LIDTGLTREEAYDTVQPKAMEAWEKQVPFRELVEAEEKITSRLSPEKIADCFDYNYHLKNVDLIFERLGLA       431
```

FIGURE 5: Pairwise alignment of CMLE and adenylosuccinate lyase (ADSL). Amino acid identities and sequence gaps are represented by (|) and (–), respectively. The shaded regions correspond to the consensus regions for class II fumarases shown in Figure 6.

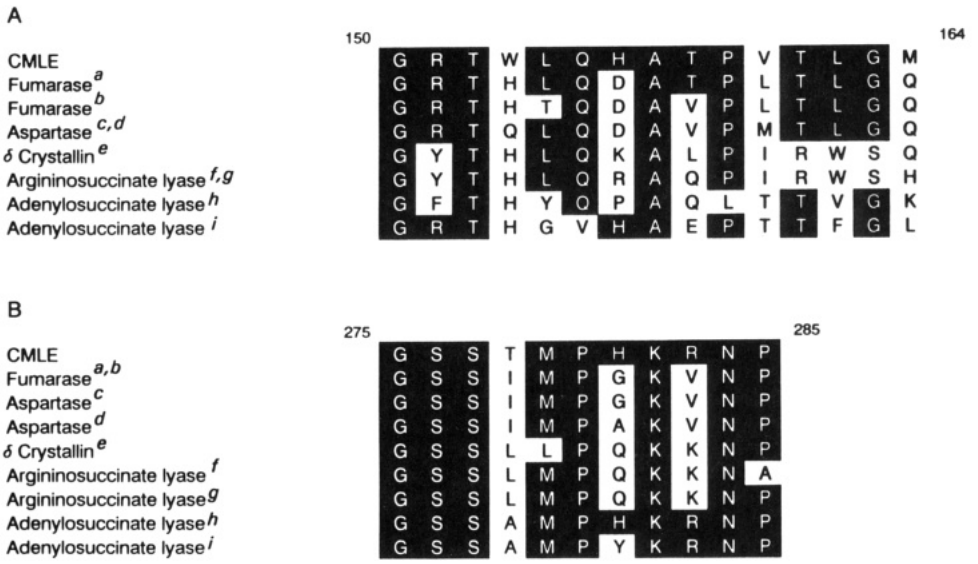


FIGURE 6: Amino acid sequence alignment of CMLE with two consensus sequences (A and B) found in the class II fumarases and related enzymes. The sequence GSxxMxxKxN (B) has also been proposed to relate the class I and class II fumarases (Woods et al., 1988). Sequences are [a]from *B. subtilis*, *E. coli*, and *S. cerevisiae*; [b]from human and pig; [c]from *P. fluorescens*; [d]from chicken and *E. coli*; [e]from chicken; [f]from *S. cerevisiae*; [g]from algae, rat, and human; [h]from *B. subtilis*; and [i]from chicken.

additions to double bonds bearing a single carboxylate group are fundamentally different from those to double bonds bearing two carboxyl groups. Recent studies have offered mechanistic explanations for this distinction (Benner et al., 1989; Gerlt & Gassman, 1992).

Recent work has revealed that MLE has a remarkably strong sequence and structural homology to mandelate racemase, which catalyzes a very different reaction (Neidhart et al., 1990; Tsou et al., 1990). The common mechanistic feature of the two enzymes may well be the stabilization of a carbanionic intermediate adjacent to a carboxyl group. The α,β-barrel motif that is shared by these enzymes has been implicated in a number of enzymes which stabilize carbanionic intermediates (Farber & Petsko, 1990).

Our present findings fully support our hypothesis that CMLE is a member of the class II fumarase family. CMLE is, then, the first member of this class that catalyzes an intramolecular variant of a fumarate reaction. The high homology between CMLE and ADSL suggested that CMLE might have ADSL activity. We have found that adenylosuccinate is not a substrate for CMLE ($<10^{-6}$ the rate of 3CM turnover and $<10^{-5}$ the rate of adenylosuccinate turnover by yeast ADSL). Conversely, 3CM is a poor substrate for ADSL. In addition, malate is not a substrate for CMLE, and it is a poor inhibitor of the CMLE reaction ($K_I \sim 2$ mM; data not shown). These data suggest that significant differences in active site structure must exist between CMLE and other family members. Crystallographic studies are currently underway with the expressed protein. To date, no X-ray structures have been reported for any fumarate-dependent enzyme.

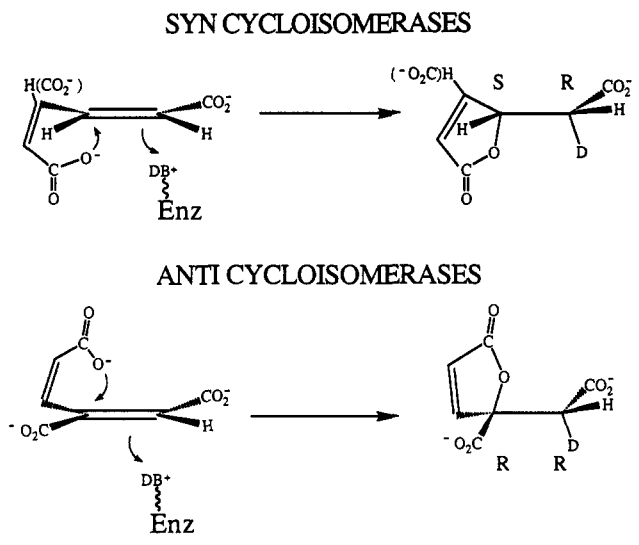


FIGURE 7: Absolute stereochemical course of MLE (*syn*-cycloisomerase including the carboxymuconate lactonizing enzyme from *N. crassa*) and of CMLE (*anti*-cycloisomerase) from *P. putida* (Chari et al., 1987a).

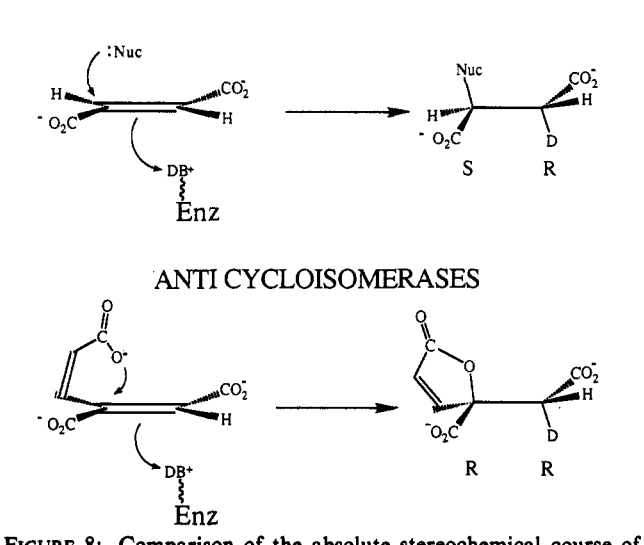


FIGURE 8: Comparison of the absolute stereochemical course of CMLE from *P. putida* to that of the class II fumarases and related enzymes.

In conclusion, MLE and CMLE represent a striking example of disparate enzyme recruitment based on fundamentally different mechanistic motifs to catalyze essentially identical reactions. The overall similarity of the catechol and protocatechuate branches of the pathway serves to reinforce the illusion of relatedness between the pair. However, the true lineages of MLE and CMLE reveal a more subtle and exciting evolutionary mechanism and are a reminder of our limited understanding of the origins of molecular recognition in enzyme catalysis.

ACKNOWLEDGMENT

We thank Professors John A. Gerlt and George L. Kenyon for stimulating discussions and Professor L. Nicholas Ornston for the gift of pPX81. We also thank Ms. Leslie Taylor of the Computer Graphics Lab at UCSF for assistance with sequence analysis software.

REFERENCES

Aimi, J., Badylak, J., Williams, J., Chen, Z. D., Zalkin, H., & Dixon, J. E. (1990) *J. Biol. Chem. 265*, 9011–9014.

Ainsworth, A. T., & Kirby, G. W. (1968) *J. Chem. Soc. C*, 1483–1487.

Altschul, S. F., & Erickson, B. W. B. (1986) *Bull. Math. Biol. 48*, 603–616.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403–410.

Beacham, I. R., Schweitzer, B. W., Warrick, H. M., & Carbon, J. (1984) *Gene 29*, 271–279.

Benner, S. A., Glasfeld, A., & Piccirilli, J. A. (1989) *Top. Stereochem. 19*, 127–207.

Bilofsky, H. S., Burks, C., Fickett, J. W., Goad, W. B., Lewitter, F. I., Rindone, W. P., Swindell, C. D., & Tung, C.-S. (1986) *Nucleic Acids Res. 14*, 1–5.

Chari, R. V. J., Whitman, C. P., Kozarich, J. W., Ngai, K.-L., & Ornston, L. N. (1987a) *J. Am. Chem. Soc. 109*, 5514–5519.

Chari, R. V. J., Whitman, C. P., Kozarich, J. W., Ngai, K.-L., & Ornston, L. N. (1987b) *J. Am. Chem. Soc. 109*, 5520–5521.

Clarke, P. H. (1984) in *Microbial Degradation of Organic Compounds* (Gibson, D. T., Ed.) Vol. 13, pp 11–27, Dekker, New York.

Debuchy, R., Purton, S., & Rochaix, J. D. (1989) *EMBO J. 8*, 2803–2809.

Ebbole, D. J., & Zalkin, H. (1987) *J. Biol. Chem. 262*, 8274–8287.

Farber, G. K., & Petsko, G. A. (1990) *Trends Biochem. Sci. 15*, 228–234.

George, D. G., Barker, W. C., & Hunt, L. T. (1986) *Nucleic Acids Res. 14*, 11–15.

Gerlt, J. A., & Gassman, P. G. (1992) *J. Am. Chem. Soc. 114*, 5928–5934.

Gottesman, M., Adhya, S., & Das, A. (1980) *J. Mol. Biol. 140*, 57–75.

Gross, M., Sweet, R., Sathe, G., Yokoyama, S., Fasano, O., Goldfarb, M., Wigler, M., & Rosenberg, M. (1985) *Mol. Cell. Biol. 5*, 1015–1024.

Hanson, K. R., & Rose, I. A. (1975) *Acc. Chem. Res. 8*, 1–10.

Hughes, E. J., Shapiro, M. K., Houghton, J. E., & Ornston, L. N. (1988) *J. Gen. Microbiol. 134*, 2877–2887.

Kinsella, B. T., & Donnan, S. (1986) *Biosci. Rep.* 921–929.

Kirby, G. W., O'Laughlin, G. J., & Robins, D. J. (1975) *J. Chem. Soc., Chem. Commun.*, 402–403.

Lawrence, C. B., & Goldman, D. A. (1988) *Comp. Appl. Biosci. 4*, 25–33.

Matsubasa, T., Takiguchi, M., Amaya, Y., Matsuda, I., & Mori, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 592–596.

Matuo, S., Tatsuno, M., Kobayashi, K., Saheki, T., Miyata, T., Iwanga, S., Amaya, Y., & Mori, M. (1988) *FEBS Lett. 234*, 395–399.

Miles, J. S., & Guest, J. R. (1985) *Nucleic Acids Res. 13*, 131–140.

Mori, M., Matsubasa, T., Amaya, Y., & Takiguchi, M. (1990) *Prog. Clin. Biol. Res. 344*, 683–699.

Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., & Petsko, G. A. (1990) *Nature 347*, 692–694.

Neidle, E. L., Hartnett, C., Bonitz, S., & Ornston, L. N. (1988) *J. Bacteriol. 170*, 4874–4880.

Nickerson, J. M., Wawrousek, E. R., Hawkins, J. W., Wakil, A. S., Wistow, G. J., Thomas, B. L., & Piatigorsky, J. (1985) *J. Biol. Chem. 260*, 9100–9105.

Nickerson, J. M., Wawrousek, E. F., Borras, T., Hawkins, J. W., Norman, B. L., Filpula, D. R., Nagle, J. W., Ally, A. H., & Piatigorsky, J. (1986) *J. Biol. Chem. 261*, 552–557.

O'Brien, W. E., McInnes, R., Kalumuck, K., & Adcock, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7211–7215.

Ornston, L. N. (1966) *J. Biol. Chem. 241*, 3787–3794.

Ornston, L. N., & Yeh, W.-K. (1982) in *Biodegradation and Detoxification of Environmental Pollutants* (Chakrabarty, A. M., Ed.) pp 105–126, CRC Press, Miami, FL.

Patel, R. N., Meagher, R. B., & Ornston, L. N. (1973) *Biochemistry 12*, 3531–3537.

Pearson, W. R. (1990) *Methods Enzymol. 183*, 63–98.

Pearson, W. R., & Lipman, D. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 2444–2448.

Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., Borras, T., Nickerson, J. M., & Wawrousek, E. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3479–3483.

Sacchettini, J. C., Frazier, M. W., Chiara, D. C., Banaszak, L. J., & Grant, G. A. (1988) *Biochem. Biophys. Res. Commun. 153*, 435–440.

Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342–1346.

Smith, R. F., & Smith, T. F. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 118–122.

Smith, R. F., & Smith, T. F. (1992) *Protein Eng. 5*, 35–41.

Stanier, R. Y., & Ornston, L. N. (1973) in *Advances in Microbial Physiology* (Rose, A., & Tempest, D., Eds.) pp 86–149, Academic Press, London.

Studier, F., & Moffatt, B. (1986) *J. Mol. Biol. 189*, 113–130.

Studier, F., Rosenberg, A., Dunn, J., & Dubendorff, J. (1990) *Methods Enzymol. 185*, 60–89.

Takagi, J. S., Ida, N., Tokushige, M., Sakamoto, H., & Shimura, Y. (1985) *Nucleic Acids Res. 13*, 2063–2074.

Takagi, J. S., Tokushige, M., & Shimura, Y. (1986) *J. Biochem. (Tokyo) 100*, 697–705.

Takiguchi, M., Matsubasa, T., Amaya, Y., & Mori, M. (1989) *Bioessays 10*, 163–166.

Teipel, J. W., Hass, G. M., & Hill, R. L. (1968) *J. Biol. Chem. 243*, 5684–5694.

Tsou, A. Y., Ransom, S. R., Gerlt, J. A., Buechter, D. D., Babbitt, P. C., & Kenyon, G. L. (1990) *Biochemistry 29*, 9856–9862.

Woods, S. A., Miles, J. S., Roberts, R. E., & Guest, J. R. (1986) *Biochem. J. 237*, 545–557.

Woods, S. A., Schwartzbach, S. D., & Guest, J. R. (1988) *Biochim. Biophys. Acta 954*, 14–26.

Wu, M., & Tzagoloff, A. (1987) *J. Biol. Chem. 262*, 12275–12282.

Yanisch-Perron, C., Viera, J., & Messing, J. (1985) *Gene 33*, 103–119.

Yeh, L. S., Elzanowski, A., Hunt, L. T., & Barker, W. C. (1988) *Comp. Biochem. Physiol. B89*, 433–437.