Crystal Structures of 2-Methylisocitrate Lyase in Complex with Product and with Isocitrate Inhibitor Provide Insight into Lyase Substrate Specificity, Catalysis and Evolution[†,‡]

Sijiu Liu,[§] Zhibing Lu,[||] Yin Han,[||] Eugene Melamud,[§] Debra Dunaway-Mariano,[||] and Osnat Herzberg*[,§]

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland 20850, and Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131

Received September 20, 2004; Revised Manuscript Received November 1, 2004

ABSTRACT: Two crystal structures of the C123S mutant of 2-methylisocitrate lyase have been determined, one with the bound reaction products, Mg^{2+}-pyruvate and succinate, and the second with a bound Mg^{2+}-(2*R*,3*S*)-isocitrate inhibitor. Comparison with the structure of the wild-type enzyme in the unbound state reveals that the enzyme undergoes a conformational transition that sequesters the ligand from solvent, as previously observed for two other enzyme superfamily members, isocitrate lyase and phosphoenolpyruvate mutase. The binding modes reveal the determinants of substrate specificity and stereoselectivity, and the stringent specificity is verified in solution using various potential substrates. A model of bound 2-methylisocitrate has been developed based on the experimentally determined structures. We propose a catalytic mechanism involving an α-carboxy-carbanion intermediate/transition state, which is consistent with previous stereochemical experiments showing inversion of configuration at the C(3) of 2-methylisocitrate. Structure-based sequence analysis and phylogenic tree construction reveal determinants of substrate specificity, highlight nodes of divergence of families, and predict enzyme families with new functions.

In this paper we report the structure of *Escherichia coli* (Cys123Ser) 2-methylisocitrate lyase (MICL)[1] mutant bound with Mg^{2+} cofactor and with the product ligands, succinate and pyruvate, or with the competitive inhibitor (2*R*,3*S*)-isocitrate. In both structures the active site loop (residues 117−132), which contributes several essential residues to the catalytic site, assumes the catalytically active, closed conformation. Previous papers have reported the structures of the loop-open conformation of the wild-type *E. coli* MICL bound with Mg^{2+} (*1*) or the wild-type *Salmonella typhimurium* MICL complexed with Mg^{2+} and pyruvate (*2*).

MICL functions in the methylcitrate cycle for α-oxidation of propionate (*3*, *4*), where it catalyzes the thermodynamically favored C−C bond cleavage reaction of (2*R*,3*S*)-2-methylisocitrate to pyruvate and succinate (Figure 1) (*5*, *6*).

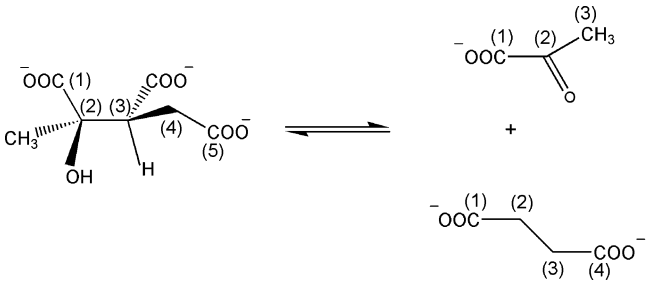

FIGURE 1: The chemical reaction catalyzed by MICL.

MICL is similar in structure to the enzyme isocitrate lyase (ICL), which catalyzes the analogous cleavage reaction of (2*R*,3*S*)-isocitrate to glyoxylate and succinate in the glyoxalate pathway (for review see (*7*)). Several crystal structures of the ICL have been reported, which show the catalytically active, loop-closed conformation as well as the loop-open conformation (*8*−*10*). Chemical modification (*11*, *12*) and/or mutagenesis studies (*13*−*17*) have been carried out on both ICL and MICL to identify active site residues essential to catalysis.

MICL and ICL adopt an α/β-barrel fold, and are members of the isocitrate lyase superfamily. PSI−BLAST (*18*) on the nonredundant protein sequence database (including all nonredundant GenBank CDS translations, RefSeq proteins, PDB, SwissProt, PIR, and PRF) identifies over 587 unique sequence homologues . A distinguishing structural characteristic of the superfamily is the exchange of the eighth α-helix between the barrels of two adjacent subunits (*19*), which has been observed in all X-ray structures except for one member (from *E. coli*) of the most divergent family,

[†] Grant sponsors: NSF Grant MCB9813271 and NIH Grant P01GM57890 (O.H.) and NIH Grant RO1GM36260 (D.D.-M.). Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Energy Research, under Contract No. W-31-109-ENG-38.

[‡] The coordinates have been deposited in the Protein Data Bank (entry codes 1oqf, 1xg3, 1xg4).

* Corresponding author. Mailing address: Center for Advanced Research in Biotechnology, 9600 Gudelsky Drive, Rockville MD 20850. Tel: 301-738-6245. Fax: 301-738-6255. E-mail: osnat@carb.nist.gov.

[§] University of Maryland Biotechnology Institute.

[||] University of New Mexico.

[1] Abbreviations: PEP, phosphoenolpyruvate; MICL, 2-methylisocitrate lyase; ICL, isocitrate lyase; OXAHYD, oxaloacetate acetylhydrolase; HEPES, *N*-(2-hydroxyethyl)-piperazine-*N*′-ethanesulfonic acid; IPTG, isopropyl-β-D-thiogalactopyranoside; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; DEAE, diethylaminoethyl.

10.1021/bi0479712 CCC: $30.25
Published on Web 02/01/2005

ketopantoate hydroxymethyl transferase ((*20, 21*), and the enzyme from *Neisseria meningitides*, PDB entry code 1O66). Known members of the superfamily catalyze an interesting variety of C−C bond and P−C bond forming or cleaving reactions, which proceed via α-carboxy-oxyanion intermediates and/or transition states (*22*). The families of enzymes acting on C−C bonds include ketopantoate hydroxymethyl transferase of the coenzyme A biosynthetic pathway (*20, 21, 23, 24*) (C−C bond formation), the α-hydroxyacid lyases, MICL and ICL (C−C bond cleavage), and oxaloacetate acetylhydrolase (*25*) (C−C bond hydrolysis). The families of enzymes catalyzing reactions of P−C bonds includes PEP mutase and carboxyPEP mutase (*26*), which function in phosphonate biosynthesis, and phosphonopyruvate hydrolase (*27*), an enzyme which functions in phosphonate degradation. Despite the divergence in the chemistry that it catalyzes, the PEP mutase active site conserves the same catalytic scaffold observed in MICL and ICL, including the flexible active site loop, which regulates solvent access (*19, 28, 29*).

Our laboratories employ a structure-mechanism based approach to identify sequence markers within the conserved catalytic scaffold of the ICL/PEP mutase superfamily for function assignment to uncharacterized family sequences. Because α-hydroxyacids are common metabolites, we anticipate the discovery of C−C bond lyases that assist in novel carbon metabolite degradation and recycling pathways. The focus of the present work was to define the structure of the MICL active site in a catalytically active conformation, and to reveal the residues functioning in substrate recognition and catalysis. In this manner, scaffold residues unique to the C−C lyases can be identified as well as scaffold residues responsible for their divergence in substrate specificity. We propose a model for the shared catalytic mechanism of MICL and ICL, identify active site structural features important to substrate discrimination, and use phylogeny inference to suggest sequences that may possess novel catalytic activities.

MATERIALS AND METHODS

E. coli MICL and ICL Gene Cloning, Overexpression, and Purification. Preparation of His-tagged recombinant *E. coli* MICL was previously reported by Brock et al (*6*), and the preparation of recombinant *E. coli* ICL was reported by Hoyt et al (*30, 31*). In this work, the MICL and ICL encoding genes were cloned without His-tag so that the kinetics are as close to the native enzymes as possible. A PCR-based strategy where *PfuTurbo* DNA polymerase (Stratagene) and synthetic oligonucleotide primers (Invitrogen) encoding *Nde*I and *Bam*HI restriction sites were used to amplify the targeted genes from the *E. coli* K-12 genomic template (ATTCC 700926). In the case of the ICL gene, an internal *Nde*I restriction site was mutated without change to the encoded amino acid. The PCR products were gel purified, digested with *Nde*I and *Bam*HI restriction enzymes (Invitrogen) and ligated to linearized pET-3c vectors (Novagen). The ligation products were used to transform *E. coli* BL21(DE3) plys competent cells (Novagen). Gene sequences were confirmed by DNA sequencing carried out by the DNA Sequencing Facility of the University of New Mexico. For protein expression, transformed cells were grown at 28 °C in Luria broth (LB) containing 50 μg/mL carbenicillin for 10 h to an $OD_{600\ nm} \approx 1$, and then induced for 5 h at 30 °C with 0.4 mM IPTG. Cells were harvested by centrifugation (6500 rpm for 15 min at 4 °C), and the cell pellets were suspended in 100 mL of ice-cold lysis buffer consisting of 50 mM K^+ HEPES (pH 7.5), 1 mM EDTA, 1 mM benzamide hydrochloride, 0.05 mg/mL trypsin inhibitor, 1 mM 1, 10-phenanthroline, 0.1 mM PMSF and 5mM DTT. The suspensions were passed twice through a French Press at 1200 psi, and the cell lysates were centrifuged at 4 °C for 45 min at 20 000 rpm. The supernatants were fractionated by ammonium sulfate induced protein precipitation. The protein pellets from the 35−70% saturated ammonium sulfate mixtures were separately dissolved in 50 mL buffer A (50 mM triethanolamine (pH 7.5), 5 mM $MgCl_2$ and 0.5 mM DTT) and dialyzed for 6 h at 4 °C against buffer A before loading onto a 3.5 × 100 cm DEAE-cellulose columns equilibrated with buffer A at 4 °C. The columns were washed with 1 L of buffer A, and then eluted with a 2 L linear gradient of 0 to 0.5 M KCl in buffer A. The column fractions were analyzed by measuring the absorbance at 280 nm and by carrying out SDS−PAGE analysis. The ICL eluted at 0.30 M KCl while the MICL eluted at 0.27 M KCl. Ammonium sulfate was added to 20% saturation. The resulting solutions were loaded onto a 2.5 cm × 50 cm Phenyl Sepharose columns equilibrated at 4 °C with 20% saturated ammonium sulfate in buffer A. The columns were washed with 200 mL of 20% saturated ammonium sulfate in buffer A and then eluted with a 800 mL linear gradient of 20% to 0% saturated ammonium sulfate in buffer A. The column fractions were analyzed by measuring the absorbance at 280 nm and by carrying out SDS−PAGE analysis. The ICL eluted at 13% and the MICL at 4% saturated ammonium sulfate. The MICL fractions were combined and concentrated at 4 °C to 31 mg/mL using an Amicon concentrator, and stored at − 80 °C. The ICL fractions were combined, dialyzed at 4 °C against 1 L buffer A, and then loaded onto a 2.5 cm × 30 cm DEAE Sepharose column equilibrated at 4 °C with buffer A. The column was washed with 100 mL buffer A and then eluted with a 600 mL linear gradient of 0 to 0.5M KCl in Buffer A. The ICL eluted at 0.32 M KCl. The fractions were combined and concentrated at 4 °C to 62 mg/mL using an Amicon concentrator and then stored at − 80 °C. The ICL (yield: 37 mg/gm wet cell) and MICL (yield 8 mg/gm wet cell) were shown to be homogeneous by SDS−PAGE analysis. Electrospray mass spectrometry of the recombinant MICL was carried out by the University of New Mexico Mass Spectroscopy Laboratory.

E. coli C123S MICL mutant was prepared by PCR using high fidelity *pfu* DNA polymerase, synthetic primers and the wild-type MICL gene in pET-3c (described above) as template. The gene sequence was verified by DNA sequencing. The mutant gene was expressed and the mutant enzyme purified in the same manner described above for the wild-type MICL.

Substrates and Inhibitors. Threo-D_S-isocitrate (i.e. (2*R*,3*S*)-isocitrate), (*S*)- and (*R*)-citramalate (i.e., (*S*)- and (*R*)-2-methylmalate), (*S*)- and (*R*)-malate, and 3-nitroproprionate were purchased from Aldrich. *Threo* (i.e., racemic mixture of (2*R*,3*S*)- and (2*S*,3*R*)-) 2-methylisocitrate and *erythro* (i.e., racemic mixture of (2*S*,3*S*)- and (2*R*,3*R*)-) 2-methylisocitrate were synthesized according to published procedures (*6*).

Enzyme Assays and Steady-State Kinetics. Formation of glyoxylate and pyruvate products in the lyase reaction mixtures was monitored spectrophotometrically ($\Delta\epsilon$ at 340

Table 1: Data-Processing Statistics

	native	Hg	Xe	C123S/isocitrate	C123S/(pyruvate + succinate)
space group	$P3_221$	$P3_221$	$P3_221$	$C2$	$C2$
cell dimensions (Å)					
a	83.0	83.0	82.8	159.1	159.4
b	83.0	83.0	82.8	84.6	86.0
c	166.9	166.6	167.0	99.8	99.6
γ or β (deg)	120.0	120.0	120.0	108.1	108.3
no. of molecules in the asymmetric unit	2	2	2	4	4
resolution range (Å)	50.0−1.86	50.0−1.86	50.0−2.30	50.0−1.55	50.0−1.90
highest resolution shell (Å)	1.93−1.86	1.93−1.86	2.38−2.30	1.61−1.55	1.97−1.90
no. of independent reflections [a]	52 853	53 985	29 927	161 348	89 039
completeness (%) [b]	93.1 (68.2)	95.3 (82.6)	99.0 (98.1)	88.8 (22.7)	88.5 (67.5)
redundancy	4.35	5.65	7.12	4.16	2.82
R_{merge} [b,c]	0.056 (0.290)	0.058 (0.284)	0.065 (0.143)	0.030 (0.186)	0.025 (0.084)
$\langle I/\delta(I)\rangle$	12.8	14.2	18.1	15.9	20.3

[a] The reflections for which $F_{obs} > 1.0\delta(F_{obs})$. [b] The values in parentheses correspond to the highest resolution shell. [c] $R_{merge} = \sum_h\sum_i|I(h)_i - \langle I(h)\rangle|/\sum_h\sum_i I(h)_i$.

nm is 6200 M^{-1} cm^{-1}) using the NADH/lactate dehydrogenase coupling system. Reaction solutions (1 mL) contained 1 mM D_S-isocitrate, *threo* 2-methylisocitrate, or *erythro* 2-methylisocitrate, (*S*)-2-methylmalate, (*R*)-2-methylmalate, (*R*)-malate, or (*S*)-malate, 5 mM $MgCl_2$, 10 units/mL lactate dehydrogenase and 0.2 mM NADH in 50 mM K^+Hepes (pH 7.5, 25 °C). Reactions were initiated by adding the ICL or MICL. High lyase concentration (20 μM) and an extended incubation period (1 h) were employed when testing activity toward the alternate substrates. The steady-state maximum velocity (V_{max}) and Michaelis constant (K_m) were determined from initial velocity (V_o) data (measured as a function of substrate concentration ([S]) varied in the range of 0.5 to 10-fold K_m) fitted with eq 1 using the computer program

$$V_o = V_{max}[S]/(K_m + [S]) \quad (1)$$

KinetAsyst 1 (Hi-Tech Scientific). The turnover-over rate (k_{cat}) was calculated from the equation $k_{cat} = V_{max}/[E]$ where the enzyme concentration, [E], was calculated using the Bio-Rad protein Microassay Kit and the enzyme molecular weight. The competitive inhibition constants (K_i) were determined by measuring the initial velocities of the MICL catalyzed reactions as a function of substrate concentration at fixed concentrations of inhibitor ([I]), then fitting the data with eq 2.

$$V_o = V_{max}[S]/(K_m(1 + ([I]/K_i) + [S])) \quad (2)$$

Crystallization and Data Collection. Crystals of wild-type MICL were obtained at room temperature by vapor diffusion in hanging drops. The protein solution contained 10 mg/mL protein, 5 mM $MgCl_2$, 5 mM pyruvate, and 5 mM succinate. The reservoir solutions contained 1.6 M ammonium sulfate and 100 mM sodium HEPES buffer (pH 6.8−7.8). The hanging drops consisted of equal volumes of protein and reservoir solutions. The crystals belong to space group $P3_221$ (Table 1). Despite the presence of products in the crystallization solution, this is the same crystal form as that reported by Grimm and colleagues (*1*) with unbound ligands. Although we determined the structure independently by Multiples Isomorphous Replacement (MIR) the details of the structure are already available and the reader is referred to the earlier publication (*1*). In this report, we use the wild-type structure determined in our laboratory for comparisons with the structures of the complexes.

Crystals of the mutant C123S MICL in complex with isocitrate were obtained at room temperature by vapor diffusion in hanging drops. The reservoir solution contained 30% poly(ethylene glycol) (PEG) 3350, 0.1 M ammonium sulfate, and 0.1 M MES (pH 6.0). The drops consisted of protein at a concentration of 11 mg/mL, 5 mM $MgCl_2$, and 5 mM *threo*-D_S-isocitrate, diluted by an equal volume of reservoir solution. The crystals belong to space group C2 (Table 1).

A crystal of C123S MICL in complex with pyruvate and succinate was obtained by overnight incubation of a C123S MICL/isocitrate complex crystal in a solution containing 35% PEG3350, 100 mM pyruvate and 100 mM succinate. These crystal also belongs to space group C2 (Table 1).

For X-ray data collection of the wild-type apo MICL, crystals were incubated in a solution of 90% saturated Li_2SO_4 for 30 s, and flash-cooled in liquid propane cooled by liquid nitrogen. The protein complex crystals were flash-cooled in their original PEG solutions used for crystallization. X-ray data were collected at 100 K. For the wild-type enzyme structure determination, Cu Kα X-rays were supplied by a Siemens rotating anode generator equipped with Harvard focusing mirrors. Diffraction intensity data were recorded on a MAR345 imaging plate detector. For the S123C mutant enzyme bound to products, X-ray data were collected on a Rigaku RU-300 rotating anode generator (Rigaku/MSC) equipped with focusing mirrors (MSC/Yale) and an R-AXIS IV++ image plate detector. X-ray data of the C123S MICL/isocitrate crystal were collected at the Advanced Photon Source on the SBC−CAT 19-BM beamline equipped with a MAR CCD detector (Argonne National Laboratory). The data were processed using the program HKL (*32*). Statistics of the data processing are provided in Table 1.

Structure Determination and Refinement. The structure of the wild-type enzyme was determined by MIR using a mercury derivative (2 mM methyl mercury acetate soaked for 2 h) and a Xe derivative (20 s incubation in high-pressure Xe chamber). Heavy atom sites were determined using the program SOLVE, version 2.05 (*33*). The phases were improved by solvent flattening and a partial model was built at 3 Å resolution with RESOLVE (*34*, *35*). The remaining model was built manually on an interactive graphics work-

Table 2: Refinement Statistics

	wild-type	C123S/isocitrate	C123S/(pyruvate + succinate)
space group	$P3_221$	$C2$	$C2$
resolution ranges (Å)	50−1.93	50−1.6	50−1.9
completeness (%)	91.9	83.9	88.3
$F \geq 2\sigma(F)$			
R_{work}[a]	0.191	0.181	0.161
R_{free}[b]	0.247	0.200	0.195
RMSD from ideal geometry			
bond length (Å)	0.011	0.007	0.007
bond angle (deg)	1.5	1.3	1.4
average B factors (Å^2)/no. of atoms			
whole	29.4/5059	23.2/9718	19.9/9815
protein	28.1/4411	22.2/8694	18.9/8702
solvent	38.4/648	32.2/982	28.1/1061
substrate/others		26.0/39	22.5/48
Mg^{2+}		26.6/3	15.9/4

[a] $R_{work} = \sum_h |F(h)_{calcd} - F(h)_{obsd}| / \sum_h F(h)_{obsd}$, where $F(h)_{calcd}$ and $F(h)_{obsd}$ are the refined observed and calculated structure factors, respectively. [b] R_{free} is calculated for a randomly selected 5.6% (wild-type crystal data), 3.0% (C123S MICL/isocitrate), and 3.6% (C123S MICL/(pyruvate + succinate) of the reflections that were omitted from refinement.

station using the structure of PEP mutase to guide the building (*19*, PDB entry code 1pym) and employing the program O (*36*).

The structure was refined at 2.5 Å with the program CNS (*37*), using cycles of simulated annealing at 3000 or 2500 K, followed by positional and temperature factor refinement. When the protein model was complete, the resolution of the data was increased to 1.93 Å. The two MICL subunits of the asymmetric unit were refined independently, and bulk solvent correction was included. Water molecules were added gradually as the refinement progressed. They were assigned in the $|F_o| - |F_c|$ difference Fourier maps (where F_o and F_c are the observed and calculated structure factors, respectively) with a 3σ cutoff level for inclusion in the model. The refinement statistics are shown in Table 2.

The structure of the C123S MICL/isocitrate complex was determined by Molecular Replacement using the program AmoRe (*38*). The wild-type structure determined in our laboratory (PDB entry code 1oqf) served as the search model (except that the solvent molecules were omitted). The resulting difference Fourier map indicated some alternative tracing, which was incorporated into the model. The map revealed density for the bound Mg^{2+}-isocitrate in three of the four MICL subunits in the asymmetric unit (A, B and D in the coordinates deposited in the PDB). The structure was refined at 1.6 Å with the program CNS (Table 2), following a protocol similar to that described above.

For the structure of the C123S MICL/product complex, the model of C123S MICL/isocitrate (omitting Mg^{2+}, isocitrate and solvent molecules) was used to calculate a difference Fourier map with the coefficients $2F_o - F_c$ and calculated phases. The map revealed density for the bound Mg^{2+}-pyruvate-succinate in subunits A, B and D, and only Mg^{2+}-pyruvate in molecule C. The structure was refined at 1.9 Å resolution with the program CNS (Table 2).

Sequence Analysis. Members of the ICL superfamily were identified by running 11 cycles of PSI−BLAST on the nonredundant protein database (*18*). The complete protein sequences of the PSI−BLAST list were retrieved from NCBI. Crystal structures of MICL, ICL, and PEP mutase were aligned using VAST (*39*). The resulting structure-based sequence alignment provided a seed alignment for multiple sequence alignment of superfamily members with experimentally confirmed function using ClustalW with default settings (*40*), and the resulting alignment was then used to align the entire superfamily sequence set. The multiple alignment was visualized with the program Belvu (written by E. L. Sonnhammer, Karolinska Institutet, Stockholm, Sweden), and incomplete sequences and obvious duplications were eliminated, yielding a set of 512 protein sequences. Manual editing to improve the automatic alignment was carried out in Jalview (*41*). Pairwise distances between each protein were calculated using PRODIST in PHYLIP (*42*) with the Jones, Taylor and Thornton model of amino acid change (*43*). These distances were used in the NEIGHBOR program of PHYLIP to construct a phylogeny tree. Bootstrapping was achieved by building 100 trees from shuffled alignments in SEQBOOT to obtain confidence limits (*44*)]. The tree was drawn using PhyloDraw (*45*).

RESULTS AND DISCUSSION

Protein Crystallization. The *E. coli* MICL gene was cloned for overexpression in *E. coli* BL21(DE3) cells. Mass spectral analysis of the purified protein provided a molecular weight of 32,003 Da, compared to the calculated weight of 32,003 Da for residues 2−229, thus showing that the N-terminal methionine residue was removed by posttranslational modification. The C123S MICL mutant was prepared by site directed mutagenesis. Trace contamination of native MICL produced by the *E. coli* cells used for overexpression was indicated by a small (<1%) but detectable level of activity in the purified mutant. The *Salmonella enterica* C123S MICL mutant prepared without contamination is inactive (*14*). Although crystals of the apo wild-type MICL were easily obtained from ammonium sulfate solution, it was the C123S MICL mutant that yielded good quality cocrystals with *threo*-D_S-isocitrate which could then be replaced by product as described in the methods section. Thus, the structures of three enzyme forms were determined and refined: wild-type MICL (at 1.93 Å resolution), C123S MICL bound with Mg^{2+} and *threo*-D_S-isocitrate (at 1.6 Å resolution), and C123S MICL bound with Mg^{2+}, pyruvate and succinate (at 1.9 Å resolution).

Overall Structures. For consistency with the amino acid numbering system used in the first MICL structure deter-

mination (*1*), the numbering system used here is based on the gene sequence including Met1. The wild-type MICL structure includes residues 2−291 of each of two subunits of the asymmetric unit. The root-mean-square deviation (RMSD) between α-carbon atom positions of the two subunits is 0.5 Å. As reported earlier (*1*), and similar to the structures of PEP mutase (*19*) and ICL (*9*, *10*), the biological unit is a tetramer (dimer of dimers) generated by the crystallographic 2-fold symmetry axis. MICL exhibits the α/β barrel fold with dimers exchanging the eighth helix (encompassing residues 238−260), the structural feature common to all but one (ketopantoate hydroxymethyl transferase) known members of the PEP mutase/ICL family. Despite the presence of Mg^{2+} in the crystallization solution, and the ability of MICL to bind Mg^{2+} under assay conditions (K_m for Mg^{2+} activation = 58 μM; (*6*)), no Mg^{2+} is evident in the electron density map. This is in contrast to the finding of Grimm and colleagues (*1*), who assigned an active site Mg^{2+} bound directly to Asp87 (instead of Asp85 as found in complexes). Another slight difference between the two structures is that in the structure determined by us, the peptide bond of Pro19 is cis, whereas it is a trans-peptide in the structure determined by Grimm and colleagues. Pro19 exhibits cis peptide also in the two complexes described below, and the associated electron density is very clear.

The MICL/isocitrate structure contains a tetramer in the asymmetric unit, and includes residues 2−288 of one subunit, residues 3−288 of the second subunit, residues 3−289 of the third subunit, and residues 3−285 of the fourth subunit. This tetramer constitutes the biological unit. Mg^{2+} ions and isocitrate molecules are observed only in the active sites of subunits A, B and D (chains are labeled as in the PDB entries). In subunits A, B, and D, the loop comprising residues 117−132 associates with the active site and adopts a conformation that sequesters the isocitrate from solvent (the closed conformation), while the same loop in subunit C, where Mg^{2+} and isocitrate are absent, adopts a conformation that exposes the active site (the open conformation). The electron density map associated with one of the active sites (molecule D) is shown in Figure 2.

The RMSD between α-carbon atom positions of subunits A and B in the asymmetric unit is 0.14 Å, and the corresponding RMSD excluding residues 117−130 between subunits C and D is 0.3 Å. The RMSD between α-carbon atom positions of subunits A and C excluding residues 117−130 is 0.3 Å, and between subunits A and D is 0.2 Å.

MICL/product structure contains the entire tetramer in the asymmetric unit, with residues 2−289 of one subunit, residues 3−288 of the second subunit, residues 3−289 of the third subunit, and residues 3−285 of the fourth subunit. The tetramer packing is the same as that of the MICL/isocitrate complex. Mg^{2+} ions and pyruvate were assigned in each of the four active sites, whereas succinate molecules were observed only in subunits A, B and D. The electron density map associated with one of the active sites (subunit D) is shown in Figure 2. The active site loops encompassing residues 117−132 of subunits A, B and D are in the closed conformation, and in subunit C this loop is in the open conformation.

For the MICL/product complex, the RMSD between α-carbon atom positions of subunits A and B in the asymmetric unit is 0.2 Å, and the corresponding RMSD

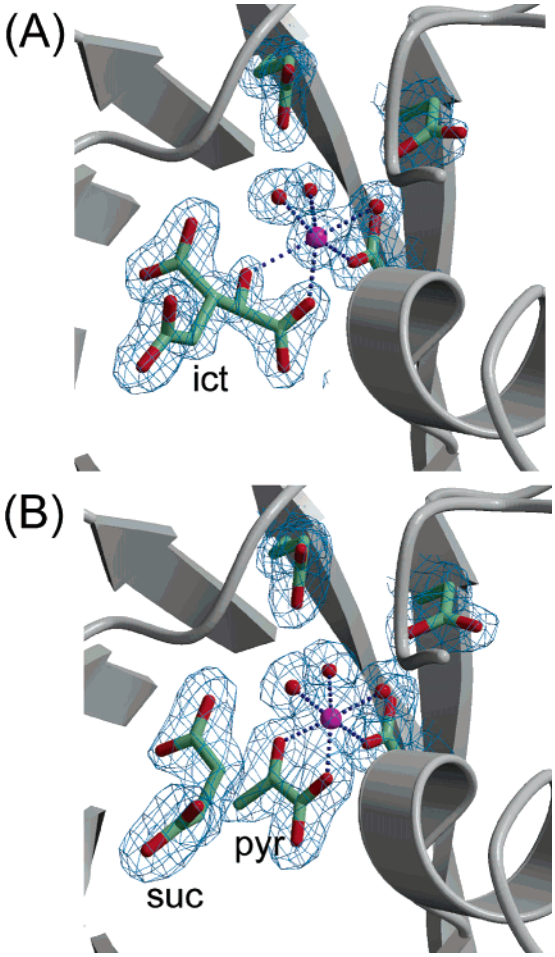

FIGURE 2: The electron density maps associated with ligand binding: (A) bound isocitrate; (B) bound product (pyruvate and succinate). Coordination of Mg^{2+} is also shown. The maps are contoured at 1.2σ level.

(excluding residues 117−130) between subunits C and D is 0.4 Å. The RMSD between α-carbon atom positions of subunits A and C (excluding residues 117−130) is 0.3 Å, and between subunits A and D is also 0.3 Å.

The RMSD between α-carbon atom positions of the MICL/isocitrate or MICL/product and unbound MICL structures (excluding residues 117−130 and 270−288) varies between 0.3 and 0.4 Å. The corresponding RMSD between MICL/isocitrate and MICL/product structures varies between 0.1 and 0.3 Å.

Conformational Flexibility of the Active Site Loop. As with PEP mutase (*19*, *28*) and ICL (*8*−*10*), MICL undergoes a mechanistically important, localized conformational transition upon inhibitor binding, which is preserved when the inhibitor is replaced by product. In the apo structure the loop comprising residues 117−132 adopts an "open" conformation that exposes the active site to bulk solvent (Figure 3). The binding of the substrate analogue, isocitrate, induces a "closed" loop conformation that sequesters the ligands from solvent. Subunit C of the C123S MICL complexes (unbound in the Mg^{2+}-isocitrate crystal, and bound with Mg^{2+}-pyruvate in the product-soaked crystal), displays a loop-open conformation (Figure 3), which differs from the loop-open conformation seen in the unliganded enzyme structure. Succinate binds on top of pyruvate, indicating that the kinetic mechanism is ordered, as it is in *E. coli* ICL, where succinate

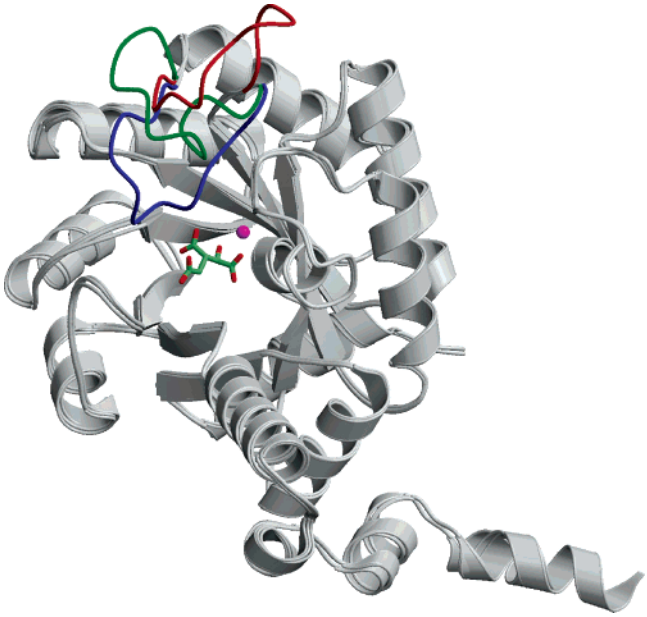

FIGURE 3: Three conformational states of MICL. The superpositioned molecules are depicted in gray color where the trace of the polypeptide chain is similar. The open conformation of the active site loop seen in the apo structure is highlighted in red. The open loop conformation seen in the MICL/isocitrate structure, where one of the tetramer molecules does not bind the inhibitor is colored green, and the closed loop conformation seen in the remaining three isocitrate-bound molecules is colored blue. Mg^{2+} is shown as a magenta sphere, and isocitrate is shown as a stick model with carbon atoms colored pale green and oxygen atoms colored red. The same two loop conformations are also observed in the MICL/product structure.

departs before glyoxylate (*46*). The synergy between succinate binding and loop closure is suggested by the observed correlation between the loop-closed structures in the Mg^{2+}-product or Mg^{2+}-isocitrate liganded enzyme, and the loop-open structures in the unliganded or Mg^{2+}-pyruvate bound enzyme. In the two lyases, as in the PEP mutase, loop closure partitions the active site from solvent while contributing substrate binding and/or catalytic residues. The loop position which is occupied by the substrate binding residue, Asn122, in PEP mutase is occupied by the catalytic cysteine residue in MICL and ICL.

Active Site. In the following discussion, 2-methylisocitrate, pyruvate and succinate atoms are numbered as shown in Figure 1.

The active site of MICL is located in a crevice at the center of the α/β barrel, close to the carboxy termini of the β-strands (Figure 3). The loop-closed, liganded MICL structures reported here offer further proof of a conserved catalytic scaffold associated with Mg^{2+} binding that serves both lyase and mutase families. Briefly, four carboxylic groups bind the Mg^{2+} by direct coordination (Asp85) or via bridging water ligands (Asp58, Asp87, Glu115) (Figure 4A and B). In the C123S MICL/ product complex (Figure 4B), pyruvate provides two additional coordinating oxygen atoms (C(1)O and C(2)O) to complete the almost perfect octahedral coordination geometry with average Mg^{2+}-ligand distance of 2.1 Å. In the C123S MICL/isocitrate complex (Figure 4A), the octahedral coordination is perturbed because the distance between Mg^{2+} and the hydroxyl group of isocitrate is 2.6−2.7 Å (depending on the molecule). Because isocitrate is not a substrate of MICL, this binding mode is not productive.

The closure of the gating loop (residues 117−132) shifts the ammonium group of Lys121 by approximately 13 Å, to within range for electrostatic interactions with two Mg^{2+} water ligands, and the carboxylate groups of Asp87 and Glu115. In PEP mutase, the same constellation of residues exists for interaction with Mg^{2+}, and the equivalent lysine residue is known to play an important role in catalysis by regulating loop closure (*28*). The loop-closed ICL structure indicates the same role for the ICL loop's lysine (*10*). Loop closure also positions Cys123 (represented by Ser123 in the mutant) in the active site, near the succinate C(2). The ICL catalytic cysteine (Cys191 in the *Mycobacterium tuberculosis* ICL is also replaced by a serine in the complex with glyoxylate and succinate) is similarly positioned by the loop closure, as is the corresponding substrate binding Asn122 in PEP mutase. In both closed loop MICL structures, the mutated residue, Ser123 forms an Asx turn (*47*) so that the hydroxyl group interacts with the backbone amide of His125. The same interaction is possible for a cysteine residue with a slight adjustment to increase the S−N distance. Such an interaction would reduce the pK_a of the cysteine, consistent with its role in proton transfer (see below). The Asx turn interaction is also present in ICL.

MICL forms additional interactions with pyruvate, analogous to those seen between ICL and glyoxylate, and between PEP mutase and oxalate or sulfopyruvate. The ligand carboxyl group is positioned close to the N-terminus of a short helix, where it interacts with the backbone NH groups of residues 46 and 47, and the hydroxyl group of Ser45. The ligand carbonyl group interacts with Arg158, an invariant residue in all ICL/PEP mutase superfamily members. Together, Mg^{2+} and Arg158 form an oxyanion hole for the carbonyl oxygen atom. The methyl group of pyruvate is located in a hydrophobic pocket formed by Phe186, Leu234, and Pro236. This pocket is unique to MICL and is not found in PEP mutase or in ICL. Another interaction, unique to the structures of MICL and ICL, is that between the hydroxyl group of Tyr43 (MICL)/Tyr89 (ICL) and the C(2) of pyruvate/glyoxylate (3.1 or 3.0 Å, respectively).

In the C123S MICL/isocitrate complex, the glyoxyl moiety is shifted considerably relative to the pyruvate ligand of the C123S MICL/product complex (Figure 4C). Superposition of the Mg^{2+}, the coordinating water molecules, and the carboxylic acid residues Asp58, Asp85, Asp87, and Glu115 in the two structures results in displacement of glyoxyl atoms from the analogous pyruvate atoms in the range of 0.5 to 1.5 Å. For comparison, the superposition places the Mg^{2+} ions much closer together, separated by only 0.1 Å. The interaction distances of the carboxyl C(1)O_2 and the C(2)O of the glyoxyl moiety with Gly46 NH and Arg158 guanidinium group, respectively, are 0.3 Å longer than those of the pyruvate. Clearly, the carboxyl and carbonyl groups of the pyruvate ligand are better positioned for interaction with the Mg^{2+} cofactor and with the electropositive enzyme groups than are the analogous carboxyl and carbonyl groups of the glyoxyl unit of the isocitrate ligand. This difference may explain why isocitrate cleavage is not catalyzed by *E. coli* MICL (present work, see below) or by *Samonella enterica* MICL (*14*) under forcing conditions (with high enzyme concentration and long reaction time).

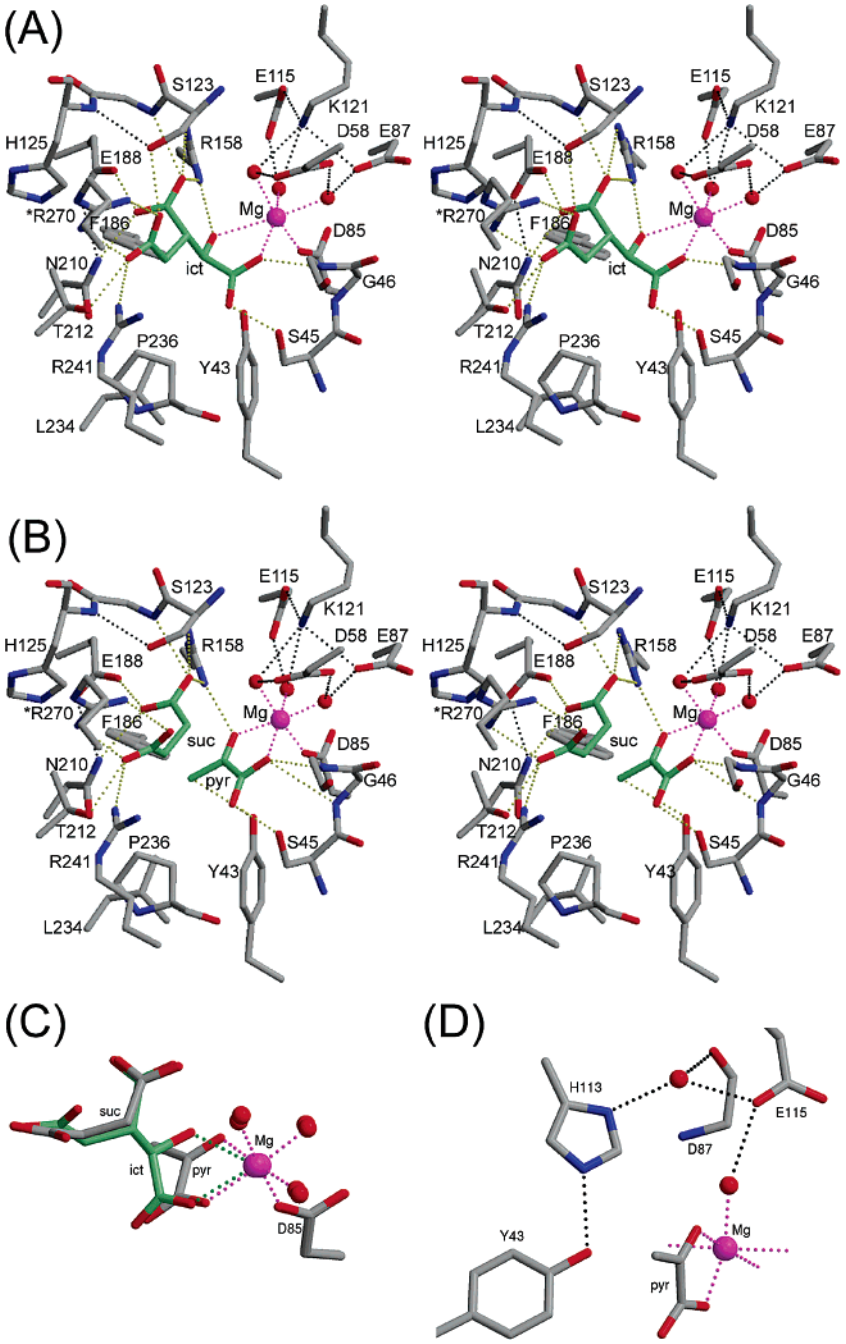

FIGURE 4: Binding modes of isocitrate and product in the active site of MICL: (A) stereoscopic representation of Mg^{2+}-isocitrate binding; (B) stereoscopic representation of Mg^{2+}-pyruvate and succinate binding; (C) superposition of the isocitrate and the product molecules, depicting the nonproductive binding mode of isocitrate—the superposition was calculated with active site residues and Mg^{2+} and its coordinating water molecules but did not include the ligands; (D) the hydrogen bond network connecting Tyr43 to Glu115. Atomic colors are as follows: oxygen − red, carbon − gray, sulfur − yellow, Mg^{2+} − magenta. Ligand's carbon atoms are colored pale green. Key interactions are highlighted with dotted lines as follows: Coordination to the Mg^{2+} is shown in magenta, electrostatic interactions with the ligand are shown in yellow, and protein interactions with Mg^{2+}-coordinated water molecules and protein−protein interactions are shown in black. Arg270 of the neighboring molecule is labeled with ∗.

The C(1) carboxyl group of the succinate product, and the analogous C(3) carboxyl group in isocitrate, interact with Arg158 and with Glu188 side chains. Interactions between two carboxylate-containing residues have been observed in other protein crystal structures (*48*), and indicate that they share a proton with an abnormally high pK_{a2}. The protein environment supports such arrangement by providing other favorable electrostatic interactions. In MICL, Glu188 shares a proton with the C(3) carboxyl group of the isocitrate and with C(4) carboxylate of succinate (Figure 4). The same succinate/isocitrate carboxyl group also engages in electrostatic interaction with the backbone NH of residue Gly124, and with the side chain of Asn210, which in turn interacts with the carboxyl group of Glu188. This same constellation of interactions is conserved in the structure of ICL bound to glyoxylate and the succinate analogue, 3-nitropropionate (*10*). We note that the crystals of MICL and ICL complexes were obtained at pH 6 and pH 8, respectively, and the optimal pH range for catalysis by *S. enterica* MICL in solution is 7−8 (*14*), indicating that the carboxyl-carboxylate interaction

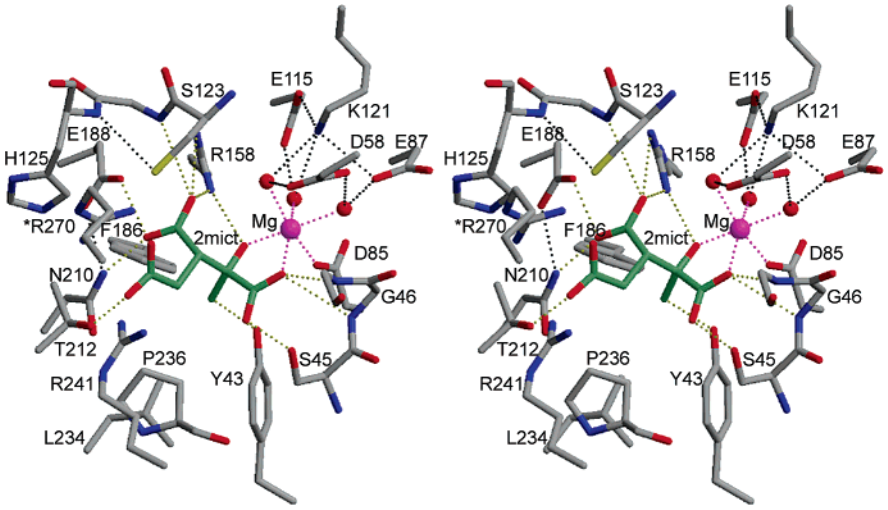

FIGURE 5: Stereoscopic view of a model of 2-methylisocitrate bound to MICL. Colors are as in Figure 4. The model was generated by mutating Ser123 to cysteine, and by adding a methyl group to isocitrate followed by shifting the molecule to optimize the pyruvyl moiety interaction with Mg^{2+}.

observed in the crystals is also possible at physiological pH.

Extending the electrostatic interaction network further, Glu188 interacts with the invariant His125 imidazole group, which site-directed mutagenesis studies carried out on the *S. enterica* MICL have shown to contribute a rate factor of 100 to catalytic turnover (*14*). As discussed below, Glu188 is a key catalytic residue, and its ionization state and exact orientation are crucial for function.

The C(4) carboxyl group of succinate bound to MICL forms two charge−charge interactions, one with Arg241 and the second with Arg270 on the N-terminal tail of the neighboring molecule that traverses the entry to the active site. Both arginine residues are conserved among the four MICLs with experimentally confirmed function, but not in ICL. The C(4) carboxyl group of succinate also interacts with Thr212, conserved in confirmed MICLs (a serine in ICLs).

A Model of Bound Substrate. A model of wild-type MICL bound with 2-methylisocitrate (Figure 5) was built by adding a methyl group bonded to C(2) of (2*R*,2*S*)-isocitrate, and moving the molecule in the active site so that the interactions of the pyruvyl moiety with Mg^{2+} and enzyme residues resemble more closely those of pyruvate in the enzyme/product complex (Figure 4A). The ligand C(2) and its methyl group move 1 Å, leading to a better fit of the methyl in the hydrophobic pocket. We propose that in this manner, the substrate is poised for the chemical reaction to take place, whereas the binding mode seen for the isocitrate inhibitor is nonproductive.

In addition to the adjustment in the position of the substrate, Ser123 was mutated back to a cysteine residue, and its side chain conformation was rotated slightly (by 20°) so that the thiol/thiolate group avoids short contacts with the backbone NH groups of Gly124 and His125, while remaining appropriately oriented for proton transfer to/from the C(2) of the succinate moiety.

Catalytic Mechanism. In developing a model for the catalytic mechanism of MICL based on structure−function data, we have also taken into consideration earlier work carried out on ICL, which we assume to have a similar catalytic mechanism because of the similarities in the active site architecture, substrate structure, and the chemistry performed.

It is evident from the structural data, that the MICL substrate is activated for C(2)OH deprotonation through binding of the C(2)O atom to the Mg^{2+} cofactor and to the Arg158 guanidinium group, and by binding of the C(2) carboxyl substituent to the Mg^{2+} and the N-terminal turn, Ser45-Gly46-Gly47, of a short helix (residues 46−54). An analogous set of interactions is present in the structure of the *M. tuberculosis* ICL/ Mg^{2+}-nitroproprionate-glyoxyl complex (*10*). In the reverse direction, leading to condensation, these interactions would serve to activate the pyruvate (MICL) or glyoxylate (ICL) carbonyl for nucleophilic addition.

The identity of the catalytic base used for C(2)OH deprotonation is not entirely clear and in Figure 6 we show one possible scenario. Specifically, orientation and distance from the substrate hydroxyl group suggest two possible water molecule candidates, one that is hydrogen bonded to the invariant Asp58, and another that is hydrogen bonded to the invariant Glu115. Both water molecules are coordinated to the Mg^{2+} cofactor and thus anchored, they may shuttle a proton from the substrate C(2)OH to a hydrogen-bonded active site carboxylate group (viz. Asp58 or Glu115). The Asp58 residue was replaced with Ala in the *S. enterica* MICL by site directed mutagenesis to show that it is essential to catalytic turnover (*14*).

Two other alternatives for a catalytic base have been considered. One in which the Arg158 guanidium group might serve as a proton shuttle from the 2-methylisocitrate C(2)-OH to the C(3)COO− substituent, thereby offsetting the negative charge that forms at C(3) as the cleavage step proceeds. The caveat is that the Arg158 guanidinium group is already positively charged, and to the best of our knowledge there are no other examples of arginine residues that mediate proton shuttle. The second alternative is that Tyr43 hydroxyl group serves as the base. The docked 2-methylisocitrate presented in Figure 5 corresponds to both C(2) methyl and hydroxyl groups not overlaping their counterpart atoms in pyruvate because C(2) of 2-methylisocitrate is tetrahedral and that of pyruvate is planar. A torsion of 35° around the C(1)−C(2) bond of 2-methylisocitrate superposes the $C(2)CH_3$ on its pyruvate counterpart, and brings the C(2) hydroxyl group within 2.7 Å from Tyr43

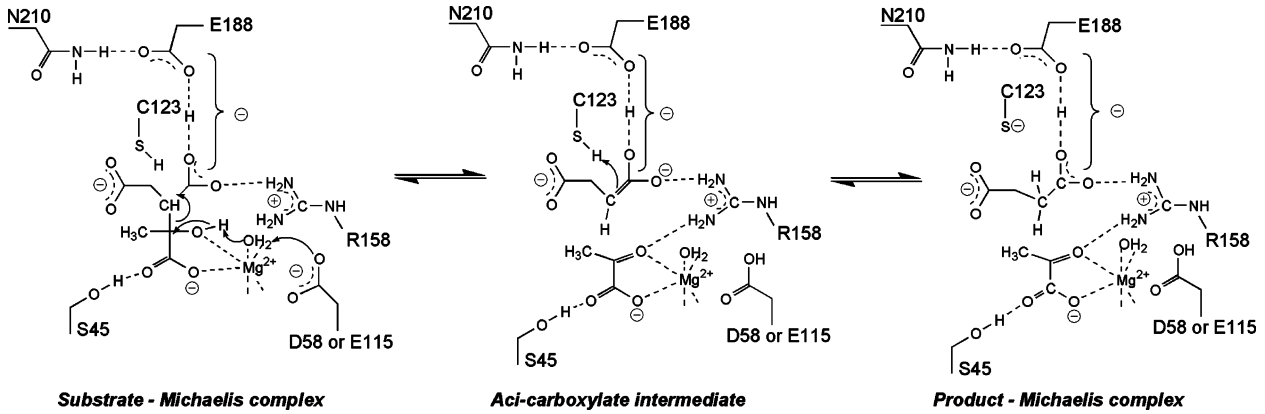

FIGURE 6: Proposed catalytic mechanism of MICL.

hydroxyl. In the X-ray structures, Tyr43 hydroxyl forms a hydrogen bond with His113 Nϵ atom (Figure 4D). The Nδ atom of His113 is hydrogen bonded to an internal water molecule, which in turn forms hydrogen bonds with Glu115 carboxylate group and Asp87 backbone carbonyl. The network of interactions provides a second route for shuttling a proton from C(2)OH to Glu115, this time via Tyr43 hydroxyl. The caveat here is that while the MICL and ICL enzymes with confirmed function contain the tyrosine-histidine pair, and both crystal structures contain the same internal water molecule, an exception occurs in the *Dianthus caryophyllus* petal death protein (PSR132, Swiss-Prot accession number Q05957, annotated as putative carboxyPEP mutase, (*49*)), where two phenylalanine residues replace the tyrosine-histidine pair. Activity studies of this enzyme revealed that in contrast to the annotation in the sequence database, the protein has no mutase activity (*50*). Instead, it exhibits a lyase activity, cleaving 2-methylmalate and 2-ethylmalate with $k_{cat}/K_m \approx 1000$ M^{-1} s^{-1} in the presence of Mg^{2+} at pH 7, and 25 °C.

The location of the catalytic acid responsible for protonation of the succinate leaving group can be inferred from the reported inversion of stereochemistry in the succinate leaving group. Specifically, ICL catalyzed cleavage of (2*R*,3*S*)-isocitrate in D_2O yields (2*S*)-[D_1]-succinate (*51*) and therefore, the catalytic acid must be positioned at the isocitrate C(3) side opposite of the departing C(2). In the MICL structures, the invariant and catalytically essential active site cysteine is ideally positioned to deliver the proton, as is the case in ICL. In both enzymes, the cysteine is activated for proton transfer by interaction with the backbone amide NH of the Asx turn. Note that the model of bound substrate shows that C(3) of the intact substrate is remote from Cys123 thiol group, at 4.5 Å. However, once the C(2)−C(3) bond is cleaved, the aci-carboxylate intermediate described below must move away from the pyruvate to the position of the observed succinate product, bringing the hydrogen accepting carbon atom within 3.2 Å from the thiol group.

Unless proton transfer from the MICL Cys123 is synchronized with the C(2)−C(3) bond cleavage of 2-methylisocitrate, a carbanion-like transition state might precede the charged intermediate, stabilized by the enzyme as the planar aci-carboxylate [C(2)═C(1)(O^-)(O^-)]. Such an intermediate would be intrinsically high in energy given that the pK_a of an acetate anion in water is 33.5 (*52*). The enzyme may offset the high energy of the ligand by providing a favorable electrostatic environment (the reader is referred to references (*52*−*58*) for discussion of enzyme stabilization of this type of charged intermediates). The MICL active site is highly polar and pre-organized for favorable electrostatic interaction with an aci-carboxylate intermediate and/or aci-carboxylate like transition state. In particular, the Glu188 carboxyl group shares a proton with the aci-carboxylate while also engaging in electrostatic interaction with the NH_2 of Asn210. Additionally, the aci-carboxylate binds to the Arg158 guanidium group and the backbone amide of Gly124. Analogous electrostatic interactions are present in ICL. It is note worthy that the Glu188 and Arg158 interaction with the putative aci-carboxylate formed in MICL, is analogous to the Glu317 and Lys164 interaction with the proposed aci-carboxylate intermediate of mandelate racemase (*54*). The intermediacy of the succinate aci-acid carboxylate was previously suggested for ICL, based on the finding the aci-carboxylate carbanion analogue, 3-nitroproprionate, binds to *Pseudomonas indigofera* ICL ($K_i = \leq 17$ nM at pH 8) 65,000 times more tightly than does succinate (*59*).

MICL and ICL Substrate Specificity. Steady-state kinetic methods were used to measure *E. coli* ICL and MICL activities with (2*R*,3*S*)-isocitrate and with racemic *threo*-2-methylisocitrate (i.e., a mixture of (2*R*,3*S*)- and (2*S*,3*R*)-2-methylisocitrate), (2*S*)- and (2*R*)-methylmalate, and (*S*)- and (*R*)-malate. Previous studies (*6*) showed that the Mg^{2+} activated *E. coli* MICL (1) catalyzes the formation of a single *threo* enantiomer, presumably (2*R*,3*S*)-2-methylisocitrate, from pyruvate and succinate ($K_{eq} \leq 0.06$ M^{-1}), (2) that the enzyme is not active toward either of the *erythro* enantiomers (i.e., (2*R*,3*R*)- and (2*S*,3*S*)-2-methylisocitrate) ($k_{cat} < 2 \times 10^{-3}$ s^{-1}), and (3) that the $k_{cat} = 12$ s^{-1} and apparent $K_m = 19$ μM for the reaction of active *threo* enantiomer at 30 °C, pH 7.

In the present investigation, MICL was shown to catalyze under exhaustive conditions (with high enzyme concentration and long reaction time) the conversion of 50% of the synthetic mixture of (2*R*,3*S*) and (2*S*,3*R*)-2-methylisocitrate with a $k_{cat} = 19$ s^{-1} and $K_m = 24$ μM. Therefore, MICL known stereospecificity toward the (2*R*,3*S*)-2-methylisocitrate enantiomer is consistent with the crystal structures of the complexes. MICL was not active with the mixture of *erythro*-enantiomers under conditions that set an upper limit of the k_{cat} of 1×10^{-5} s^{-1}. Likewise, reaction of *E. coli* MICL with (2*R*,3*S*)-isocitrate, under forcing conditions failed to

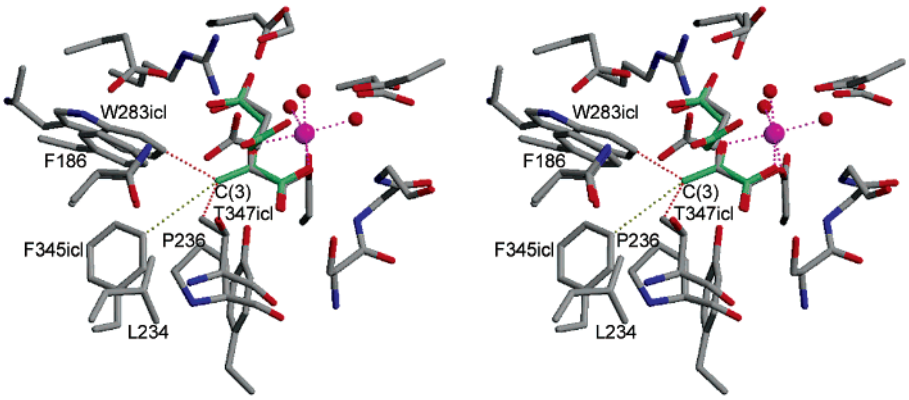

FIGURE 7: Stereoscopic view of ICL/ Mg^{2+}-glyoxylate-succinate aligned with MICL/Mg^{2+}-pyruvate-succinate, focusing on the hydrophobic pocket. The superposition was generated to optimize overlap between Mg^{2+}, the Mg^{2+} coordinating adducts, Arg158, Asp58, Asp85, and Glu115, and the ICL counterparts. Colors are as in Figure 4, except that carbon atoms of the glyoxylate and succinate in the ICL structure are colored gray to differentiate them from pyruvate and succinate in the MICL structure. For clarity, only three ICL residues are depicted, those of the hydrophobic pocket. Their relationship to the superposed pyruvate C(3) are shown in dotted lines, with red dotted lines corresponding to short van der Waals contacts.

produce the glyoxylate and succinate products. The k_{cat} for this reaction is therefore less than the detection limit of 1×10^{-5} s^{-1}. A similar observation was reported for the *S. enterica* MICL (*14*). Thus, the methyl group at C(2) of the isocitrate skeleton, poised in R configuration, is an absolute requirement for MICL catalytic turnover. Binding affinity is not the basis for the substrate specificity: the K_i measured for competitive inhibition by (2*R*,3*S*)-isocitrate is 400 ± 50 μM. As was described in the previous section, isocitrate binds to the C123S MICL in an orientation which miss-aligns the glyoxylate moiety with the Mg^{2+} cofactor and enzyme catalytic groups. Such unproductive binding could account for the stringent substrate specificity.

E. coli ICL catalyzed the conversion of (2*R*,3*S*)-isocitrate to gyloxylate and pyruvate with a $k_{cat} = 100$ s^{-1} and $K_m = 18$ μM. In contrast, no activity was observed toward *threo* or *erythro*-2-methylisocitrate ($k_{cat} < 1 \times 10^{-5}$ s^{-1}). On the other hand, the ICLs *from N. crassa, P. indigofera, Chlorella vulgaris,* and *Saccharomyces cerevisae* have reported activity toward *threo*-2-methylisocitrate, which varies from 5 to 50-fold less than the maximal activity observed toward (2*R*,3*S*)-isocitrate (*5, 59*). Thus, the degree of substrate specificity toward the C(2) methyl substituted isocitrate appears to vary from one ICL to another.

Last, *E. coli* ICL and MICL activities with the natural products malate and 2-methylmalate (citramalate) were tested to assess the importance of the succinate moiety of the respective natural substrates in productive binding. Neither lyase catalyzed the cleavage of (*R*)-malate, (*S*)-malate, (*R*)-2-methylmalate or (*S*)-2-methylmalate under forcing conditions.

Structure and Sequence Patterns of Lyase Substrate Specificity. The mechanism of MICL and ICL discrimination between isocitrate and 2-methylisocitrate is of particular interest because it, along with sequence divergence, provides insight into the evolution of lyases within the ICL/PEP mutase superfamily for function in different metabolic pathways. MICL and ICL function in two different pathways that branch from the citric acid cycle. These pathways are not found in all organisms nor do they necessarily occur together. Thus, ICL and MICL are the products of enzyme evolution directed at specialization.

The requirement for the (2*R*,3*S*) configuration for the MICL substrate 2-methylisocitrate and the ICL substrate isocitrate is easily understood in the context of the numerous interactions between the enzyme and the C(2) and C(3) substituents described previously. These interactions anchor the substrates in the correct orientation to trigger the proton transfer and the C−C bond cleavage.

The hydrophobic pocket surrounding the pyruvate C(2)-CH_3 in MICL is absent in ICL (Figure 7). The superposition shows that the C(2) methyl group of pyruvate (and therefore also of 2-methylisocitrate) cannot be accommodated in the hydrophobic pocket of ICL because it forms short van der Waals contacts with Thr347 (2.6 Å) and Trp283 (3.1 Å), a conclusion also reached by Grimm et al (*1*). The extra room that accommodates the C(2)methyl group of 2-methylisocitrate in MICL allows an isocitrate ligand to slide over into a nonproductive binding mode, explaining the competitive inhibition of MICL by isocitrate, rather than catalytic turnover.

In this work, the *E. coli* ICL was shown to be inactive toward 2-methylisocitrate, consistent with the prediction of specificity by steric restrictions. On the other hand, the ICLs from *N. crassa, P. indigofera, C. vulgaris,* and *S. cerevisae,* which reportedly are reasonably active with 2-methylisocitrate (*5, 59*), conserve the binding pocket triad of the *E. coli* and *M. tuberculosis* ICLs. Thus, substrate specificity in ICL does not appear to derive from a single determinant. It has been previously pointed out that the binding pocket triad of MICL is not conserved in *S. cerevisiae,* wherein Phe186 is replaced with a tryptophan (*1*). Likewise, we find that some sequences of the MICL sequence family substitute cysteine for Pro236 and/or isoleucine for Leu234 (Figure 8A). However, substrate specificity has not been tested in these enzymes and therefore the significance of these substitutions cannot yet be determined.

From a protein engineering perspective, the short contact with Thr347 is the consequence of main chain differences between the two structures, thus simply mutating Pro236 of MICL to a threonine is not expected to convert MICL to an ICL. Conversely, an ICL mutation of T347P would not create a MICL. The replacements that may achieve reversal of substrate specificity are P236L or P236I in MICL (introduc-

(A)

```
MICL_E.coli              19 PLQIVGTINANHALLAQRAG.YQAIYLSGGGVAAGS.....LGLPDLGISTLDDVLTDIRRITDVC.............................SLPLLVDADIGFGSSAFNVARTVKS 103
MICL_S.typhimurium       19 PLQIVGAINANHALLAQRAG.YQAIYLSGGGVAAGS.....LGLPDLGISTLDDVLTDIRRITDVC.............................PLPLLVDADIGFGSSAFNVARTVKS 103
MICL_B.sacchari          22 PLQVVGAITAYAAKLAEQTG.FKAVYLSGGGVAANS.....LGVPDLGISTMDDVLIDARRITDAV.............................DIPLMVDIDTGWG.GAFNIARTIKS 105
MICL_C.glutamicum        26 IARLPGAFSPLIARSIEEAG.FEGVYVSGAVIAAD......LALPDIGLTTLTEVAHRARQIARVT.............................DLGVLVDADTGFG.EPMSAARTVAE 108
MICL_C.glutamicum        27 IARMPGAFSPLAARAIQEAG.FEGVYVSGAVVAAD......LALPDIGLTTLTEVAHRSRQIARVT.............................DLPVLVDADTGFG.EPMSAARTVSE 109
ICL_M.Tuberculosis       66 WVNALGALTGNMAVQQVRAG.LKAIYLSGWQVAGDAN.LSGHTYPDQSLYPANSVPQVVRRINNALQRADQIAKIEG..........DTSVENWLAPIVADGEAGFG.GALNVYELQKA 171
ICL_C.glutamicum         67 YINALGALTGNQAVQQVRAG.LKAVYLSGWQVAGDAN.LSGHTYPDQSLYPANSVPSVVRRINNALLRSDEIARTEG..........DTSVDNWVVPIVADGEAGFG.GALNVYELQKA 172
ICL_M.loti               58 FVNALGAMSGNQAMQQVRAG.LKAIYLSGWQVAADAN.TASAMYPDQSLYPANAAPELVKRINRTLQRADQIETSEG..........NGLSVETWFAPIVADAEAGFG.GPLNAFEIMKA 164
ICL_E.coli               65 YINSLGALTGGQALQQAKAG.IEAVYLSGWQVAADAN.LAASMYPDQSLYPANSVPAVVERINNTFRRADQIQWSAGIEP.......GDPRYVDYFLPIVADAEAGFG.GVLNAFELMKA 174
ICL1_S.cerevisiae        81 VSKTFGALDPVQISQMAKY..LDTIYISGWQCSSTAS.TSNEPGPDLADYPMDTVPNKVEHLFKAQLFHDRKQLEARSKAKSQEELDEMGAPIDYLTPIVADADAGHG.GLTAVFKLTKM 196
ICL2_S.cerevisiae       107 PLHTLGVIDPVQMSQLARCRNIKVAYISGWACSSTLVGSTNEVSPDFGDYPYDTVPNQVERIFKAQQLHDRKAFLEASIKG........STPVDYLKPIIADADMGHG.GPTTVMKVAKL 217
CPPM_S.hygroscopicus     18 ILVVPSAYDALSAKVIQQAG.FPAVHMTGSGTSAS.....MLGLPDLGFTSVSEQAINLKNIVLTV.............................DVPVIMDADAGYG.NAMSVWRATRE 101
PSR132_D.caryophyllus    40 SVLMPGVQDALSAAVVEKTG.FHAAFVSGYSVSAAM.....LGLPDFGLLTTTEVVEATRRITAAAP.............................NLCVVVDGDTGGG.GPLNVQRFIRE 124
OXAHYD_B.fuckeliana      39 LIVCPGVYDGISARVALQVG.FPALYMTGAGTTASR.....LGMADLGIAQLSDMKDHAEMIANLDP.............................YGPPLIADMDTGYG.GPLIIDKAVKS 124
OXAHYD_Aspergillus       60 LLVCPGVYDGLSARIAINLG.FKGMYMTGAGTTASR.....LGMADLGLAHIYDMKTNAEMIANLDP.............................YGPPLIADMDTGYG.GPLMVARSVQQ 145
PEPM_T.pyriformis        28 LSFIMEAHNGLSAAIVEETG.FKGIWGSGLSISAA......MGVRDSNEASYTQVLEVLEFMSDRT.............................KIPILLDGDTGYG.NYNNARRLVKK 110
PEPM_M.edulis            20 LEFIMEAHNGLSARIVQEAG.FKGIWGSGLSVSAQ......LGVRDSNEASWTQVVEVLEFMSDAS.............................DVPILLDADTGYG.NFNNARRLVRK 102
PPYRH_Variovorax.sp      16 LFTAMAAHNPLVAKLAEQAG.FGGIWGSGFELSAS......YAVPDANILSMSTHLEMMRAIASTVS.............................IPLIADIDTGFG.NAVNVHYVVPQ  98

MICL_E.coli             104 MIKAGAAGLHIEDQVGA.KRCGHRP..NKAIVSKEEMVDRIRAAVDAK..TDPDFVIMARTDA 161
MICL_S.typhimurium      104 IAKAGAAALHIEDQVGA.KRCGHRP..NKAIVSKEEMVDRIRAAVDAR..TDPNFVIMARTDA 161
MICL_B.sacchari         106 FIKAGVAAVHLEDQVGQ.KRCGHRP..GKEVVSTDEMVDRVKAAVDAR..TDDGFVIMARTDA 163
MICL_C.glutamicum       109 LEDAGVAGCHLEDQVNP.KRCGHLD..GKEVVRTDVMVRRIAAAVSAR..RDPNFVICARTDA 166
MICL_C.glutamicum       110 LEDAGVAGCHLEDQVNP.KRCGHLD..GKEVVGTDIMVRRIAAAVNER..RDEQFVICARTDA 167
ICL_M.Tuberculosis      172 LIAAGVAGSHWEDQLASEKKCGHLG..GKVLIPTQQHIRTLTSARLAADVADVPTVVIARTDA 232
ICL_C.glutamicum        173 MIAAGAAGTHWEDQLASEKKCGHLG..GKVLIPTQQHIRTLNSARLAADVANTPTVVIARTDA 233
ICL_M.loti              165 FIEAGAAGVHYEDQLASEKKCGHLG..GKVLIPTAAHIRNLNAARLAADVMGTPTLVVARTDA 225
ICL_E.coli              175 MIEAGAAAVHFEDQLASVKKCGHMG..GKVLVPTQEAIQKLVAARLAADVTGVPTLLVARTDA 235
ICL1_S.cerevisiae       197 FIERGAAGIHMEDQTSTNKKCGHMA..GRCVIPVQEHVNRLVTIRMCADIMHSDLIVVARTDS 257
ICL2_S.cerevisiae       218 FAEKGAAGIHLEDQMVGGKRCGHLS..GAVLVPTATHLMRLISTRFQWDIMGTENLVIARTDS 278
CPPM_S.hygroscopicus    102 FERVGIVGYHLEDQVNP.KRCGHLE..GKRLISTEEMTGKIEAAVEAR..EDEDFTIIARTDA 159
PSR132_D.caryophyllus   125 LISAGAKGVFLEDQVWP.KKCGHMR..GKAVVPAEEHALKIAAAREAI..GDSDFFLVARTDA 182
OXAHYD_B.fuckeliana     125 YIRAGVAGFHIEDQION.KRCGHLQ..GKKVVPAEEYYMRIRAAKAAKEAMNSDIVLIARTDA 184
OXAHYD_Aspergillus      146 YIQAGVAGFHIEDQION.KRCGHLA..GKRVVTMDEYLTRIRAAKLTKDRLRSDIVLIARTDA 205
PEPM_T.pyriformis       111 LEQRSIAGVCLEDKIFP.KRNSLLDDGRQELAPINEFVAKIKACKDTQ..QDADFQVVARVEA 170
PEPM_M.edulis           103 LEDRGVAGACLEDKLFP.KTNSLHDGRAQPLADIEEFALKIKACKDSQ..TDPDFCIVARVEA 162
PPYRH_Variovorax.sp      99 YEAAGASAIVMEDKTFP.KDTSLRTDGRQELVRIEEFQGKIAAATAAR..ADRDFVVIARVEA 158

MICL_E.coli             170 AIERAQAYVEAGAEMLFPEA.ITELAMYRQFA.....DAVQVPILANITEFGAT.........PLFTTDELRSAHVAMALYPLSAFRAMNRAAEHVYNVLRQEGTQKSVIDTMQTRNELY 274
MICL_S.typhimurium      170 ALDRAQAYVDAGAGDMLFPEA.ITELSMYRRFA.....DVAQVPILANITEFGAT.........PLFTTDELRSAHVAMALYPLSAFRAMNRAAEKVYTVLRQEGTQKNVIDIMQTRNELY 274
MICL_B.sacchari         172 AIERAVAYVEAGADMIFPEAMKT.LDDYRRFR.....AAVKVPILANLTEFGST.........AFFTTEELKSANVDIALYCCGAYRAMNAAALNFYETVKRDGTQKAAVSTMQTREDLY 276
MICL_C.glutamicum       175 AIERAKAYLDAGADMIFTEALHS.EADFRYFR....HAIPDALLLANMTEFGKT.........TLLSADVLEEIGYNAVIYPVTTLRIAMGQVEQALAEIKEHGTQEGWLDRMQHRSRLY 280
MICL_C.glutamicum       176 AIERAKAYADAGADMIFTEALYS.PADFEKFR.....AAVDIPLLANMTEFGKT.........ELLPAQLLEDIGYNAVIYPVTLLRIAMGQVEQALGDIANTGIQTDWVDRMQHRSRLY 280
ICL_M.Tuberculosis      269 CIARAKAYAP.FADLIWMETGTPDLEAARQFSEAVKAEYPDQMLAYNCSPSFNWKKH.LDDATIAKFQKELAAMGFKFQFITLAGFHALNYSMFDLAYGYAQNQMSAYVELQEREF.... 386
ICL_C.glutamicum        270 CIARAKSYAP.YADMIWMETGTPDLELAKKFAEGVRSEFPDQLLSYNCSPSFNWSAH.LEADEIAKFQKELGAMGFKFQFITLAGFHSLNYGMFDLAYGYARGEMTSFVDLQNREF.... 387
ICL_M.loti              264 CIARAVAYAP.HADLIWCETSKPDLAQAKKFAEGVRRHHPGKLLAYNCSPSFNWKKN.LDDATIAKFQRELGAMGYKFQFITLAGFHQLNYGMFELARGYKERQMAAYSELQEAEF.... 381
ICL_E.coli              272 AISRGLAYAP.YADLVWCETSTPDLELARRFAQAIHAKYPGKLLAYNCSPSFNWQKN.LDDKTIASFQQQLSDMGYKFQFITLAGIHSMWFNMFDLANAYAQGEGMKHYVEKVQQP.... 389
ICL1_S.cerevisiae       392 SIMRARAFAP.YADLVWMESNYPDFQQAKEFAEGVKEKFPDQWLAYNLSPSFNWPKA.MSVDEQHTFIQRLGDLGYIWQFITLAGLHTNALAVHNFSRDFAKDGMKAYAQNVQQRE.... 509
ICL2_S.cerevisiae       414 AIRRSLVFAP.YSDMIWLETKTPDLEQARSFSRKIHKQLPATKLVYNLSPSFNWSAHGFDDKALKSFVWDLAKEGFTLQLVSLAGLHSDGVSFWELANSFQSDGMKAYVEKVQKRE.... 532
CPPM_S.hygroscopicus    168 AIRRSREYVAAGADCIFLEAMLD.VEEMKRVR.....DEIDAPLLANMVEGGKT.........PWLTTKELESIGYNLAIYPLSGWMAAASVLRKLFTELREAGTTQKFWDDMGLKMSFA 272
PSR132_D.caryophyllus   191 GIRRANLYKEAGADATFVEAPAN.VDELKEVS.....AKTKGLRIANMIEGGKT.........PLHTPEEFKEMGFHLIAHSLTAVYATARALVNIMKILKEKGTTRDDLDQMATFSEFN 295
OXAHYD_B.fuckeliana     193 CIKRLKVAREMGADVGLLEGYTSKEMAAKTVK.....EFAPWPILLNMVENGST.........PIITTKEAQEMGFRIMIFSFAALAPAYLAIQETFLRLKRDGVVGTPKNLTPKALF.. 299
OXAHYD_Aspergillus      214 CIRRLKAARDLGADVGLLEGFTSKEMARRCVQ.....DLAPWPLLLNMVENGAG.........PVISVDEARERMGFRIMIFSFACITPAYMGITAALERLKKDGVVGLPEGMGPKKLF.. 319
PEPM_T.pyriformis       180 ALKRAEAYRNAGADAILMHSKLKEPSEIEAFMKEWKN...RSPVIIVPTNY............HTVPTDTFRKWGVNMVIWANHNMRACVSAMQETSRRIYEDESLVNVEPKVAKVKEVF 284
PEPM_M.edulis           172 ALKRAEAYRNAGADAILMHSKKADPSDIEAFMKAWNN...QGPVVIVPTKY............YKTPTDHFRDMGVSMVIWANHNLRASVSAIQQTTKQIYDDQSLVNVEDKIVSVKEIF 276
PPYRH_Variovorax.sp     168 AVRRGQAYEEAGADAILIHSRQKTPDEILAFVKSWPG...KVPLVLVPTAYP............QLTEADIAALSKVGIVIYGNHAIRAAVGAVREVFARIRRDGGIREVDAALPSVKEII 273
```

(B)

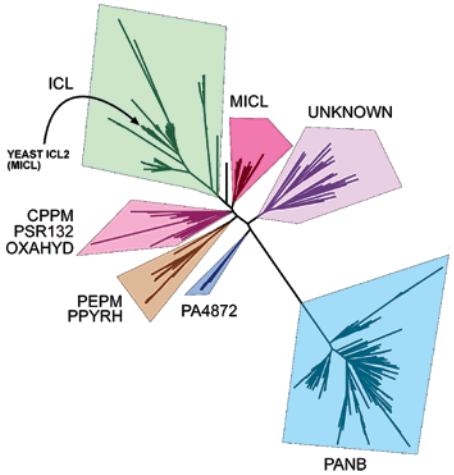

FIGURE 8: Sequence analyses of the ICL superfamily: (A) Structure-based multiple sequence alignment of superfamily members with confirmed function. For brevity, sequences close to the active site region are included. Abbreviated protein names and the organism names are listed. For each block, the residue numbers at the beginning and end of the block are listed. Invariant residues or function discriminating residues that have been discussed in the text are highlighted in yellow. (B) Unrooted neighbor-joining phylogeny tree of the ICL superfamily. The tree was generated as described in the methods section. It includes 512 sequences, and therefore only major nodes and global clusters are depicted. The labeling of the families is based on the location of proteins that have experimentally confirmed function. Abbreviations used: MICL, 2-methylisocitrate lyase; ICL, isocitrate lyase; CPPM, carboxyPEP mutase; OXAHYD, oxaloacetate acetylhydrolase; PEPM, PEP mutase; PPYRH, phosphonopyruvate hydrolase; PANB, ketopantoate hydroxymethyl transferase.

ing a bulkier side chain), and T347A in ICL (reducing the size of the side chain). Along this line, Thr347 is replaced by a serine in the ICL2 of *S. cerevisae* that encodes for a mitochondrial MICL (*60*) and is inactive toward isocitrate (*61*) (Figure 8A). Because the amino acid sequence of ICL2 is so closely related to that of the true *S. cerevisae* ICL, ICL1 (same length and 38% sequence identity), it is tempting to speculate that compared with a threonine side chain, the missing methyl group in the serine side chain opens a niche to accommodate the C(2)methyl of 2-methylisocitrate.

While discrimination between ICL and MICL based on conservation of the amino acids that define the hydrophobic

pocket may not be obvious because of backbone differences, there is a sequence fingerprint that enables discrimination between the two enzyme families. Neither *E. coli* ICL or MICL recognized malate or 2-methyl malate as substrates. Thus, both CH_2COO- moieties of the succinate unit of isocitrate and 2-methylisocitrate are required for catalytic turnover. As discussed in previous sections, the contacts made between the C(1) carboxyl group and ICL or MICL residues are extensive and similar. In contrast, the contacts that the C(4) carboxyl group makes with MICL are different than those made with ICL. In MICL, the interactions are with Arg241, Arg270, and Thr212, residues that are conserved in all MICL sequences with confirmed function (Figure 8A), as well as in the genomic sequences that form the MICL group in the phylogeny tree (Figure 8B). Tyr274, a residue that interacts with Arg270 is also conserved within this cluster. For ICL, the interactions with the succinate C(4) carboxyl group are mediated by His193, Asn313, Ser315, Ser317, and Thr347 (numbering according to the PDB entry of *M. tuberculosis* ICL). These residues are conserved in ICL with confirmed function as well as in most of the clustered genomic sequences corresponding to the ICL family. As His123 and Asn313 are invariant also in MICL sequences, it is the presence of Ser315, Ser317, and Thr347 that provides the fingerprint for recognizing the ICL enzymes. The fungi ICLs in which Thr347 is replaced by a serine (10 members) are the exception to the rule, and in one case, that of the *S. cerevisiae* ICL2, an MICL activity was demonstrated (*60*). These proteins are located within the ICL cluster in the phylogeny tree (Figure 8). Thus, it appears that there is a second pathway for the evolution of MICL enzymes that are more closely related to ICLs.

Functional Divergence of the Superfamily. Because the substrates and the C−C cleavage reactions that they undergo are so similar, we might expect the primary structures of MICL and ICL to be more closely related to one another than to other members of the enzyme superfamily. This however, is not the case. MICL shares significantly greater sequence identity with the P−C bond forming/cleaving carboxyPEP mutase enzymes (30−37%) than it does with ICL (23−32%) (*22*). Moreover, the sequence motif of the active site loop ([K(R,K)CGH] conserved among the MICL, ICL, and oxaloacetate acetylhydrolase (OXAHYD) enzymes is found in the carboxyPEP mutase, but not in the PEP mutase or phosphonopyruvate hydrolase. The catalytic Glu188 in MICL or its counterpart in ICL is also conserved in carboxyPEP mutase. Such similarities in sequence patterns have lead to confusion in the annotation of many newly determined sequences. Structure-based sequence analysis helps resolves such ambiguities, as we have seen for MICL and ICL.

The phylogeny tree (Figure 8B) shows clear separation between several function groups, yet it also reveals a region of ambiguity of function and regions where the function has not been identified.

The family of the ketopantoate hydroxymethyl transferase enzymes is clearly distant from the remaining families. In particular, the sequences contain a diverged Mg^{2+} binding fingerprint. Consistent with the sequence divergence, the three-dimensional structure is less related to the structures of other superfamily members, with monomers that form a decameric assembly rather than tetramers, which in the enzyme from *M. tuberculosis* exhibit swapped helices but not in the enzyme from *E. coli* (*20*, *21*, *62*).

ICL and MICL enzymes form two well separated groups in the tree, but the ICL2 type MICL and its homologues in fungi are contained within the ICL family cluster. In addition to the conservation of the Mg^{2+} binding fingerprint and the substrate-binding arginine residue (Arg158 in MICL), all members of the two clusters exhibit the characteristic cysteine-containing active site capping loop sequence, the catalytic glutamate (Glu188 in MICL), and the active site tyrosine (Tyr43 in MICL) (Figure 8A). The differences in substrate binding motifs between ICL and MICL have been discussed in the previous section, and in addition, ICL sequences are approximately 150 amino acid residues longer than those of MICL.

The family adjacent to MICL in the phylogeny tree includes several enzymes with different function: carboxyPEP mutase, PSR132, and OXAHYD. The cluster shares the same MICL/ICL sequence fingerprint listed above except that Tyr43 is not conserved in carboxyPEP mutase and PSR132, and in some sequences Glu188 is replaced by aspartate. The functional basis for the presence of the cysteine and glutamate residues in carboxyPEP mutase is not yet understood, as the enzyme catalyzes a very different chemical reaction than that of the lyases (P−C bond formation). It would be interesting to establish whether the Cys-Glu catalytic apparatus is utilized here in cleavage of the carboxyphosphoryl CO_2 group. Overall, the group exhibits a diverse functional specialization and it is yet to be determined whether some of its members have different undiscovered functions.

The P−C bond breaking family that includes PEP mutase and phosphonopyruvate hydrolase is well separated from other families. For the PEP mutase, the active site cysteine is replaced by asparagine, which the structure shows to be involved in loop closure and desolvation of the substrate (*19*, *29*). The phosphonopyruvate hydrolase sequence contains a threonine in the same position whose functional role is not yet understood.

The phylogenic tree reveals two distinct clusters of proteins of unknown function. One cluster includes PA4872 from *Pseudomonas aeruginosa*, currently under investigation in our laboratory. The second, larger family, does not include any protein with experimentally confirmed function. Some of the sequences are annotated as "PEP mutase and related enzymes", some "probable carboxyPEP mutase", and some are annotated hypothetical proteins. These two clusters indicate that the function diversity of the superfamily may be much broader than is currently known.

ACKNOWLEDGMENT

We thank Drs. John Moult and Arlin Stoltzfus for useful discussions. We thank the staff of the SBC beamlines at the APS for help with data collection.

REFERENCES

1. Grimm, C., Evers, A., Brock, M., Maerker, C., Klebe, G., Buckel, W., and Reuter, K. (2003) Crystal structure of 2-methylisocitrate lyase (PrpB) from *Escherichia coli* and modelling of its ligand bound active centre, *J. Mol. Biol. 328*, 609−621.
2. Simanshu, D. K., Satheshkumar, P. S., Savithri, H. S., and Murthy, M. R. (2003) Crystal structure of *Salmonella typhimurium*

2-methylisocitrate lyase (PrpB) and its complex with pyruvate and Mg(2+), *Biochem. Biophys. Res. Commun. 311*, 193−201.
3. Horswill, A. R., and Escalante-Semerena, J. C. (1999) Salmonella typhimurium LT2 catabolizes propionate via the 2-methylcitric acid cycle, *J. Bacteriol. 181*, 5615−5623.
4. Brock, M., Maerker, C., Schutz, A., Volker, U., and Buckel, W. (2002) Oxidation of propionate to pyruvate in *Escherichia coli*. Involvement of methylcitrate dehydratase and aconitase, *Eur. J. Biochem. 269*, 6184−6194.
5. McFadden, B. A., Rose, I. A., and Williams, J. O. (1972) Production of pyruvate and succinate by action of isocitrate lyase on -methylisocitrate, *Arch. Biochem. Biophys. 148*, 84−88.
6. Brock, M., Darley, D., Textor, S., and Buckel, W. (2001) 2-Methylisocitrate lyases from the bacterium *Escherichia coli* and the filamentous fungus Aspergillus nidulans: characterization and comparison of both enzymes, *Eur. J. Biochem. 268*, 3577−3586.
7. Vanni, P., Giachetti, E., Pinzauti, G., and McFadden, B. A. (1990) Comparative structure, function and regulation of isocitrate lyase, an important assimilatory enzyme, *Comp. Biochem. Physiol. B. 95*, 431−458.
8. Britton, K. L., Abeysinghe, I. S., Baker, P. J., Barynin, V., Diehl, P., Langridge, S. J., McFadden, B. A., Sedelnikova, S. E., Stillman, T. J., Weeradechapon, K., and Rice, D. W. (2001) The structure and domain organization of *Escherichia coli* isocitrate lyase, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 1209−1218.
9. Britton, K., Langridge, S., Baker, P. J., Weeradechapon, K., Sedelnikova, S. E., De Lucas, J. R., Rice, D. W., and Turner, G. (2000) The crystal structure and active site location of isocitrate lyase from the fungus Aspergillus nidulans, *Struct. Fold. Des. 8*, 349−362.
10. Sharma, V., Sharma, S., Hoener zu Bentrup, K., McKinney, J. D., Russell, D. G., Jacobs, W. R., Jr., and Sacchettini, J. C. (2000) Structure of isocitrate lyase, a persistence factor of *Mycobacterium tuberculosis*, *Nat. Struct. Biol. 7*, 663−668.
11. Ko, Y. H., and McFadden, B. A. (1990) Alkylation of isocitrate lyase from *Escherichia coli* by 3-bromopyruvate, *Arch. Biochem. Biophys. 278*, 373−380.
12. Roche, T. E., McFadden, B. A., and Williams, J. O. (1971) Modification of the active site of isocitrate lyase from Pseudomonas indigofera, *Arch. Biochem. Biophys. 147*, 192−200.
13. Diehl, P., and McFadden, B. A. (1994) The importance of four histidine residues in isocitrate lyase from *Escherichia coli*, *J. Bacteriol. 176*, 927−931.
14. Grimek, T. L., Holden, H., Rayment, I., and Escalante-Semerena, J. C. (2003) Residues C123 and D58 of the 2-methylisocitrate lyase (PrpB) enzyme of Salmonella enterica are essential for catalysis, *J. Bacteriol. 185*, 4837−4843.
15. Rehman, A., and McFadden, B. A. (1997) Cysteine 195 has a critical functional role in catalysis by isocitrate lyase from *Escherichia coli*, *Curr. Microbiol. 35*, 267−269.
16. Rehman, A., and McFadden, B. A. (1997) Lysine 194 is functional in isocitrate lyase from *Escherichia coli*, *Curr. Microbiol. 35*, 14−17.
17. Rehman, A., and McFadden, B. A. (1997) Serine319 and 321 are functional in isocitrate lyase from *Escherichia coli*, *Curr. Microbiol. 34*, 205−211.
18. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI−BLAST: a new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
19. Huang, K., Li, Z., Jia, Y., Dunaway-Mariano, D., and Herzberg, O. (1999) Helix swapping between two alpha/beta barrels: crystal structure of phosphoenolpyruvate mutase with bound Mg(2+)-oxalate, *Struct. Fold. Des. 7*, 539−548.
20. Chaudhuri, B. N., Sawaya, M. R., Kim, C. Y., Waldo, G. S., Park, M. S., Terwilliger, T. C., and Yeates, T. O. (2003) The crystal structure of the first enzyme in the pantothenate biosynthetic pathway, ketopantoate hydroxymethyltransferase, from M tuberculosis, *Structure (Camb.) 11*, 753−764.
21. von Delft, F., Inoue, T., Saldanha, S. A., Ottenhof, H. H., Schmitzberger, F., Birch, L. M., Dhanaraj, V., Witty, M., Smith, A. G., Blundell, T. L., and Abell, C. (2003) Structure of E. coli ketopantoate hydroxymethyl transferase complexed with ketopantoate and Mg2+, solved by locating 160 selenomethionine sites, *Structure (Camb.) 11*, 985−996.
22. Jia, Y., Lu, Z., Huang, K., Herzberg, O., and Dunaway-Mariano, D. (1999) Insight into the mechanism of phosphoenolpyruvate mutase catalysis derived from site-directed mutagenesis studies of active site residues, *Biochemistry 38*, 14165−14173.
23. Sugantino, M., Zheng, R., Yu, M., and Blanchard, J. S. (2003) *Mycobacterium tuberculosis* ketopantoate hydroxymethyltransferase: tetrahydrofolate-independent hydroxymethyltransferase and enolization reactions with alpha-keto acids, *Biochemistry 42*, 191−199.
24. Schmitzberger, F., Smith, A. G., Abell, C., and Blundell, T. L. (2003) Comparative analysis of the *Escherichia coli* ketopantoate hydroxymethyltransferase crystal structure confirms that it is a member of the (betaalpha)8 phosphoenolpyruvate/pyruvate superfamily, *J. Bacteriol. 185*, 4163−4171.
25. Ruijter, G. J., van de Vondervoort, P. J., and Visser, J. (1999) Oxalic acid production by *Aspergillus niger*: an oxalate-nonproducing mutant produces citric acid at pH 5 and in the presence of manganese, *Microbiology 145* (Pt 9), 2569−2576.
26. Hidaka, T., Seto, H., and Imai, S. (1989) Biosynthetic mechanism of C−P bond formation; isolation of carboxyphosphonoenolpyruvate and its conversion to phosphinopyruvate, *J. Am. Chem. Soc. 111*, 8012−8013.
27. Kulakova, A. N., Wisdom, G. B., Kulakov, L. A., and Quinn, J. P. (2003) The purification and characterization of phosphonopyruvate hydrolase, a novel carbon−phosphorus bond cleavage enzyme from Variovorax sp Pal2, *J. Biol. Chem. 278*, 23426−23431.
28. Liu, S., Lu, Z., Han, Y., Jia, Y., Howard, A., Dunaway-Mariano, D., and Herzberg, O. (2004) Conformational flexibility of PEP mutase, *Biochemistry 43*, 4447−4453.
29. Liu, S., Lu, Z., Jia, Y., Dunaway-Mariano, D., and Herzberg, O. (2002) Dissociative phosphoryl transfer in PEP mutase catalysis: structure of the enzyme/sulfopyruvate complex and kinetic properties of mutants, *Biochemistry 41*, 10270−10276.
30. Hoyt, J. C., Robertson, E. F., Berlyn, K. A., and Reeves, H. C. (1988) *Escherichia coli* isocitrate lyase: properties and comparisons, *Biochim. Biophys. Acta 966*, 30−35.
31. Matsuoka, M., and McFadden, B. A. (1988) Isolation, hyperexpression, and sequencing of the aceA gene encoding isocitrate lyase in *Escherichia coli*, *J. Bacteriol. 170*, 4528−4536.
32. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
33. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 849−861.
34. Terwilliger, T. C. (2001) Maximum-likelihood density modification using pattern recognition of structural motifs, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 1755−1762.
35. Terwilliger, T. C. (2000) Maximum-likelihood density modification, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 56*, 965−972.
36. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, G. J. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. Sect. A 47*, 110−119.
37. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
38. Navaza, J. (2001) Implementation of molecular replacement in AMoRe, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 57*, 1367−1372.
39. Gibrat, J. F., Madej, T., and Bryant, S. H. (1996) Surprising similarities in structure comparison, *Curr. Opin. Struct. Biol. 6*, 377−385.
40. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice, *Nucleic Acids Res. 22*, 4673−4680.
41. Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004) The Jalview Java alignment editor, *Bioinformatics 20*, 426−427.
42. Felsenstein, J. (1989) PHYLIP - - Phylogeny Inference Package, *Cladistics 5*, 164−166.
43. Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992) The rapid generation of mutation data matrices from protein sequences, *Comput. Appl. Biosci. 8*, 275−282.
44. Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap, *Evolution 39*, 783−791.

45. Choi, J. H., Jung, H. Y., Kim, H. S., and Cho, H. G. (2000) PhyloDraw: a phylogenetic tree drawing system, *Bioinformatics 16*, 1056−1058.
46. Ko, Y. H., and McFadden, B. A. (1990) The inhibition of isocitrate lyase from *Escherichia coli* by glyoxylate, *Curr. Microbiol. 21*, 313−315.
47. Rees, D. C., Lewis, M., and Lipscomb, W. N. (1983) Refined crystal structure of carboxypeptidase A at 1.54 A resolution, *J. Mol. Biol. 168*, 367−387.
48. Sawyer, L., and James, M. N. (1982) Carboxyl-carboxylate interactions in proteins, *Nature 295*, 79−80.
49. Wang, H., Brandt, A. S., and Woodson, W. R. (1993) A flower senescence-related mRNA from carnation encodes a novel protein related to enzymes involved in phosphonate biosynthesis, *Plant Mol. Biol. 22*, 719−724.
50. Lu, Z. (2003) Diversification of catalytic function within the isocitrate lyase enzyme superfamily. Ph.D. Thesis, University of New Mexico, Albuquerque.
51. Sprecher, M., Berger, R., and Sprinson, D. B. (1964) Stereochemical Course of the Isocitrate Lyase Reaction, *J. Biol. Chem. 239*, 4268−4271.
52. Richard, J. P., Williams, G., O'Donoghue, A. C., and Amyes, T. L. (2002) Formation and stability of enolates of acetamide and acetate anion: an Eigen plot for proton transfer at alpha-carbonyl carbon, *J. Am. Chem. Soc. 124*, 2957−2968.
53. Cleland, W. W., and Kreevoy, M. M. (1994) Low-barrier hydrogen bonds and enzymic catalysis, *Science 264*, 1887−1890.
54. Mitra, B., Kallarakal, A. T., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1995) Mechanism of the reaction catalyzed by mandelate racemase: importance of electrophilic catalysis by glutamic acid 317, *Biochemistry 34*, 2777−2787.
55. Gerlt, J. A., and Gassman, P. G. (1993) Understanding the rates of certain enzyme-catalyzed reactions: proton abstraction from carbon acids, acyl-transfer reactions, and displacement reactions of phosphodiesters, *Biochemistry 32*, 11943−11952.
56. Warshel, A., Papazyan, A., and Kollman, P. A. (1995) On low-barrier hydrogen bonds and enzyme catalysis, *Science 269*, 102−106.
57. Warshel, A. (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites, *J. Biol. Chem. 273*, 27035−27038.
58. Schutz, C. N., and Warshel, A. (2004) The low barrier hydrogen bond (LBHB) proposal revisited: the case of the Asp... His pair in serine proteases, *Proteins 55*, 711−723.
59. Schloss, J. V., and Cleland, W. W. (1982) Inhibition of isocitrate lyase by 3-nitropropionate, a reaction-intermediate analogue, *Biochemistry 21*, 4420−4427.
60. Luttik, M. A., Kotter, P., Salomons, F. A., van der Klei, I. J., van Dijken, J. P., and Pronk, J. T. (2000) The *Saccharomyces cerevisiae* ICL2 gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-coenzyme A metabolism, *J. Bacteriol. 182*, 7007−7013.
61. Heinisch, J. J., Valdes, E., Alvarez, J., and Rodicio, R. (1996) Molecular genetics of ICL2, encoding a nonfunctional isocitrate lyase in *Saccharomyces cerevisiae*, *Yeast 12*, 1285−1295.
62. Chaudhuri, B. N., Sawaya, M. R., Kim, C. Y., Waldo, G. S., Park, M. S., Terwilliger, T. C., and Yeates, T. O. (2003) The crystal structure of the first enzyme in the pantothenate biosynthetic pathway, ketopantoate hydroxymethyltransferase, from M tuberculosis, *Structure (Camb.) 11*, 753−764.

BI0479712