

Crystal Structure of the Petal Death Protein from Carnation Flower^{†,‡}Alexey Teplyakov,[§] Sijiu Liu,[§] Zhibing Lu,^{||} Andrew Howard,[⊥] Debra Dunaway-Mariano,^{||} and Osnat Herzberg^{*,§}

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, Rockville, Maryland 20850, Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, and Biological, Chemical, and Physical Sciences, Illinois Institute of Technology, Chicago, Illinois 60616

Received September 2, 2005; Revised Manuscript Received October 26, 2005

ABSTRACT: Expression of the PSR132 protein from *Dianthus caryophyllus* (carnation, clover pink) is induced in response to ethylene production associated with petal senescence, and thus the protein is named petal death protein (PDP). Recent work has established that despite the annotation of PDP in sequence databases as carboxyphosphoenolpyruvate mutase, the enzyme is actually a C–C bond cleaving lyase exhibiting a broad substrate profile. The crystal structure of PDP has been determined at 2.7 Å resolution, revealing a dimer-of-dimers oligomeric association. Consistent with sequence homology, the overall α/β barrel fold of PDP is the same as that of other isocitrate lyase/PEP mutase superfamily members, including a swapped eighth helix within a dimer. Moreover, Mg^{2+} binds in the active site of PDP with a coordination pattern similar to that seen in other superfamily members. A compound, covalently bound to the catalytic residue, Cys144, was interpreted as a thiohemiacetal adduct resulting from the reaction of glutaraldehyde used to cross-link the crystals. The Cys144-carrying flexible loop that gates access to the active site is in the closed conformation. Models of bound substrates and comparison with the closed conformation of isocitrate lyase and 2-methylisocitrate lyase revealed the structural basis for the broad substrate profile of PDP.

The glyoxylate cycle was discovered in bacteria (1) as a means of converting two-carbon compounds for synthesis of other cell constituents. The same cycle was subsequently found in plants to convert acetyl-CoA from fat degradation into succinate to fuel the tricarboxylic acid cycle (2). The glyoxylate cycle operates in all cells that have the capacity to convert acetate to carbohydrates, including bacteria, plants, fungi, and animals. A key enzyme of the glyoxylate cycle is isocitrate lyase (ICL)¹ which cleaves isocitrate into succinate and glyoxylate. A reaction analogous to that catalyzed by ICL occurs in the metabolism of propionyl-CoA via the 2-methyl isocitrate cycle (3). The corresponding 2-methylisocitrate lyase (MICL) cleaves 2-methylisocitrate into succinate and pyruvate. The amino acid sequences of the two enzymes are homologous and belong to a protein superfamily that also includes phosphoenolpyruvate mutase (PEPM),

carboxyphosphoenolpyruvate mutase (carboxyPEPM), phosphoenolpyruvate hydrolase, oxaloacetate acetylhydrolase, and ketopantoate hydroxymethyl transferase. Even though the reactions catalyzed by these enzymes are different (Figure 1), they are all proposed to proceed via oxyanion intermediates and/or transition states derived from α -hydroxy- or α -ketocarboxylate substrates (4). The functional diversity within the ICL/PEPM superfamily appears to be even greater than we presently know, because sequence analysis has revealed a number of uncharacterized proteins with potentially different activities (5). The members of the superfamily with known structures, which include PEPM, ICL, MICL, and ketopantoate hydroxymethyltransferase, share a common overall fold and catalytic scaffold (5–15).

One of the members of the ICL/PEPM superfamily is the 318-amino acid residue PSR132 protein encoded by a gene of *Dianthus caryophyllus* (carnation, clover pink), which is induced in response to ethylene production associated with petal senescence (programmed cell death) (16), and hence, the gene product is named petal death protein (PDP). In the preceding paper, we report that this protein is a homotetramer of 34 kDa subunits (17). A BLAST sequence search (18) on the nonredundant sequence database reveals that two homologues (possibly paralogs) of PDP with 64 and 57% identical sequences (including 289 aligned residues) exist in *Arabidopsis thaliana*. All other ICL/PEPM superfamily homologues exhibit a level of sequence identity with PDP of $\leq 44\%$. The closest amino acid sequence of a functionally characterized enzyme is that of carboxyPEPM from *Streptomyces hygroscopicus* (19), which is 39% identical to that of PDP. The current annotation of PDP as a putative carboxyPEPM (Swiss-Prot entry Q05957) was based solely

[†] Supported by NSF Grant MCB9813271 (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 W-31-109-ENG-38.

[‡] The coordinates have been deposited in the Protein Data Bank (entry 1ZLP).

* To whom correspondence should be addressed: Center for Advanced Research in Biotechnology, 9600 Gudelsky Dr., Rockville, MD 20850. Telephone: (301) 738-6245. Fax: (301) 738-6255. E-mail: osnat@carb.nist.gov.

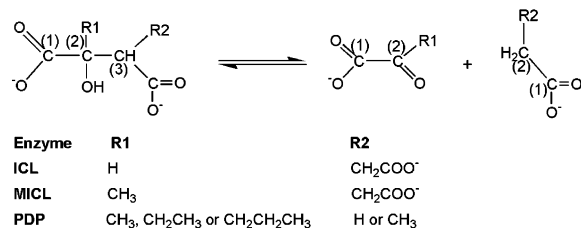
[§] University of Maryland Biotechnology Institute.

^{||} University of New Mexico.

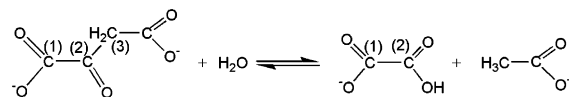
[⊥] Illinois Institute of Technology.

¹ Abbreviations: PDP, petal death protein; PEP, phosphoenolpyruvate; MICL, 2-methylisocitrate lyase; ICL, isocitrate lyase; PEPM, phosphoenolpyruvate mutase; OXAHYD, oxaloacetate acetylhydrolase; IPTG, isopropyl β -D-thiogalactopyranoside; DTT, dithiothreitol; PEG, polyethylene glycol; MAD, multiwavelength anomalous diffraction; SeMet, selenomethionine; rmsd, root-mean-square deviation.

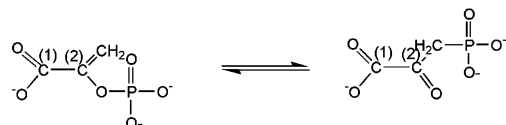
Lyases



OAH



PEP mutase



P-pyr hydrolase



CPEP mutase

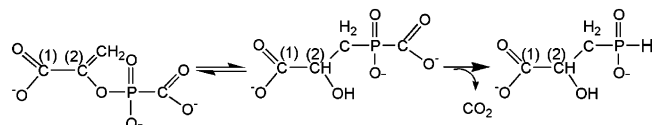


FIGURE 1: Reactions catalyzed by ICL/PEP mutase superfamily members.

on the sequence comparison. The work described in the preceding paper showed that PDP is not carboxyPEPM (17). Rather, in the presence of a divalent metal ion such as Mg^{2+} , the PDP catalyzes cleavage of the C(2)–C(3) bond in (2R)-alkyl malate derivatives to form alkyl carboxylates and α -ketocarboxylates as products (Figure 1). In addition, the PDP catalyzes the same reaction as oxaloacetate acetylhydrolase (20, 21), a fungal enzyme which occupies the same branch on the ICL/PEPM superfamily phylogeny tree (5). This particular C(2)–C(3) cleavage reaction, which produces oxalate rather than an α -keto acid, is likely to occur from the oxaloacetate hydrate and proceeds through the same mechanistic pathway followed by the other α -hydroxymalate substrates. In contrast to the substrates described above, the PDP shows only the slightest activity toward the (2R,3S)-isocitrate and (2R,3S)-2-methyl isocitrate, the substrates of ICL and MICL, respectively (17).

The crystal structure determination of PDP complements the recent functional studies. The structure reveals the active site architecture, the relationship to other superfamily members of known structure, and the basis for the broad profile substrate recognition.

MATERIALS AND METHODS

Cloning, Expression, and Purification. Recombinant PDP was prepared as described elsewhere (17). Selenomethionine

(SeMet) protein was prepared from the *Escherichia coli* B834(DE3)plysS cells transformed with pET3a-PSR132 plasmid DNA. The cells were grown in Growth Medium X (22) containing 50 $\mu\text{g}/\text{mL}$ carbenicillin to an OD_{600} of ~ 0.8 , and then induced with 0.4 mM IPTG for 5 h at 35 °C. The protein purification procedures were the same as those used to purify the wild-type protein (17), except that buffer A [50 mM triethanolamine (pH 7.5), 5 mM MgCl_2 , and 0.5 mM DTT] was modified to contain 10 mM DTT.

Crystallization and Structure Determination. PDP crystals were grown by vapor diffusion in hanging drops at 4 °C. Drops containing 1:1 volumes of protein in buffer A and reservoir solutions were equilibrated against the reservoir solution consisting of 0.1 M Tris-HCl (pH 7.5) and 12% PEG 20000. The crystals belong to space group $P3_121$ with the following unit cell parameters: $a = b = 156.2 \text{ \AA}$ and $c = 75.1 \text{ \AA}$. There are two protein monomers in the asymmetric unit. The calculated crystal specific volume, V_M , is $3.9 \text{ \AA}^3/\text{Da}$, corresponding to a solvent content of 68%. Despite much effort, a protocol to flash-cool the crystals was not found; therefore, they were cross-linked with glutaraldehyde. A drop of 2 μL of glutaraldehyde was placed on a microbridge (Hampton) and allowed to diffuse for 6 h at 4 °C. Then the crystal was transferred into a mother liquor containing 20% PEG 20000 in 0.1 M Tris-HCl (pH 7.5), soaked overnight, and flash-cooled in liquid propane. The same procedure was used for the wild-type and SeMet-containing protein crystals.

X-ray data were collected at 100 K at IMCA-CAT beamline 17-ID at the Advanced Photon Source (Argonne, IL) equipped with an ADSC CCD detector. The data from a SeMet protein crystal were measured at the wavelengths corresponding to the peak and the inflection point of the Se absorption edge, as well as at a high-energy remote wavelength. All data were processed with HKL2000 (23). Statistics are given in Table 1.

The structure was determined at 3.4 $\text{\AA}$ resolution by the multiwavelength anomalous diffraction (MAD) method, exploiting the Se absorption edge. Six of 18 potential selenium sites were located by the Shake-and-Bake method (24) and were used for phasing with MLPHARE (25) as implemented in CCP4 (26). Interestingly, all six sites corresponded to one of the two independent protein molecules in the asymmetric unit, although in retrospect both molecules have similar temperature factors. The phases were improved by density modification using DM (27), as implemented in CCP4, which included solvent flattening, histogram matching, and phase extension to 2.7 $\text{\AA}$ resolution using the wild-type protein crystal data. The resulting electron density map was still of poor quality, but was sufficient for locating several α -helices and part of the β -sheet. The phases were further improved by applying 2-fold noncrystallographic averaging which was based on the symmetry exhibited by the secondary structure units. This partial model enabled an approximate placing of the PEPM structure (PDB entry 1MUM) in the electron density, which was then used to guide the model building. The atomic model of PDP was completed over several cycles of manual modifications using O (28) and automatic refinement using REFMAC (29). The TLS option was activated at the end of refinement, thus reducing the working R -factor and R_{free} by almost 2%. Each protein monomer was treated as a TLS rigid group. Water molecules were added to the model on the basis of the $F_o - F_c$

Table 1: X-ray Data and Refinement Statistics for the Wild-Type and SeMet Protein Crystals

	wild type	peak	inflection	remote
wavelength (Å)	1.2755	0.9792	0.9799	0.9724
resolution (Å)	2.7	3.4	3.4	3.4
no. of unique reflections	26327 (2104) ^a	13833	13971	13725
completeness (%)	90.1 (66.7) ^a	91.9	92.8	91.2
redundancy	10.2 (9.0) ^a	5.2	5.1	5.2
R_{merge}^b	0.056 (0.215) ^a	0.062	0.064	0.067
$\langle I/\sigma(I) \rangle$	41.5 (5.9) ^a	22.5	19.4	19.2
R_{cryst}^c	0.195			
R_{free}^d	0.250			
no. of protein atoms	4,326			
no. of Mg ²⁺ ions	2			
no. of thiohemiacetal	2			
glutaraldehydes				
no. of water molecules	69			
rmsd from ideal geometry				
bond lengths (Å)	0.011			
bond angles (deg)	1.42			
mean B -factors,	83			
model (Å ²)				
mean B -factors,	63			
Wilson plot (Å ²)				

^a Statistics for the high-resolution shell (2.8–2.7 Å) are in parentheses. ^b $R_{\text{merge}} = \sum_{hkl} (\sum_j |I_j - \langle I \rangle|) / \sum_j I_j$ for equivalent reflections (Bijvoet pairs separated). ^c $R_{\text{cryst}} = \sum_{hkl} ||F_o| - |F_c|| / \sum_{hkl} |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. ^d R_{free} is computed for 7.8% of reflections that were randomly selected and omitted from the refinement.

difference Fourier electron density map when a protein hydrogen bond partner was available within 3.5 Å.

RESULTS AND DISCUSSION

Quality of the Model. The atomic model of PDP includes two protein monomers, each spanning residues 28–311 of the 318 amino acids listed in the Swiss-Prot database sequence (entry Q05957). Mass spectrometry together with N-terminal sequencing showed that the three N-terminal residues of the protein were cleaved (17). No electron density was associated with residues 4–27 at the N-terminus and residues 312–318 at the C-terminus; thus, they were assumed to be disordered and not included in the model. The electron density for the rest of the polypeptide backbone is clear and unambiguous. One residue in each molecule, Asp109, exhibits unusual main chain torsion angles: ψ and $\varphi = 48^\circ$ and -126° , respectively. An equivalent residue in the related structures of MICL and PEPM exhibits this conformation as well, and in all structures, this residue is associated with the magnesium binding site. Pro141 was modeled in the cis conformation. The model also includes 72 water molecules. A magnesium ion is bound in the active site of each molecule, and Cys144 is modified by glutaraldehyde to form the thiohemiacetal adduct. Cys144 might have reacted with the glutaraldehyde during the cross-linking step. A DTT molecule (present in the crystallization solution) was considered instead of the thiohemiacetal adduct but was rejected because its size and shape did not fit the electron density.

Refinement statistics are provided in Table 1. The crystallographic R -factor is 0.195 for the refined data in the resolution range of 10–2.7 Å, and R_{free} is 0.250 for reference reflections that were not included in the refinement. The atomic B -factors are high, ~ 80 Å² on average for both protein subunits, consistent with the low resolution of the

diffraction and the requirement to cross-link the crystals because of their intolerance to common flash-cooling protocols. The B -factor value estimated from the Wilson plot (30) is also high (63 Å²). Each protein subunit was treated as a rigid body group in the TLS refinement. The L -tensor is significantly non-zero for both subunits, indicating the anisotropic character of the crystal disorder.

Superposition of the two independent molecules in the asymmetric unit results in a root-mean-squares deviation (rmsd) of 0.35 Å for all 284 C α atoms. This value may be considered as an upper estimate of the coordinate error of the model. The average coordinate error based on the maximum likelihood statistics is 0.25 Å.

Molecular Topology. PDP has a topology of an (α/β)₈ barrel (Figure 2A). The notable deviation from the canonical α/β -barrel fold arises from the swapping of the eighth helix between two protein monomers (Figure 2B). This is a characteristic feature that distinguishes the ICL/PEPM superfamily. In addition to the α/β -barrel scaffold, there are four α -helical insertions in each subunit. The N-terminal α -helix (α_N) caps the barrel on its N-terminal side. The α -helix spanning residues 66–74 (α_A) follows the second β -strand and is located on the C-terminal edge of the barrel. A pair of helices at the C-terminus of the polypeptide chain (α_{C1} and α_{C2}) protrudes away from the protein subunit and forms intersubunit contacts by traversing the entrance of the active site of the neighboring molecule.

Two protein monomers form a tightly associated dimer (Figure 2B). Two such dimers form a tetramer with an almost exact 222 symmetry. The subunit interface area in the dimer is 3500 Å², which is 25% of the monomer surface. The swapped helices and the ensuing C-terminal helices, α_{C1} and α_{C2} , mediate much of this interface. The area buried between two dimers in the tetramer is 2600 Å², or 1300 Å² per subunit, which is not as extensive as the area buried in the dimer interface. In the crystal form presented here, the dimers obey exact crystallographic symmetry, whereas their association in the tetramer is pseudosymmetrical. A similar quaternary structure has been observed for other ICL/PEPM superfamily members regardless of the crystal form. Some of them (PDB entries 1O5Q, 1IGW, and 1F8I) were crystallized with the whole tetramer in the asymmetric part, whereas others (PDB entries 1PYM, 1MUM, 1OQF, and 1F61) have a tetramer sitting on the crystal dyad. In one case (ICL from *Aspergillus nidulans*, PDB entry 1DQU), the molecular 222 symmetry coincides with the crystal symmetry.

The DALI (31) search based on the monomer structure has yielded a number of α/β -barrel structures with the top hits corresponding to ICL/PEPM superfamily members. PDP is best superimposed on *E. coli* MICL (1XG3) and *Mytilus edulis* PEPM (1PYM) with rmsd's of 1.5 and 2.0 Å, respectively, for 280 common C α atoms. The overall structural similarity to these proteins is consistent with the sequence homology and establishes PDP as a member of the superfamily. We note that this level of structure similarity was insufficient to determine the PDP structure by molecular replacement using either MICL or PEPM as the search models.

PDP Active Site. The active site in α/β -barrel proteins is typically located at the C-terminal edge of the barrel (32). This is also true for PDP and the other enzymes of the ICL/

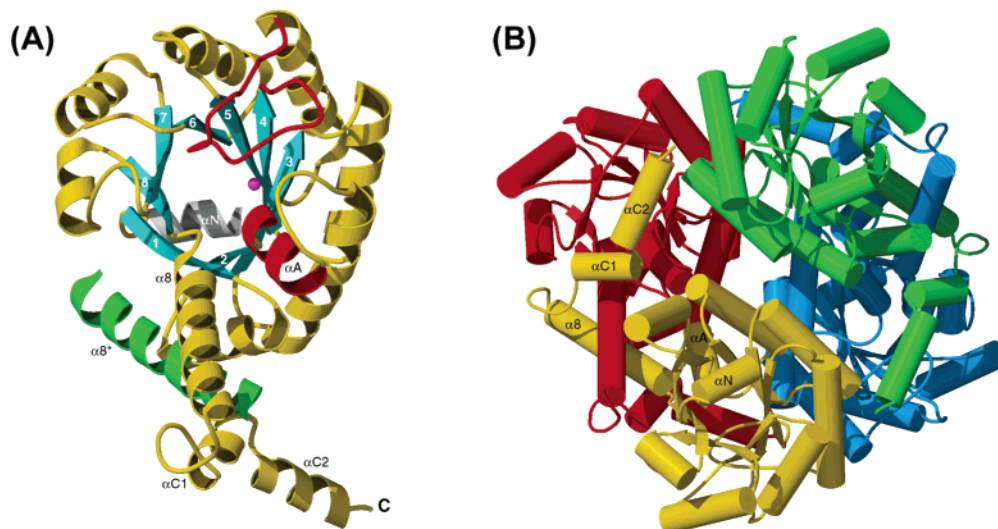


FIGURE 2: Structure of PDP. (A) Ribbon representation of the monomer fold. The eighth helix of the α/β barrel is swapped between subunits of a dimer. It is shown in a different color (green) for the neighboring molecule and is labeled with an asterisk. The α -helices and β -strands forming the α/β -barrel scaffold are numbered 1–8 sequentially. The N-terminal α -helix is labeled αN ; the C-terminal α -helices are labeled $\alpha C1$ and $\alpha C2$, and the short α -helix associated with the active site is labeled αA . The latter helix and the active site gating loop are colored red. Mg^{2+} is depicted as a magenta sphere. (B) Tetrameric association (dimer of dimers). Helices are depicted as cylinders. Each protein subunit is colored differently, so the swapped α -helices are highlighted.

PDP_D.caryophyllus	40	SVLMFGVQDALSAVVEKTGFHAAVSGYVSAAAM...LGLPDLGLLTTEVEATRITAAAP...NLCVVVDGTTGGG.GPLNVQRFIRE	124
MICL_E.coli	19	PLQIVGTINANHALLAQAGYQAIYLSGGGVAAGS...LGLPDLGLISTLDDVLTDIRRITDVC...SLPLLVDADIGFSSAFNVARTVKS	103
MICL_S.typhimurium	19	PLQIVGAINGANHALLAQAGYQAIYLSGGGVAAGS...LGLPDLGLISTLDDVLTDIRRITDVC...PLPLLVDADIGFSSAFNVARTVKS	103
ICL_M.Tuberculosis	66	WVNALGALTQNMVAQQVRAGLKATYLSGWQVAGDANLSGHTYPSQSLYPANSVPQVVRINNALQRAQIAKIEG...DTSVENWLAIVADGEAGFG.GALNVIELQKA	171
ICL_E.coli	66	YINSLGALTGGQALQAKAGIEAVYLSGWQVAGDANLAASMYPSQSLYPANSVPQVVRINNALQRAQIAKIEG...DTSVENWLAIVADGEAGFG.GALNVIELQKA	174
CPPM_S.hygrosopicus	18	ILVVPAYSADLSAKVQQAGFPVHMTSGTSAS...MLGLPDLGFTTSVSEQAINKNIVLTV...DVPVIMDADAGYG.NAMSVWRATRE	101
OXAHYD_B.fuckeliana	39	LIVCPGVYDGLSARVALQVGFPALYMTGAGTTASR...LGMADLGIATQLSDMKHAEIANLDP...YGPPLIADMTGYG.GPLIIDKAVKS	124
PEPM_M.edulis	20	LEFMEAHNGLSARIVQAGFGKINGSCLSVSAQ...LGVRSNEASWTQVVEVLEFMSDAS...DVPILLDADTGYG.NFNRRRLVRK	102
PPYRH_Variovorax.sp	16	LFTAMAHNPLVAKLAEQAGFGGINGSCLFELSAS...YAVPDLANILSMSTHLEMMRAITSVS...IPLIADITDGYG.NAVNVHYVVPQ	98

PDP_D.caryophyllus	125	LISAGAKGVFLSDQVWP.KKCGHMR..GKAVVPAAEEHALKIAAREAI..GDSDFFLVARTDA	182
MICL_E.coli	104	MLKAGAAGLHIEDQVGA.KRCGHRP..NKAIVSKEEMVDRIARAADAK..TDPDFVIMARTDA	161
MICL_S.typhimurium	104	IAKAGAAYLHIEDQVGA.KRCGHRP..NKAIVSKEEMVDRIARAADAK..TDPDFVIMARTDA	161
ICL_M.Tuberculosis	172	LIAAGVAGSHWEDQLASEKCKCHLGL..GKVLPTQHQIRLTLSARLAADVADVPTVVIARTDA	232
ICL_E.coli	175	MIEAGAAVAHFDQLASVKKCGHMG..GKVLVPTQEQIAIKLVAARLAADVGTGPTLLVARTDA	235
CPPM_S.hygrosopicus	102	FERVIGVYHLEDQVWP.KRCGHLE..GKRLISTEEMTKIEAAVEAR..EDEDFTIARTDA	159
OXAHYD_B.fuckeliana	125	YIRAGVAGFHLIEDQIQN.KRCGHQLQ..GKKVVPAAEYMYMIRAKAAKEAMNSDIVLARTDA	184
PEPM_M.edulis	103	LEDGRVAGACLEDKLF.P.KTNSLHDGRAQLADIEEFALKIKACKDSQ..TDPDFCIVARVEA	162
PPYRH_Variovorax.sp	99	YEAAGASAVMEEDKTFP.KDTSLETRDGRQELVRIEEFQGGKIAAATAAR..ADRFVVIARVEA	158

PDP_D.caryophyllus	191	GIRANLYKEAGADATFVEAPAN.VDELKEVS...AKTKGLRIANMIEGGKT...PLHTPEEFKEMGFHLIAHSLTAVYATARALVNIMKILKEGTRRDDLDQMATFSEFN	295
MICL_E.coli	170	ATERAQAYVEAGAEMLFPEA.ITELAMYRQFA...DAVQVPILANITEFGAT...PLFTTDELRSAHVAMALYELSAFRAMNRAAEHVNVLRQEGTKQSVIDTMQTRNELY	274
MICL_S.typhimurium	170	ALDRAQAYVDAGADMFLPEA.ITELSMYRRFA...DVAQVPILANITEFGAT...PLFTTDELRSAHVAMALYELSAFRAMNRAAEKVYTVLRQEGTKQNVIDIMQTRNELY	274
ICL_M.Tuberculosis	269	CIARAKAYAP.FADLIWNETSTPDLEAARFSEAVKAEYPPQMLAYNCSPS.FNWKHLDDATIAKFOKELAAAGFKFQITLAGHFNALNSMFDLAYGYAQNQMSAYVELQEREF...	386
ICL_E.coli	272	AISRGLAYAP.YADLIWNETSTPDLEAARFSEAVKAEYPPQMLAYNCSPS.FNWKHLDDATIAKFOKELAAAGFKFQITLAGHFNALNSMFDLAYGYAQNQMSAYVELQEREF...	389
CPPM_S.hygrosopicus	168	AIRSRREYVAAGADCFLEAMLD.VEEMKRVK...DEIDAPLLANMVEGGKT...PWLITKLEISIGNLALYELSGWMAAAVSLRKLFTLRLKRDGVTGPKNLTIPKALF...	299
OXAHYD_B.fuckeliana	193	CIKRLKVAEMGADVGLLEGYTSKMAAKTVK...EFAPWPLLANMVEGGKT...PIITTKAQEMGFRIMLPSFAALAPALYAIQETFLRLKRDGVTGPKNLTIPKALF...	299
PEPM_M.edulis	172	ALKRAEAYRNAGADAILMHSKKADPSDIEAFMKAWN...QGPFVVIPTYK...YKPTTDFHFRDMGVSMVIWANHNLRSASVIAQQTQTKQYIDDDQLSVNVEDKIVSVKEIF	276
PPYRH_Variovorax.sp	168	AVRRGQAYEEAGADAILHSRQKTPDEILAFVKSWPQ...KVPLVLVPTAYP...QLTEADTAALSQVGIYVGNHAIARAAGVAVREVFARIRRDGGIREVDAAALPSVKEII	273

FIGURE 3: Structure-based multiple-sequence alignment of ICL/PEPM superfamily members with known structure and/or confirmed function. For brevity, segments remote from the active site region are omitted. Abbreviated protein names and the organism names are listed. Abbreviations: PDP, petal death protein; MICL, 2-methylisocitrate lyase; ICL, isocitrate lyase; CPPM, carboxyPEP mutase; OXAHYD, oxaloacetate acetylhydrolase; PEPM, PEP mutase; PPYRH, phosphonopyruvate hydrolase. For each block, the residue numbers at the beginning and end of the block are listed. Active site residues are highlighted in yellow. The catalytic cysteine and glutamic acid residues characteristic of the lyases are marked with asterisks. Residues involved in substrate discrimination are marked with squares.

PEPM superfamily. Residues forming the active site are located on all β -strands except $\beta 1$, and on two loops following strands $\beta 2$ (containing helix αA) and $\beta 4$. A number of crystal complexes of these proteins with inhibitors and reaction products have been reported (5, 6, 9, 11, 14). These studies have established the residues involved in catalysis and substrate binding. The multiple-sequence alignment shown in Figure 3 highlights these residues.

Because the activity of all members of the ICL/PEPM superfamily is Mg^{2+} -dependent (or Mn^{2+}), the PDP was crystallized in the presence of 5 mM Mg^{2+} . Indeed, the electron density map revealed the metal bound in the active site, 2.2 Å from the carboxylic group of Asp107 (Figure 4A). In other superfamily members, the metal coordinates four carboxylic groups of aspartic and glutamic acid residues

either directly or via water molecules. In the absence of any substrate or inhibitor, the residue that is coordinated directly to Mg^{2+} is that equivalent to either Asp107 or Asp109 of PDP. When the substrate or inhibitor binds, Mg^{2+} always interacts directly with the aspartic acid equivalent to Asp107. In complexes, Mg^{2+} has an octahedral coordination with two positions occupied by the carboxyl and hydroxyl oxygen atoms of pyruvate or glyoxylate or the corresponding atoms of inhibitors. Without bound ligand, two water molecules often replace these positions. The limited resolution of the diffraction of the PDP crystal did not allow an unambiguous positioning of water molecules in the coordination sphere of Mg^{2+} , although some residual density around Mg^{2+} suggests their presence in the structure. Moreover, the dispositions of Asp79, Asp107, Asp109, and Glu136 are

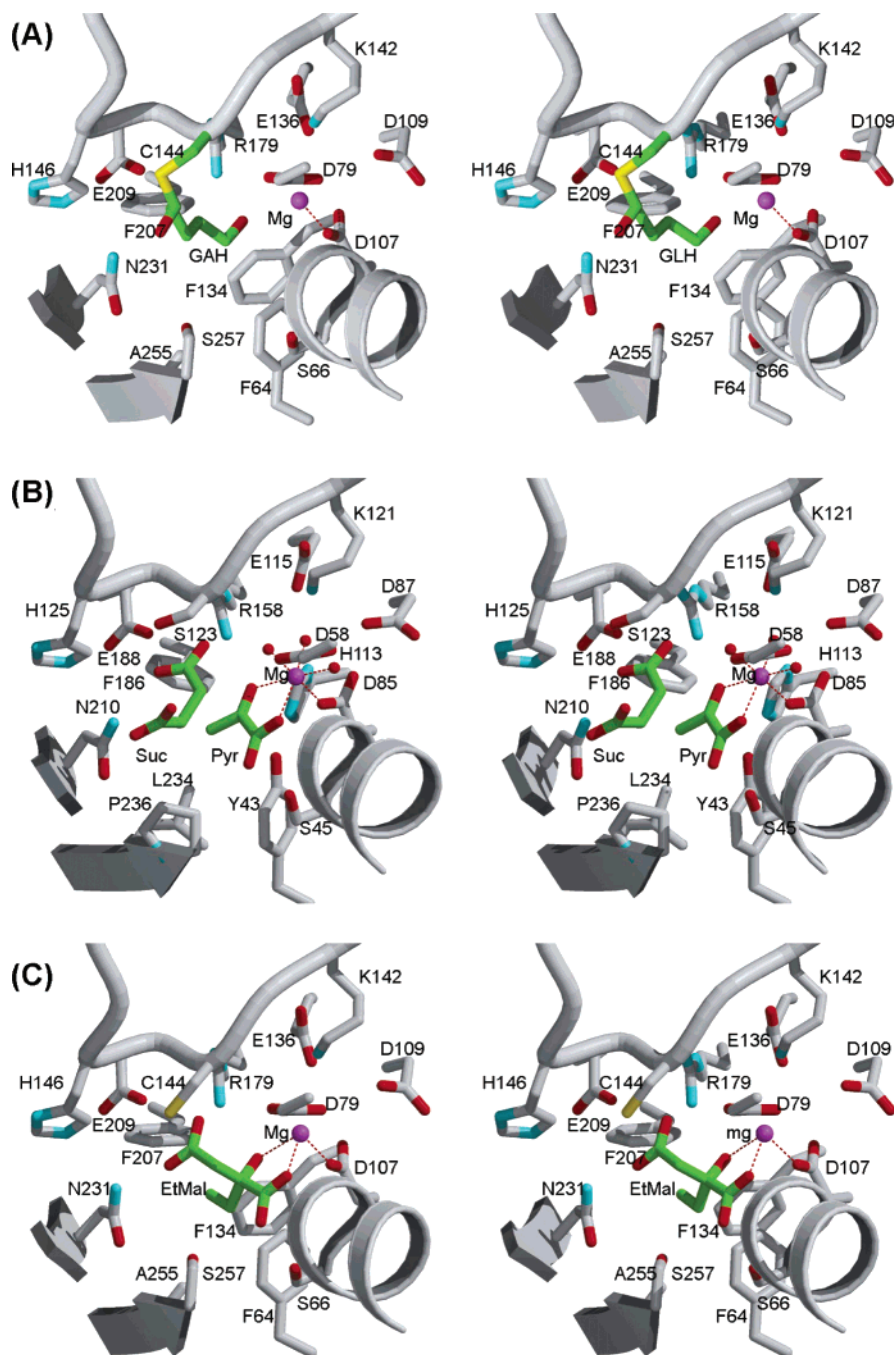


FIGURE 4: Stereoscopic view of the active site of PDP and MICL. (A) PDP with bound Mg^{2+} and glutaraldehyde in the form of thiohemiacetal (labeled GLH). The gating loop and the substrate carboxylate-binding α -helix, αA , are shown as a coil and a ribbon, respectively. Atomic colors for protein side chains are as follows: gray for carbon, red for oxygen, blue for nitrogen, and yellow for sulfur. For the glutaraldehyde moiety, carbon atoms are colored green. Mg^{2+} is depicted as a magenta sphere. (B) C123S mutant of MICL bound to Mg^{2+} and products, pyruvate (labeled Pyr) and succinate (labeled Suc). The coloring scheme is that from panel A. (C) A model of PDP bound to Mg^{2+} and 2-ethyl malate. The model was built on the basis of the structure of the MICL–product complex. Coloring scheme as in panel A.

similar to those of the Mg^{2+} -coordinating residues seen in other superfamily members (compare, for example, panels A and B of Figure 4).

Next to the metal site is the substrate binding pocket of PDP, which is occupied in both molecules in the asymmetric unit by an unknown ligand. The electron density protrudes from Cys144 on the loop following $\beta 4$, a residue analogous to the catalytic cysteine in ICL and MICL (Figure 3). The electron density was best fitted as a glutaraldehyde molecule that forms a thiohemiacetal adduct with the $S\gamma$ atom of Cys144 (Figure 4A). This interpretation appears to be most plausible because the crystals were treated with glutaralde-

hyde prior to flash-cooling, and no substrate or substrate analogue was added to the crystallization solution. The glutaraldehyde moiety is in an extended conformation so that the free aldehyde group farthest from the Cys144 $S\gamma$ atom interacts with Arg179 (equivalent to the catalytic Arg158 in MICL). Superposition of the PDP structure on the MICL enzyme–product complex (5) shows that the glutaraldehyde is located where the MICL products bind (Figure 4A,B).

Cys144 is located on a flexible loop that follows strand $\beta 4$, which in ICL, MICL, and PEPM adopts either an “open” or a “closed” conformation depending on the absence or presence of the substrate or inhibitor. The open conformation

allows substrate to bind and products to be released. The closed conformation is essential for active site desolvation and catalysis. In the PDP structure, the loop (residues 138–153) adopts a closed conformation that is most similar to that seen in MICL (5). In the closed conformation, the gating loop is “locked” by additional interactions with the C-terminal segment (α C1 and α C2) of the neighboring subunit that traverses the entry to the active site (Figure 2B). Closure of the loop in the presence of the Cys144/glutaraldehyde thiohemiacetal is reminiscent of the loop closure associated with the modification of the ICL catalytic cysteine by bromopyruvate (9).

Catalytic Activity and Substrate Specificity. Phylogeny analysis of the ICL/PEPM superfamily revealed several clusters of closely related proteins within the superfamily (5). PDP belongs to a cluster that also includes carboxyPEPM (sequence 40% identical to that of the *Streptomyces viridochromogenes* enzyme) and oxaloacetate acetylhydrolase (sequence 32% identical to that of the *Aspergillus niger* enzyme), two enzymes with very different activities. Because of the high level of sequence identity, PDP was proposed to be a carboxyPEPM (16), yet detailed activity studies have recently established that PDP has no carboxyPEPM or PEPM activity (17). Instead, the enzyme exhibits a Mg^{2+} -dependent lyase activity toward a range of substrates. The highest activity is with (2R,3S)-2-ethyl-3-methyl malate and with oxaloacetate, which we assume is bound as the C(2) hydrate (17). Several possible substrates were modeled into the PDP active site in place of the thiohemiacetal ligand to generate a picture of the enzyme–substrate complex [Figure 4C depicts the model of the complex with (2R)-ethyl malate]. These models were used to analyze the PDP catalytic scaffold in the context of the reported structures of other ICL/PEPM superfamily members.

PDP possesses the two catalytic residues that are conserved in ICL, MICL, carboxyPEPM, and oxaloacetate acetylhydrolase (Figure 3). Using the PDP amino acid numbering, these are Cys144 (the likely proton source for the acicarboxylate anion leaving group) and Glu209 [which in its protonated state shares the proton with the C(3)COO[−] group] (5, 9). In addition, the residues common to both the C–C and P–C bond forming and breaking superfamily members (with the exception of ketopantoate hydroxymethyltransferase) are conserved in PDP: Arg179 [which binds the C(3)COO[−] and C(2)O groups in the lyases, and the PO₃[−] and C(2)O groups in PEPM] and the four carboxylic acids associated with the Mg^{2+} binding site, Asp79, Asp107, Asp109, and Glu136 (Figures 3 and 4). The gating-loop lysine is also present (Lys142), implying that loop opening and closure play the same mechanistic role.

It is the active site architecture that underlies the differences in substrate specificity and catalytic efficiencies of PDP, MICL, and ICL. Indeed, much of the PDP broad substrate specificity reported in the preceding paper (17) can be rationalized in structural terms. Both ICL and MICL from *E. coli* exhibit strict specificity toward their substrate and high catalytic efficiency (k_{cat}/K_m values of 5.6×10^6 and $7.8 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$, respectively). PDP on the other hand exhibits minimal activity with the ICL and MICL substrates, (2R,3S)-isocitrate and (2R,3S)-2-methyl isocitrate ($k_{cat} \sim 1 \times 10^{-4} \text{ s}^{-1}$), and a modest but physiologically significant level of activity toward (2R,3S)-2-ethyl-3-methyl malate (k_{cat}

$= 8.4 \text{ s}^{-1}$, $K_m = 530 \mu\text{M}$, and $k_{cat}/K_m = 2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) and oxaloacetate ($k_{cat} = 2.7 \text{ s}^{-1}$, $K_m = 130 \mu\text{M}$, and $k_{cat}/K_m = 2 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$) (17). Replacing the C(2) ethyl group with a methyl or propyl group reduces both K_m and k_{cat} , leading to a catalytic efficiency approximately 10-fold lower than that with an ethyl group. On the other hand, (2R)-malate is not a substrate for PDP ($k_{cat} < 1 \times 10^{-5} \text{ s}^{-1}$), which underscores the importance of the C(2) substituent in productive binding. PDP does not act on isocitrate derivatives containing an acetyl group on C(3), and instead prefers a methyl group (17). Malate derivatives without substituents on C(3) are also substrates, although the catalytic efficiency is reduced [for example, $k_{cat}/K_m = 2 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ for (2R)-ethyl malate].

To explain the structural basis for discrimination between isocitrate and 2-methyl isocitrate by ICL and MICL, Grimm et al. proposed that a triad of residues [Phe186, Leu234, and Pro236 in MICL, and the equivalent residues, Trp283, Phe345, and Thr347, in ICL (Figure 3)] forms a hydrophobic pocket which in MICL can accommodate a methyl group but in ICL can only accommodate a hydrogen atom (10). With the structure determination of MICL bound to the product and to the isocitrate inhibitor, we realized that the size of the hydrophobic pocket depends to a large extent on the backbone placement of these residues, and therefore, the residues need not necessarily be conserved within each enzyme group (5). For PDP, the “specificity triad” comprises Phe207, Ala255, and Ser257 (Figure 3). Modeling of 2-ethyl malate in the PDP active site (Figure 4C) and superposition of MICL, ICL, and PDP revealed that the PDP hydrophobic pocket is the largest of the three, and thus can accommodate an ethyl group at C(2). Propyl and methyl groups also fit, yet the kinetic characterization suggests that the PDP pocket is optimal for the size of an ethyl group. Obviously, hydrogen at C(2) could be accommodated as well, but the empty space may result in an unproductive binding mode, similar to the unproductive binding mode of (2R,3S)-isocitrate in MICL (5). This is consistent, although not dictated by, with the low rate of catalysis of (2R)-malate reported in ref 17.

The PDP pocket is more hydrophobic than those of MICL and ICL because the Tyr-His pair (Tyr43 and His113 in *E. coli* MICL and Tyr90 and His281 in *M. tuberculosis* ICL) is replaced with two phenylalanine residues in PDP, Phe64 and Phe134 (Figure 3). The Phe134 side chain adopts an alternate conformation with respect to that of the analogous histidine side chain (Figure 4B,C), thus providing a larger binding pocket.

The PDP specificity triad can accommodate a hydroxyl group at C(2), as in oxaloacetate hydrate, which is appropriately oriented to form a hydrogen bond with the Ser257 hydroxyl group. Alternatively, a water molecule might occupy the site adjacent to Ser257, in which case it could mount a nucleophilic attack on C(2) following the C(2)–C(3) bond cleavage to produce oxalate. Whether the oxaloacetate binds in the form of C(2) hydrate or as a nonhydrated compound is related to the extent of similarity between the lyase and oxaloacetate acetylhydrolase catalytic mechanisms. The solvent structure is uncertain because of the low resolution of the PDP structure determination, which prevents us from distinguishing between the two possibilities. Regardless of the mechanism, the catalytic activity of PDP toward hydrolytic cleavage of oxaloacetate is within 1 order of

magnitude of that of the *Botryotinia fuckeliana* oxaloacetate acetylhydrolase ($k_{\text{cat}} = 15.5 \pm 0.2 \text{ s}^{-1}$, $K_{\text{m}} = 65 \pm 3 \text{ }\mu\text{M}$, and $k_{\text{cat}}/K_{\text{m}} = 2.4 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$), an enzyme that, in contrast to PDP, exhibits minimal activity toward (2R,3S)-2-ethyl-3-methyl malate (17). The fungal oxaloacetate acetylhydrolase also contains a serine residue at the same position as Ser257 of PDP. However, the use of an isoleucine in the place of Ala255 of the PDP may reduce the size of the binding pocket, thereby impairing productive binding of (2R,3S)-2-ethyl-3-methyl malate.

The C(3) acetate group of isocitrate and 2-methyl isocitrate renders these compounds poor substrates of PDP. Comparison of the PDP structure with those of MICL and ICL shows that the acetate group forms favorable electrostatic interactions with protein groups [Arg241, Arg270, and Thr212 in MICL and His193, Asn313, Ser315, Ser317, and Thr347 in ICL (Figure 3)]. In PDP, the position of the hydrophobic side chains of Ile233, Leu258, and Tyr68 cannot be adjusted because of their crowded environment, and these side chains would clash with a substrate's C(3) acetate group. Consequently, only a methyl group or a hydrogen atom can be accommodated. Comparison of the kinetic constants measured for (2R)-methyl malate ($k_{\text{cat}} = 0.2 \text{ s}^{-1}$, $K_{\text{m}} = 290 \text{ }\mu\text{M}$, and $k_{\text{cat}}/K_{\text{m}} = 7 \times 10^2 \text{ M}^{-1} \text{ s}^{-1}$) and (2R,3S)-dimethyl malate ($k_{\text{cat}} = 0.04 \text{ s}^{-1}$, $K_{\text{m}} = 26 \text{ }\mu\text{M}$, and $k_{\text{cat}}/K_{\text{m}} = 1 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$) indicates that the C(3) methyl might facilitate binding at the expense of catalysis [the opposite trend than that seen in increasing the size of the C(2) substituent (17)]. The absence of PDP substrate activity in the (3S)-isopropyl malate and (3S)-butyl malate is suggestive of too much bulk at C(3) and/or not enough bulk at C(2) in these ligands (17).

From inspection of the PDP active site, we do not gain a clear idea of what changes might be made in the malate scaffold to improve substrate activity with the PDP. The 100–1000-fold lower catalytic efficiency of PDP compared with those of MICL and ICL remains a puzzle. The replacement of a Tyr-His pair found in all other lyases with Phe64 and Phe134 in PDP might contribute to its low activity (Figure 3). The tyrosine hydroxyl group interacts with the substrate, and we previously proposed that the Tyr-His pair may assist proton shuttle (5). It would be interesting to mutate these residues in the future and to determine the effect of the mutations on catalysis.

It is possible the true substrate of PDP has not been identified, although the functional groups of such a substrate must be similar to those of substrates identified in the substrate screen (17). Nevertheless, in contrast to ICL and MICL, PDP exhibits both the broad substrate profile and the lower enzymatic efficiency typical of a scavenger enzyme. Programmed petal death involves numerous degradation processes that enable accumulation of nutrients for seed development. The demonstration that PDP catalyzes C–C bond cleavage of several substrates leading to products commonly used in metabolic pathways (i.e., pyruvate, 2-ketobutarate, acetate, and propionic acid), together with the insight about the origin of broad specificity provided by the structural work, suggests that this lyase might by design function in metabolite recycling. In addition, the PDP-catalyzed conversion of oxaloacetate to oxalate might be important for the regulation of Ca^{2+} levels in senescent cells because oxalate plays a role in sequestering and storing Ca^{2+} in plants.

ACKNOWLEDGMENT

We thank the staff of IMCA-CAT at the APS for help during data collection. The IMCA-CAT facility was supported by the companies of the Industrial Macromolecular Crystallographic Association through a contract with the Illinois Institute of Technology. We are grateful to the Keck Foundation for providing funds to purchase of X-ray equipment at the Center for Advanced Research in Biotechnology.

REFERENCES

- Kornberg, H. L., and Krebs, H. A. (1957) Synthesis of cell constituents from C2-units by a modified tricarboxylic acid cycle, *Nature* 179, 988–91.
- Kornberg, H. L., and Beevers, H. (1957) The glyoxylate cycle as a stage in the conversion of fat to carbohydrate in castor beans, *Biochim. Biophys. Acta* 26, 531–7.
- Tabuchi, T., and Serizawa, N. (1975) A hypothetical cyclic pathway for the metabolism of odd-carbon *n*-alkanes or propionyl-CoA via seven-carbon tricarboxylic acids in yeasts, *Agric. Biol. Chem.* 39, 1055–61.
- Seidel, H. M., and Knowles, J. R. (1994) Interaction of inhibitors with phosphoenolpyruvate mutase: Implications for the reaction mechanism and the nature of the active site, *Biochemistry* 33, 5641–6.
- Liu, S., Lu, Z., Han, Y., Melamud, E., Dunaway-Mariano, D., and Herzberg, O. (2005) Crystal structures of 2-methylisocitrate lyase in complex with product and with isocitrate inhibitor provide insight into lyase substrate specificity, catalysis and evolution, *Biochemistry* 44, 2949–62.
- Huang, K., Li, Z., Jia, Y., Dunaway-Mariano, D., and Herzberg, O. (1999) Helix swapping between two α/β barrels: Crystal structure of phosphoenolpyruvate mutase with bound Mg^{2+} -oxalate, *Struct. Folding Des.* 7, 539–48.
- 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. Folding Des.* 8, 349–62.
- Britton, K. L., Abeyasinghe, 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. D* 57, 1209–18.
- 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, 663–8.
- 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, 609–21.
- 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.
- 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* 11, 753–64.
- 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 hydroxymethyltransferase complexed with ketopantoate and Mg^{2+} , solved by locating 160 selenomethionine sites, *Structure* 11, 985–96.
- Liu, S., Lu, Z., Jia, Y., Dunaway-Mariano, D., and Herzberg, O. (2002) Dissociative phosphoryl transfer in PEP mutase catalysis: Structure of the enzyme/sulfolpyruvate complex and kinetic properties of mutants, *Biochemistry* 41, 10270–6.
- Liu, S., Lu, Z., Han, Y., Jia, Y., Howard, A., Dunaway-Mariano, D., and Herzberg, O. (2004) Conformational flexibility of PEP mutase, *Biochemistry* 43, 4447–53.

16. 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, 719–24.
17. Lu, Z., Feng, X., Song, L., Han, Y., Kim, A., Herzberg, O., Woodson, W. R., Martin, B. M., Mariano, P. S., and Dunaway-Mariano, D. (2005) Diversity of Function in the Isocitrate Lyase Enzyme Superfamily: The *Dianthus caryophyllus* Petal Death Protein Cleaves α -Keto and α -Hydroxycarboxylic Acids, *Biochemistry* 44, 16365–16376.
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, 3389–402.
19. Pollack, S. J., Freeman, S., Pompliano, D. L., and Knowles, J. R. (1992) Cloning, overexpression and mechanistic studies of carboxyphosphoenolpyruvate mutase from *Streptomyces hygroscopicus*, *Eur. J. Biochem.* 209, 735–43.
20. Lenz, H., Wunderwald, P., and Eggerer, H. (1976) Partial purification and some properties of oxalacetase from *Aspergillus niger*, *Eur. J. Biochem.* 65, 225–36.
21. Ruijter, G. J., van de Vondervoort, P. J., and Visser, J. (1999) Oxalic acid production by *Aspergillus niger*: An oxalate-non-producing mutant produces citric acid at pH 5 and in the presence of manganese, *Microbiology* 145 (Part 9), 2569–76.
22. Gerchman, S. E., Graziano, V., and Ramakrishnan, V. (1994) Expression of chicken linker histones in *E. coli*: Sources of problems and methods for overcoming some of the difficulties, *Protein Expression Purif.* 5, 242–51.
23. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276, 307–26.
24. Weeks, C. M., and Miller, R. (1999) Optimizing Shake-and-Bake for proteins, *Acta Crystallogr. D* 55 (Part 2), 492–500.
25. Otwinowski, Z. (1991) in *Isomorphous replacement and anomalous scattering* (Wolf, W., Evans, P. R., and Leslie, A. G. W., Eds.) pp 80–6, Daresbury Laboratory, Warrington, U.K.
26. Collaborative Computational Project No. 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D* 50, 760–3.
27. Cowtan, K. (1994) “DM”: An automated procedure for phase improvement by density modification, *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* 31, 34–8.
28. 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. A* 47, 110–9.
29. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood methods, *Acta Crystallogr. D* 53, 240–55.
30. Wilson, A. J. C. (1949) The probability distribution of X-ray intensities, *Acta Crystallogr.* 2, 318–21.
31. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices, *J. Mol. Biol.* 233, 123–38.
32. Brändén, C.-I. (1980) Relation between structure and function of α/β -proteins, *Q. Rev. Biophys.* 13, 317–38.

BI051779Y