PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

STRUCTURE NOTE

Crystal structure of a putative methylmalonyl-coenzyme a epimerase from *Thermoanaerobacter tengcongensis* at 2.0 Å resolution

Liang Shi,[†] Pu Gao,[†] Xiao-Xue Yan, and Dong-Cai Liang[*]

National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, People's Republic of China

Key words: methylmalonyl-coenzyme A; MMCE; X-ray crystallography.

INTRODUCTION

Methylmalonyl-CoA is an important metabolic intermediate that exists in several key degradative and biosynthetic pathways.[1] The pathway responsible for the degradation of branched amino acids and odd chain fatty acids involves three steps. In the first step, a biotin-dependent carboxylation produces the *S*-epimer of methylmalonyl-CoA. The third step, the B_{12}-dependent methylmalonyl-CoA mutase converts the *R*-epimer of methylmalonyl-CoA to succinyl-CoA.[1] This key epimerization between the two steps is carried out by the essential enzyme methylmalonyl-CoA epimerase (MMCE: EC number 5.1.99.1), which catalyzes the conversion of (*2S*)-methylmalonyl-CoA to (2R)-methylmalonyl-CoA. In prokaryotes, MMCE participated in autotrophic CO_2 fixation via the 3-hydroxypropionate pathway, in propionate fermentation, the regeneration of glyoxylate and in the biosynthesis of polyketide antibiotics.[2–6] In animals, defects of this pathway can result in a buildup of methylmalonic acid, which causes severe acidosis and also damages the central nervous system.[7,8] In this study, we determined the crystal structure of TTE0360 from *Thermoanaerobacter tengcongensis* to 2.0 Å resolution using single-wavelength anomalous dispersion (SAD) by short cryo-soaking with a high concentration of potassium iodide.[9,10] In GenBank database, the gene *tte0360* is predicted to be glyoxalase I, a member of glyoxalase superfamily. However, the TTE0360 structure is very different than the released structures of glyoxalase I. A search for the 3D homologs using DALI[11] shows that TTE0360 has significant structural similarity with the MMCE from *Propionibacterium shermanii*,[12] which is also a member of glyoxalase superfamily. The fold type of the TTE0360 molecule belongs to the vicinal-oxygen-chelate superfamily (VOC).[13] The similarities of structure folds, dimer formation, and active sites between TTE0360 and MMCE, suggest that the protein TTE0360 may be the MMCE from *T. tengcongensis*.

MATERIALS AND METHODS

Cloning, expression, and purification

The *tte0360* gene (NCBI Accession No. AE008691) was amplified by PCR from *Thermoanaerobacter tengcongensis* genomic DNA.[14] The PCR product was subcloned into pHAT-2 vector by *Nco*I and *Hind* III restriction sites. This plasmid was transformed into BL21 (DE3) *Escherichia coli* cells and the cells were grown in LB media supplemented with 100 mg/mL ampicillin to a log phase (OD_{600} = 0.6). The expression of MMCE was induced by the addition of 0.5 m*M* IPTG at 37°C, for 8 h. Cells were harvested and

Grant sponsor: National Program on key Basic Research Project (973 Program); Grant number: 2006CB806501; Grant sponsor: National Protein Project; Grant numbers: 2006CB910902, 2007CB914302.
[†]Liang Shi and Pu Gao contributed equally to this work.
*Correspondence to: Dong-Cai Liang, National Laboratory of Biomacromolecules, Institute of Biophysics, CAS 15 Datun Road, Chaoyang District, Beijing 100101, People's Republic of China. E-mail: dcliang@sun5.ibp.ac.cn
Received 25 May 2009; Revised 19 June 2009; Accepted 24 June 2009
Published online 7 July 2009 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/prot.22528

resuspended in buffer A (50 m*M* Tris, pH 8.0, and 500 m*M* NaCl), and were lysed by sonication. The lysate was clarified by centrifugation and purified by passage through a nickel-affinity column, and a further purification step was performed by size exclusion chromatography using a Superdex 200 column (Amersham). The pure protein was concentrated to 20 mg/mL for crystallization with buffer B (5 m*M* Tris, pH 8.0, and 300 m*M* NaCl).

Crystallization and diffraction data collection

The crystals of recombinant MMCE were grown at 20°C using the hanging-drop, vapor-diffusion method. Drops consisted of 2 μL of protein solution and 2 μL of mother liquor (0.1 *M* citric acid, pH 2.6, 1.6 *M* $(NH_4)_2SO_4$). Crystals suitable for X-ray diffraction studies were obtained after 7–8 days growth. Native and derivative Crystals were soaked in 2 *M* Li_2SO_4 for 2 min before data collection and were flash-frozen in liquid nitrogen. The native data were collected on Beamline 1W2B at the Beijing Synchrotron Radiation Facility using a Mar555 flatpanel detector. The derivative data were collected on a Rigaku FR-E X-ray generator and the Rigaku R-AXIS IV++ image plate detector. Data were integrated and scaled with HKL2000.[15] The statistics of data collection are summarized in Table I. The Matthews coefficient (V/M) of 2.0 $Å^3$ Da^{-1} for both crystals suggests the presence of a monomer in an asymmetric unit with an estimated solvent content of 39% (v/v).[16]

Structure determination and refinement

The heavy atom derivatives were obtained by soaking the native crystals in the mother liquid supplemented with 0.5 *M* KI for 2 min. Three iodine heavy atom sites with occupancy beyond 20% (54, 47, and 32% respectively) in each asymmetric unit were determined using PHENIX.[17] By using the initial phase from PHENIX, automatic model building was performed with ARP/wARP[18] at 2.0 Å resolution. Model refinement was performed in CNS,[19] and the program COOT[20] was used for inspection and manual improvement of the model. Within the resolution range 8–2.0 Å, the native structure was refined to a final R_{work} = 21.8% and R_{free} = 24.7%, and confirmed to have good stereochemistry from the Ramachandran plot calculated by PROCHECK.[21] The statistics of the refinement and the stereochemistry of the final model are summarized in Table I. The coordinates and structure factors of TTE0360 were deposited into RCSB Protein Data Bank with accession code 3GM5.

Table I
Data Collection, Phasing and Refinement Statistics for TTE0360 Structures

	Native	I-derivative
Data collection statistics		
Wavelength (Å)	1.0000	1.5418
Resolution (Å)	20–2.0	20–2.2
The highest resolution shells (Å)	2.05–2.00	2.25–2.20
Space group	$P4_12_12$	$P4_12_12$
Unit cell parameters	a = 54.02 Å	a = 53.60 Å
	c = 104.80 Å	c = 105.66 Å
No. unique reflections	11119	9379
Redundancy	12.4 (9.0)	13.4 (10.2)
Completeness (%)	99.9 (99.4)	100 (99.2)
R_{merge} (%)	7.0 (24.5)	6.0 (12.6)
Mean $I/\sigma(I)$	37.0 (7.6)	36.1 (10.3)
Refinement statistics		
Resolution (Å)	8–2.0	
Total No. of reflections used	10849	
Reflections in working set	10318	
Reflections in test set	531	
Nonhydrogen atoms		
Protein	1250	
Water	129	
SO_4^{2-}	2	
Citric acid	1	
R_{work} (%)	21.8	
R_{free} (%)	24.7	
RMS deviation from ideal geometry		
Bond length (Å)	0.009	
Bond angle (°)	1.61	
Overall average B factor ($Å^2$)	26.6	
Main chain B factor ($Å^2$)	21.8	
Side chain B factor ($Å^2$)	28.5	
Water B factor ($Å^2$)	38.1	
Ramachandran plot		
Most favorable (%)	89.6	
Allowed (%)	9.7	
Generously allowed (%)	0.7	

RESULTS AND DISCUSSION

Monomer structure

The crystal structure of TTE0360 was determined to 2.0 Å resolution using the SAD method (Table I). The TTE0360 molecule exhibits a compact α/β structure with dimensions of about 51 × 35 × 36 Å [Fig. 1(A)]. The asymmetric unit includes a monomer with 159 amino acid residues. The final model includes 151 amino acid residues (150 amino acid residues coded by the *tte0360* gene and a serine from the His-tag), 1 citric acid molecule (CIT), 2 sulfate ions, and 129 water molecules. The fold type of TTE0360 monomer belongs to the VOC superfamily. One TTE0360 molecular is made up of two repeat modules, and each of these modules is a βαβββ motif. All the β strands of each module are arranged in the 1-4-3-2 order (such as β1-β4-β3-β2; β5-β8-β7-β6 in Figure 1(A,B), and the two insert α helixes are α1 and α4. The β sheets of the two modules pack edge-to-edge, and the first strand of the N-terminal module (β1) was determined to be antiparallel to the first strand of the C-terminal module (β5) [Fig. 1(B)]. These eight strands

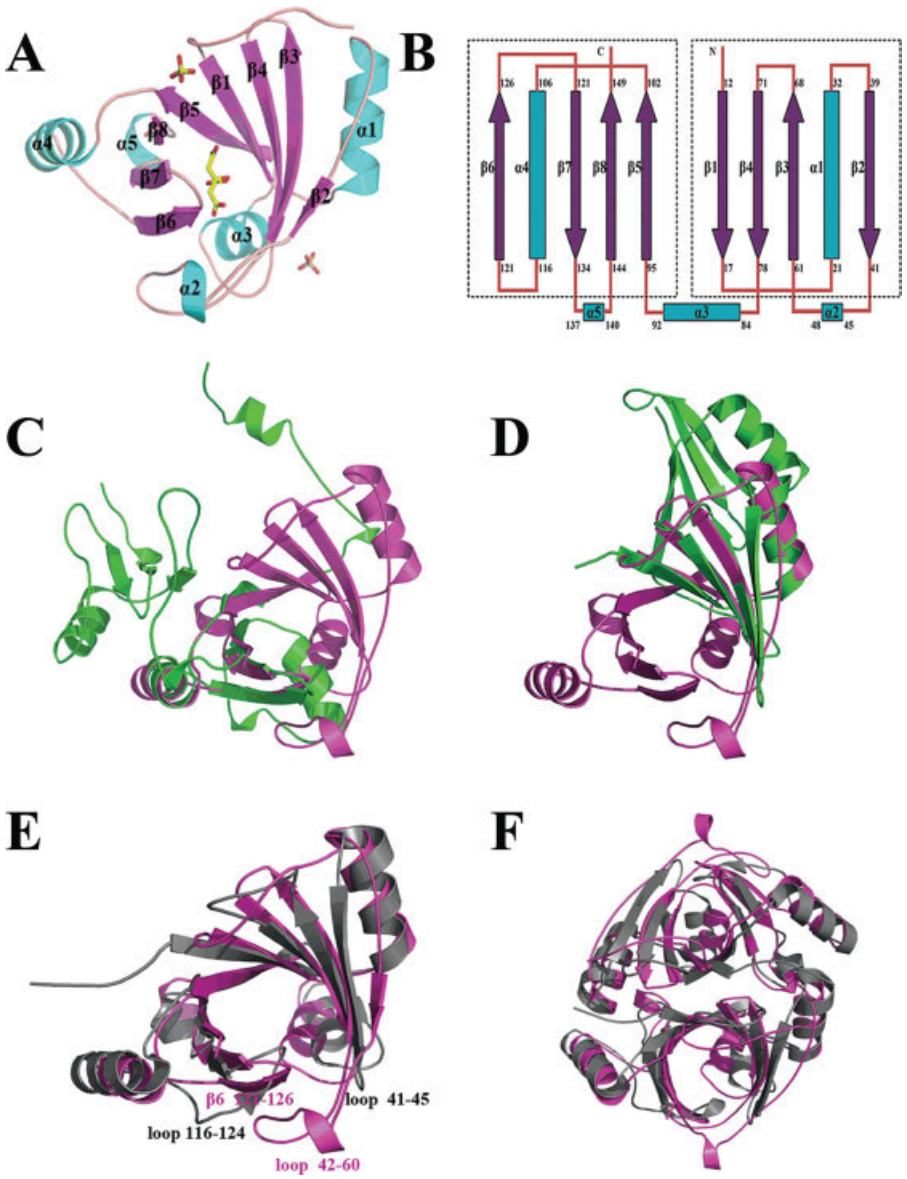

Figure 1

Crystal structure of TTE0360. **A**: The monomer fold of TTE0360 represented as a ribbon diagram. α-Helices, β-sheets, and loops are colored cyan, magenta, and light pink, respectively. The bound citric acid molecule and two sulfate ions are shown in stick form in yellow. **B**: The monomer fold of TTE0360 represented as a topology diagram, showing the two βαββββ modules. α-Helices, β-sheets, and loops are colored cyan, magenta, and light pink, respectively. **C**: Superimposition of the TTE0360 monomer (purple) and the human_Glo I monomer (green). **D**: Superimposition of the TTE0360 monomer (purple) and the *E.coli*_Glo I monomer (green). **E**: Superimposition of the TTE0360 monomer (purple) and the psMMCE monomer (gray). Some significant differences are shown in some of the loops. **F**: Superimposition of the TTE0360 dimer (purple) and the psMMCE dimer (gray).

were found to create a cleft to form the core of the TTE0360 protein monomer. There are other three α helixes, besides the two βαββββ motif: α2 is between β2 and β3; α5 is between β7 and β8; α3 is between β4 and β5 and links the two modules [Fig. 1(B)]. A long loop (Thr42-Arg60) between β2 and β5 in which a short helix (α2) is located was observed, and it was found to compose a region sticking out. This region and β6 of the C-terminal module were shown to seal the cleft from the side. A helix (α3) was found to be located in the bottom of the cleft, closing the cleft from the back. In addition, a citric acid molecular was found to be bound in the cleft of TTE0360 and was confirmed by omit map. This citric acid makes a number of well-defined hydrogen bonds with the nearby residues and a water molecular to form a stable network.

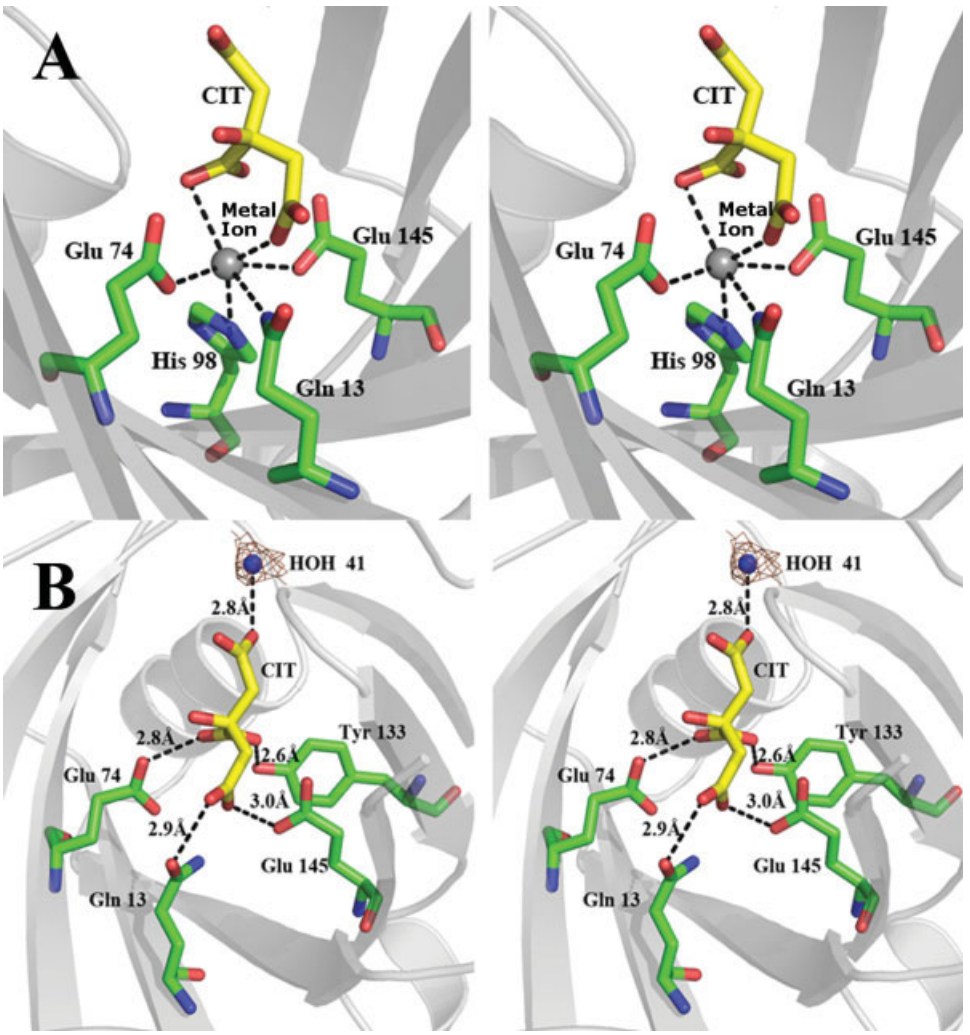

Figure 2

Active site of TTE0360. **A**: Stereo view of the metal ion binding site of TTE0360. The four conserved residues, the citric acid, and the modeled metal ion are colored green, yellow, and gray, respectively. The black dash lines represent the putative six-membered chelate ring. **B**: Stereo view of the citric acid binding site of TTE0360. The amino acid residues, the citric acid, and the water molecular are colored green, yellow, and blue respectively. The black dash lines represent the network of hydrogen bonds.

Based on primary sequence analysis in NCBI, the TTE0360 protein was predicted to be a member of the Glyoxalase superfamily, glyoxalase I (Glo I). Several structures of glyoxalase I have been solved, and the functions of these proteins have been proved by functional experiments, including human glyoxalase I,[22] mouse glyoxalase I,[23] *E. coli* glyoxalase I,[24] and *L. major* glyoxalase I.[25] However, using the released crystal structures of glyoxalase I as molecule replacement models, we could not solve the protein structure of TTE0360. We solved the structure using the SAD method. We found that the structure of TTE0360 and the earlier released structures of glyoxalase I are quite distinct. The three-dimensional structural comparison Z score of TTE0360 and human Glo I, mouse Glo I, *E. coli* Glo I, *L.* major Glo I were 3.7, 3.6, 4.6, and 4.7 respectively. As shown in Figure 1(C,D), the structures of TTE0360 and glyoxalase I are very different. On the other hand, analysis by the program DALI showed significant similarities between TTE0360 and methylmalonyl-CoA epimerase (MMCE) from *Propionibacterium shermanii* (PDB id: 1j4c), giving a Z score of 12.7. The root mean square deviation between TTE0360 and MMCE from *P. shermanii* (psMMCE) was found to be 2.2 Å for 145 Cα atom pairs. As shown in Figure 1(E), the crystal structures of TTE0360 and psMMCE are similar in the protein core region with differences in some loops. A remarkable difference was found at the loop (Thr42-Arg60) in

TTE0360, which was much longer than the corresponding loop (Asn41-Gly45) in psMMCE. In psMMCE, two short loops (Asn41-Gly45 and Asp116-Gly124) were shown to seal the side of the active pocket. Because the two loops are remote from each other, they can only close the active pocket roughly. However, in TTE0360 the loop (Thr42-Arg60) and the β6 strand (Gly121-Phe126) are much closer, and Ala48 and Asn53 on the loop form hydrogen bonds with Asp125 and Ile121 on the β6 strand respectively. In TTE0360, the loop (Thr42-Arg60) and the β6 strand (Gly121-Phe126) close the active pocket compactly.

Dimer structure

The dimerizations of TTE0360 and psMMCE are very similar. The psMMCE is a dimer in both solution and crystal.[12] The data of size exclusion chromatography and analytical ultracentrifugation show that TTE0360 is also a dimer in solution. Although there is only one molecule in an asymmetric unit of TTE0360 crystal, the two molecules related by the two-fold axis of the space-group $P4_12_12$ can form a tight crystallographic dimer. Superimposition of the dimers of TTE0360 and psMMCE shows significant similarities between the two structures, as shown in Figure 1(F). The two TTE0360 monomers pack back to back, stacking with the β strands as well as psMMCE. Their active sites open outward in opposite directions, and thus each monomer active site is completely independent. TTE0360 at the interaction surface, there are many hydrophobic residues just as those in psMMCE, including Val12, Ile14, Ile16, Val18, Phe31, Phe32, Pro70, Leu71, Leu73, Ile96, Ile99, Phe101, Leu113, Leu140, Val142, and Ile144. Dimerization buries 1204 $Å^2$ of the surface, which represents 14% of the surface of each monomer (8800 $Å^2$). The resemblance of dimerization between TTE0360 and MMCE also suggests that TTE0360 may be the MMCE from *T. tengcongensis*.

Active site

MMCE is a metal-dependent enzyme. In the active site of psMMCE, a metal binding site (His12, Gln65, His91, and Glu141) binds the coenzyme Co^{2+}.[12] Superimposition of the structures of TTE0360 and MMCE shows that the four residues Gln13, Glu74, His98, and Glu145 of TTE0360 correspond to the metal binding site. According to the common catalytic mechanism of VOC superfamily members, the two axial active site residues are both glutamic acids, and they are also involved in the catalytic reaction.[13] Although no bound metal ion is present in TTE0360 structure, a typical MMCE metal binding site is formed by these four residues. The distance between the OE1 of Glu74 and the OE1 of Glu145, which could be regarded as the axial residues of the active site, was determined to be 5.1 Å. There is enough space to hold a bound metal ion among the side chains of these residues. We modeled a metal ion into the active site of TTE0360. The distances between the metal ion and the four residues Gln13, Glu74, His98, and Glu145 are 2.8, 2.5, 2.2, and 2.7 Å, respectively, indicating that this metal ion is quite suitable in the active site of TTE0360 [Fig. 2(A)]. In addition, a citric acid molecule was found to be bound in the active site of TTE0360. It is not the substrate of MMCE, but it can bind to the side chains of Gln13, Glu74, Tyr 133, and Glu145 and a water molecular, and makes a number of well-defined hydrogen bonds to form a stable network [Fig. 2(B)]. Mccarthy and coworkers modeled 2-methylmalonate above the binding site of Co^{2+} in psMMCE, and found two oxygens on C1 and C3 of the substrate could form chelate bonds with the metal ion.[12] In TTE0360 structure, two oxygens on C1 and C6 of the citric acid can also form chelate bonds with the modeled metal ion, with the distances 2.2 Å and 2.8 Å, respectively [Fig. 2(A)]. We can predict that the citric acid can compete for the binding site at the substrate binding site with 2-methylmalonate, and it may be a universal substrate of these proteins. Citric acid plays a similar or opposite role when taking the place of 2-methylmalonate remains to be proved.

To summarize, we determined the crystal structure of TTE0360 from *T. tengcongensis* to 2.0 Å resolution. Based on primary sequence analysis in NCBI, the TTE0360 protein was predicted to be glyoxalase I. But the superimposition of the two structures shows that they are very different. We found significant similarities between TTE0360 and a MMCE from *Propionibacterium shermanii* using the program DALI. By superimposition and careful analysis, TTE0360 and psMMCE were found to have very similar monomer structure and almost the same dimerization. Also, the metal binding sites in the cleft of the two structures are conserved. A metal ion was modeled into the metal binding site of TTE0360, and it can form a six-coordination with the four conserved residues and two oxygens of the citric acid. The others, a citric acid molecule binds to the substrate binding site. It may be the competitive analog of 2-methylmalonate. Comparing with the results from primary sequence analysis, the three-dimensional structural informations suggest that TTE0360 may be the MMCE from *Thermoanaerobacter tengcongensis*.

ACKNOWLEDGMENTS

The authors thank professor Run-Sheng Chen of the Institute of Biophysics (CAS, China) for providing the *T. tengcongensis* strains, and professor Yuhui Dong at the Institute of High Energy Physics (CAS, China) for his assistance in data collection.

REFERENCES

1. Kaziro Y, Ochoa S. The metabolism of propionic acid. Adv Enzymol Relat Areas Mol Biol 1964;26:283–378.
2. Allen SH, Kellermeyer RW, Stjernholm RL, Wood HG. Purification and properties of enzymes involved in the propionic acid fermentation. J Bacteriol 1964;87:171–187.
3. Herter S, Farfsing J, Gad'on N, Rieder C, Eisenreich W, Bacher A, Fuchs G. Autotrophic CO(2) fixation by *Chloroflexus aurantiacus*: study of glyoxylate formation and assimilation via the 3-hydroxypropionate cycle. J Bacteriol 2001;183:4305–4316.
4. Korotkova N, Chistoserdova L, Kuksa V, Lidstrom ME. Glyoxylate regeneration pathway in the methylotroph *Methylobacterium extorquens* AM1. J Bacteriol 2002;184:1750–1758.
5. Marsden AF, Caffrey P, Aparicio JF, Loughran MS, Staunton J, Leadlay PF. Stereospecific acyl transfers on the erythromycin-producing polyketide synthase. Science 1994;263:378–380.
6. Swick RW, Wood HG. The role of transcarboxylation in propionic acid fermentation. Proc Natl Acad Sci USA 1960;46:28–41.
7. Brusque AM, Borba Rosa R, Schuck PF, Dalcin KB, Ribeiro CA, Silva CG, Wannmacher CM, Dutra-Filho CS, Wyse AT, Briones P, Wajner M. Inhibition of the mitochondrial respiratory chain complex activities in rat cerebral cortex by methylmalonic acid. Neurochem Int 2002;40:593–601.
8. Kolker S, Schwab M, Horster F, Sauer S, Hinz A, Wolf NI, Mayatepek E, Hoffmann GF, Smeitink JA, Okun JG. Methylmalonic acid, a biochemical hallmark of methylmalonic acidurias but no inhibitor of mitochondrial respiratory chain. J Biol Chem 2003;278:47388–47393.
9. Dauter Z, Dauter M. Anomalous signal of solvent bromides used for phasing of lysozyme. J Mol Biol 1999;289:93–101.
10. Dauter Z, Dauter M, Rajashankar KR. Novel approach to phasing proteins: derivatization by short cryo-soaking with halides. Acta Crystallogr 2000;56:232–237.
11. Holm L, Sander C. Touring protein fold space with Dali/FSSP. Nucl Acids Res 1998;26:316–319.
12. Mccarthy AA, Baker HM, Shewry SC, Patchett ML, Baker EN. Crystal structure of methylmalonyl-coenzyme A epimerase from P. shermanii: a novel enzymatic function on an ancient metal binding scaffold. Structure 2001;9:637–646.
13. Armstrong RN. Mechanistic diversity in a metalloenzyme superfamily. Biochemistry 2000;39:13625–13632.
14. Bao Q, Tian Y, Li W, Xu Z, Xuan Z, Hu S, Dong W, Yang J, Chen Y, Xue Y, Xu Y, Lai X, Huang L, Dong X, Ma Y, Ling L, Tan H, Chen R, Wang J, Yu J, Yang H. A complete sequence of the *T. tengcongensis* genome. Genome Res 2002;12:689–700.
15. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol 1997;276:307–326.
16. Matthews BW. Solvent content of protein crystals. J Mol Biol 1968;33:491–497.
17. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, Mccoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr 2002;58:1948–1954.
18. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nat Struct Biol 1999;6:458–463.
19. Brunger AT, Adams PD, Clore GM, Delano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr 1998;54:905–921.
20. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr 2004;60:2126–2132.
21. Laskowski RA, Macarthur MW, Moss DS, Thornton JM. Procheck–a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
22. Cameron AD, Olin B, Ridderstrom M, Mannervik B, Jones TA. Crystal structure of human glyoxalase I–evidence for gene duplication and 3D domain swapping. EMBO J 1997;16:3386–3395.
23. Kawatani M, Okumura H, Honda K, Kanoh N, Muroi M, Dohmae N, Takami M, Kitagawa M, Futamura Y, Imoto M, Osada H. The identification of an osteoclastogenesis inhibitor through the inhibition of glyoxalase I. Proc Natl Acad Sci USA 2008;105:11691–11696.
24. He MM, Clugston SL, Honek JF, Matthews BW. Determination of the structure of *Escherichia coli* glyoxalase I suggests a structural basis for differential metal activation. Biochemistry 2000;39:8719–8727.
25. Ariza A, Vickers TJ, Greig N, Armour KA, Dixon MJ, Eggleston IM, Fairlamb AH, Bond CS. Specificity of the trypanothione-dependent Leishmania major glyoxalase I: structure and biochemical comparison with the human enzyme. Mol Microbiol 2006;59:1239–1248.