Accelerated Publications

Evolution of Enzymatic Activities in the Enolase Superfamily: Stereochemically Distinct Mechanisms in Two Families of *cis*,*cis*-Muconate Lactonizing Enzymes[†,‡]

Ayano Sakai,[§] Alexander A. Fedorov,[||,⊥] Elena V. Fedorov,[||] Alexandra M. Schnoes,[#] Margaret E. Glasner,[#] Shoshana Brown,[#] Marc E. Rutter,[+] Kevin Bain,[+] Shawn Chang,[+] Tarun Gheyi,[+] J. Michael Sauder,[+] Stephen K. Burley,[+] Patricia C. Babbitt,[*,#] Steven C. Almo,[*,||,⊥] and John A. Gerlt[*,@]

Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois 61801, Department of Biochemistry, Albert Einstein College of Medicine, Bronx, New York 10461, New York SGX Research Center for Structural Genomics, Albert Einstein College of Medicine, Bronx, New York 10461, Department of Biopharmaceutical Sciences, School of Pharmacy, and California Institute for Quantitative Biomedical Research, University of California, 1700 Fourth Street, San Francisco, California 94158, New York SGX Research Center for Structural Genomics, SGX Pharmaceuticals, Inc., 10505 Roselle Street, San Diego, California 92121, and Departments of Biochemistry and Chemistry, University of Illinois, Urbana, Illinois 61801

Received December 14, 2008; Revised Manuscript Received January 15, 2009

ABSTRACT: The mechanistically diverse enolase superfamily is a paradigm for elucidating Nature's strategies for divergent evolution of enzyme function. Each of the different reactions catalyzed by members of the superfamily is initiated by abstraction of the α-proton of a carboxylate substrate that is coordinated to an essential Mg^{2+}. The muconate lactonizing enzyme (MLE) from *Pseudomonas putida*, a member of a family that catalyzes the *syn*-cycloisomerization of *cis*,*cis*-muconate to (4*S*)-muconolactone in the β-ketoadipate pathway, has provided critical insights into the structural bases for evolution of function within the superfamily. A second, divergent family of homologous MLEs that catalyzes *anti*-cycloisomerization has been identified. Structures of members of both families liganded with the common (4*S*)-muconolactone product (*syn*, *Pseudomonas fluorescens*, gi 70731221; *anti*, *Mycobacterium smegmatis*, gi 118470554) document that the conserved Lys at the end of the second β-strand in the $(\beta/\alpha)_7\beta$-barrel domain serves as the acid catalyst in both reactions. The different stereochemical courses (*syn* and *anti*) result from different structural strategies for determining substrate specificity: although the distal carboxylate group of the *cis*,*cis*-muconate substrate attacks the same face of the proximal double bond, opposite faces of the resulting enolate anion intermediate are presented to the conserved Lys acid catalyst. The discovery of two families of homologous, but stereochemically distinct, MLEs likely provides an example of "pseudoconvergent" evolution of the same function from different homologous progenitors within the enolase superfamily, in which different spatial arrangements of active site functional groups and substrate specificity determinants support catalysis of the same reaction.

The mechanistically diverse enolase superfamily was recognized by the discovery that the structures of *cis*,*cis*-muconate lactonizing enzyme (MLE)[1] and mandelate racemase (MR) from *Pseudomonas putida* are remarkably similar, despite sharing only 26% sequence identity (*1*). Although MLE and MR catalyze different overall reactions, each is initiated by general base-catalyzed abstraction of the α-proton of the carboxylate substrate that is coordinated to an essential Mg^{2+}. Without significant stabilization by the Mg^{2+}, the resulting enediolate intermediates would be too unstable to be kinetically competent.

The polypeptide chains of members of the enolase superfamily are organized in two domains (*2, 3*): (1) a $(\beta/\alpha)_7\beta$-

[†] This research was supported by NIH Grant R01 GM-60595 (to P.C.B.), NIH Grant P01 GM-71790 (to J.A.G., P.C.B., and S.C.A.), NIH Grant U54 GM-74945 (to S.K.B. and S.C.A.), and NSF Grant 0640476 (to P.C.B.). Molecular graphics images were produced using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH Grant P41 RR-01081).

[‡] The X-ray coordinates and structure factors for the MLE from *Mycobacterium smegmatis* liganded with Mg^{2+}, the MLE from *M. smegmatis* liganded with Mg^{2+} and (4*S*)-muconolactone in two space groups, the MLE from *Pseudomonas fluorescens* liganded with Mg^{2+}, and the MLE from *P. fluorescens* liganded with Mg^{2+} and (4*S*)-muconolactone have been deposited in the Protein Data Bank (entries 3DG3, 3DG6, 3DG7, 3CT2, and 3DGB, respectively).

* To whom correspondence should be addressed. J.A.G.: Department of Biochemistry, University of Illinois, 600 S. Mathews Ave., Urbana, IL 61801; phone, (217) 244-7414; fax, (217) 244-6538; e-mail, j-gerlt@uiuc.edu. S.C.A.: Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, NY 10461; phone, (718) 430-2746; fax, (718) 430-8565; e-mail, almo@aecom.yu.edu. P.C.B.: Department of Biopharmaceutical Sciences, University of California, San Francisco, 1700 S. 4th St., San Francisco, CA 94158; phone, (415) 476-3784; fax, (415) 514-4797; e-mail, babbitt@cgl.ucsf.edu.

[§] Center for Biophysics and Computational Biology, University of Illinois.

[||] Department of Biochemistry, Albert Einstein College of Medicine.

[⊥] New York SGX Research Center for Structural Genomics, Albert Einstein College of Medicine.

[#] University of California.

[+] SGX Pharmaceuticals, Inc.

[@] Departments of Biochemistry and Chemistry, University of Illinois.

[1] Abbreviations: AEE, L-Ala-D/L-Glu epimerase; CCM, *cis*,*cis*-muconate; LB, Luria-Bertani broth; MLE, muconate lactonizing enzyme; MR, mandelate racemase; NSAR, *N*-succinylamino acid racemase; OSBS, *o*-succinylbenzoate synthase; rmsd, root-mean-square deviation; TIM, triosephosphate isomerase.

10.1021/bi802277h CCC: $40.75
Published on Web 01/16/2009

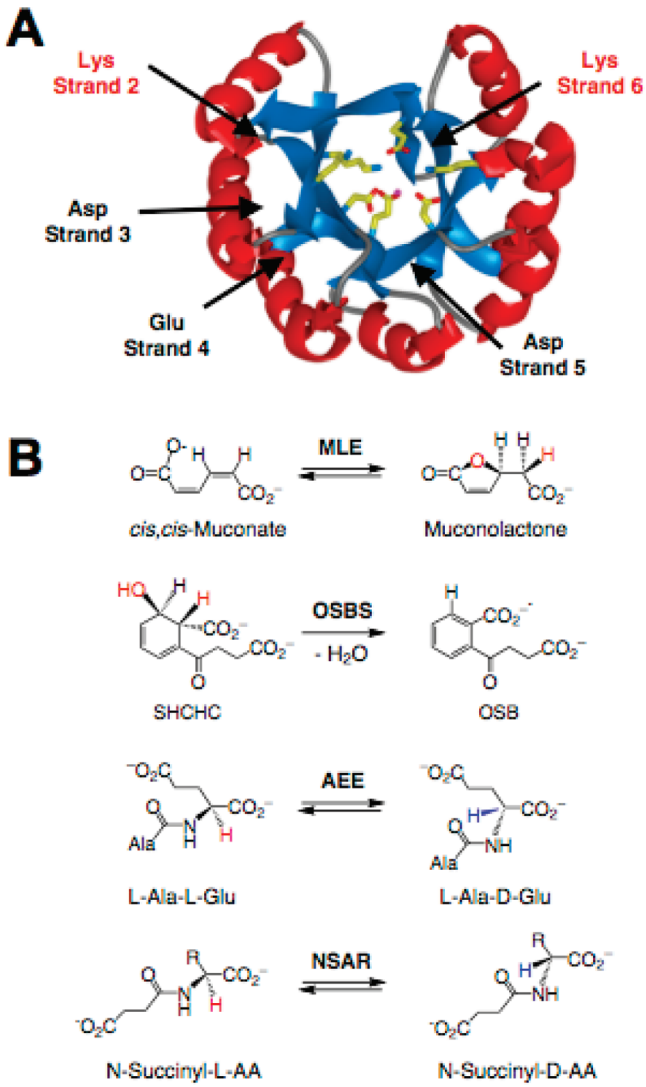

FIGURE 1: Structure and functions in the MLE subgroup. (A) $(\beta/\alpha)_7\beta$-barrel domain of the MLE subgroup showing the positions of the conserved Asp/Glu ligands for the essential Mg^{2+} at the ends of the third, fourth, and fifth β-strands and the Lys acid/base catalysts at the ends of the second and sixth β-strands. (B) Reactions catalyzed by members of the MLE subgroup.

barrel domain (modified TIM barrel) that contains the ligands for the Mg^{2+} as well as general acid/base catalysts at the C-terminal ends of the β-strands and (2) an $\alpha+\beta$ capping domain formed from N-terminal and, sometimes, C-terminal segments of the polypeptide chain that closes over the mouth of the barrel domain and contains residues responsible for determining substrate specificity. This large superfamily can be partitioned into subgroups on the basis of phylogenetic analyses, the identities of the ligands for the essential Mg^{2+}, and the identities and positions of the general acid/base catalysts. Functionally diverse subgroups include (1) the MLE subgroup in which the acid/base catalysts are Lys residues at the ends of the second and sixth β-strands (Figure 1A) and (2) the MR subgroup in which one of the acid/base catalysts is a His-Asp dyad at the ends of the seventh and sixth β-strands. Within the superfamily, the only conserved residues are ligands for the Mg^{2+} at the ends of the third, fourth, and fifth β-strands, thereby establishing the importance of stabilization of the enediolate intermediate.

Four reactions have been identified for members of the MLE subgroup (Figure 1B): (1) *syn*-cycloisomerization catalyzed by MLE (*4*−*6*), (2) *syn*-dehydration catalyzed by *o*-succinylbenzoate synthase (OSBS) (*7*), (3) dipeptide epimerization catalyzed by L-Ala-D/L-Glu epimerase (AEE) (*8*) and other recently identified epimerases with divergent substrate specificities (*9*), and (4) racemization catalyzed by *N*-succinylamino acid racemases (NSARs) (*10, 11*). Flexible 20s and 50s loops in the capping domain complete the active site cavity in which the substrate is coordinated to the Mg^{2+} and positioned proximal to the acid/base catalysts. Divergent evolution within the MLE subgroup follows the strategy which holds that different reactions can be accomplished by changing the identity of the carboxylate anion that binds vis-à-vis the conserved acid/base catalysts by altering the identities and locations of the residues that form the active site cavity (*12*−*14*).

A phylogenetic tree that depicts the relationships between sequence and function in the MLE subgroup is shown in Figure 2. The OSBSs catalyze *syn*-dehydration, as dictated by the structure of the substrate (*7*). The sequences are divergent; several subfamilies have been identified in which the common substrate specificity is determined by different structural motifs (*15*). The Lys at the end of the second β-strand is positioned proximal to the abstracted α-proton and the OH leaving group.

Several subfamilies of dipeptide epimerases have been identified. The AEE from *Bacillus subtilis* is specific for L-Ala-D/L-Glu, with Arg 24 in the 20s loop determining the specificity for the Glu residue (*8*). Homologous dipeptide epimerases, with sequence-divergent 20s and 50s loops in the capping domain, catalyze epimerization of other dipeptides (*9*). The epimerases share a DxD motif at the end of the eighth β-strand, with the Asp residues hydrogen-bonded to the substrate's ammonium group. The Lys at the end of the second β-strand is the (*R*)-specific acid/base; the Lys at the end of the sixth β-strand is the (*S*)-specific acid/base.

Two families of NSARs have been identified. The first (NSAR-1 in Figure 2) includes the NSAR from *Geobacillus kaustophilus* that catalyzes racemization of hydrophobic *N*-succinylamino acids (*10*) and the promiscuous NSAR/OSBS from *Amycolatopsis* (*16, 17*); the second (NSAR-2 in Figure 2) includes the NSAR from *Bacillus cereus* that catalyzes racemization of *N*-succinyl Arg and *N*-succinyl Lys (*11*). The specificity determinants differ in the two families (*11, 17*), but, in each, the *N*-succinyl moiety is proximal to the end of the eighth β-strand, the Lys at the end of the second β-strand is the (*R*)-specific acid/base catalyst, and the Lys at the end of the sixth β-strand is the (*S*)-specific acid/base catalyst. Despite these shared features, the families evolved from different, but homologous, progenitors, providing an example of "pseudoconvergent" evolution of function within the enolase superfamily; i.e., these proteins evolved from different intermediate ancestors derived from the common progenitor of the enolase superfamily (*11*). This contrasts with "true" convergent evolution of the same function from unrelated progenitors (analogues) (*15, 18*).

Although the MLE from *P. putida* is the structural paradigm for the MLE subgroup, its structure−function relationships are uncertain (*6*). Addition of the distal carboxylate group across the proximal double bond proceeds with a *syn*-stereochemistry to produce (4*S*)-muconolactone, with the 5-pro-*R* hydrogen derived from solvent (*4, 5*). A potential Lys acid catalyst is located on either face of the

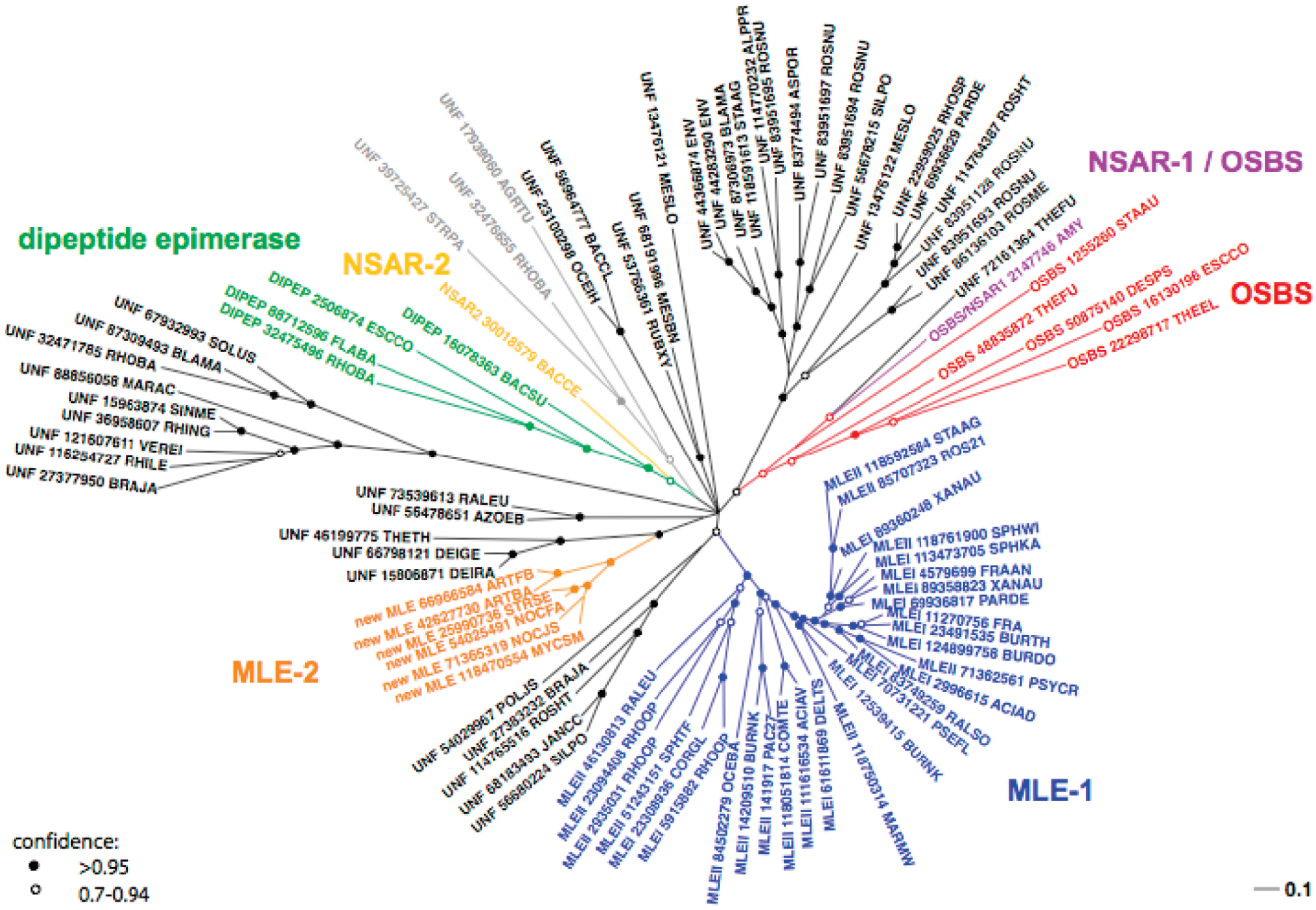

FIGURE 2: Phylogenetic tree of the MLE subgroup constructed as described in Materials and Methods.

active site at the end of the second or sixth β-strand, but its identity is uncertain because no product-liganded structure has been available.

Herein, we report the structure of the (4*S*)-muconolactone-liganded complex of the MLE from *Pseudomonas fluorescens* that catalyzes *syn*-cycloisomerization (in the same branch of the tree and 76% sequence identity with the MLE from *P. putida*; MLE-1 in Figure 2). We also report functional and structural characterization of the MLE from *Mycobacterium smegmatis*, a member of a second family of MLEs[2] (MLE-2 in Figure 2), that catalyzes *anti*-cycloisomerization, although it also produces (4*S*)-muconolactone. In both families, the Lys at the end of the second β-strand serves as the acid catalyst; the different stereochemical courses result from distinct substrate binding modes. In the *syn*-family, a solvent-derived proton is delivered to the 5-pro-*R* position of the product; in the *anti*-family, a solvent-derived proton is delivered to the 5-pro-*S* position. The possible pseudoconvergent evolution to yield MLEs with stereochemically distinct mechanisms reinforces our conclusion that evolution of function within the enolase superfamily often occurs via changes in specificity-determining residues and not changes in the identities or positions of the acid/base catalysts.

MATERIALS AND METHODS

Genomic Context. All MLE subgroup sequences except those belonging to the dipeptide epimerase, OSBS, and NSAR families were used as queries in Microbes Online (http://www.microbesonline.org/) to identify adjacent genes that are likely to be regulated as operons. The relationships among the MLE subgroup proteins and all other proteins identified in Microbes Online were assessed using Cytoscape (*19*) to determine whether the MLE subgroup proteins share the same operon contexts, which would suggest that they share the same functions.

Phylogenetic Inference. All members of the MLE subgroup were aligned using Muscle version 3.52 (*20*), and the alignment was manually adjusted on the basis of a structural alignment (*15*). Phylogenetic reconstruction was performed using Bayesian and distance methods. Bayesian trees were constructed with MrBayes version 3.1.1 (*21, 22*) under the WAG amino acid substitution model (*23*) using a gamma distribution to approximate rate variation among sites.

Cloning, Expression, and Purification of MLE-2 from M. smegmatis (gi 118470554). The gene encoding MLE-2 from *M. smegmatis* was PCR amplified from *M. smegmatis* strain MC2 155 genomic DNA (kindly provided by W. R. Jacobs, Jr., Howard Hughes Medical Institute and Albert Einstein College of Medicine) using platinum *Pfx* DNA polymerase (Invitrogen). The PCR mixture (50 μL) contained 60 ng of

[2] gi 42627730, *Arthrobacter* sp. BA-5-17; gi 116670480, *Arthrobacter* sp. FB24; gi 145221949, *Mycobacterium gilvum* PYR-GCK; gi 118470554, *Mycobacterium smegmatis* strain MC2 155; gi 126434001, *Mycobacterium* sp. JLS; gi 108798346, *Mycobacterium* sp. KMS; gi 108789346, *Mycobacterium* sp. MCS; gi 54025491, *Nocardia farcinica* IFM 10152; gi 119715080, *Nocardioides* sp. JS614; gi 134100914, *Saccharopolyspora erythraea* NRRL 2338; gi 25990736, *Streptomyces setonii*.

Table 1: X-ray Data Collection and Refinement Statistics for Crystals of MLEs from *M. smegmatis* and *P. fluorescens*

	MLE-MS·Mg^{2+}	MLE-MS·Mg^{2+}·CCM	MLE-MS·Mg^{2+}·CCM	MLE-PF·Mg^{2+}	MLE-PF·Mg^{2+}·CCM
			Data Collection		
space group	*I*422	*I*422	*C*2	$P42_12$	*I*422
no. of molecules in a.u.	1	1	4	2	1
unit cell					
a (Å)	123.62	123.84	171.05	135.93	133.84
b (Å)			124.26		
c (Å)	117.59	117.37	117.48	82.97	103.37
β (deg)			133.4		
resolution (Å)	25.0−1.6	25.0−1.6	25.0−2.0	25.0−1.8	25.0−1.7
no. of unique reflections	56648	59180	116011	71185	51329
completeness (%)	94.4	98.5	96.5	98.4	99.7
R_{merge}	0.087	0.098	0.076	0.081	0.074
average I/σ	25.9	39.1	15.2	14.7	22.7
			Refinement		
resolution (Å)	25.0−1.6	25.0−1.6	25.0−2.0	25.0−1.8	25.0−1.7
R_{cryst}	0.189	0.185	0.188	0.215	0.189
R_{free}	0.204	0.201	0.203	0.222	0.209
rmsd for bonds (Å)	0.005	0.005	0.006	0.005	0.005
rmsd for angles (deg)	1.2	1.3	1.3	1.2	1.2
no. of atoms					
protein	2710	2797	11188	5408	2787
water	298	299	789	414	236
Mg^{2+}	1	1	4	2	1
inhibitor	0	40	10		
PDB entry	3DG3	3DG6	DG7	3CT2	3DGB

DNA, 5 μL of 10× *Pfx* amplification buffer, 1 mM $MgSO_4$, 0.4 mM dNTP, 20 pmol of each primer, and 1 unit of platinum *Pfx* polymerase. The gene was amplified and cloned into the pET-17b vector (Novagen). The MLE was expressed in *Escherichia coli* strain BL21(DE3). Transformed cells were grown at 37 °C for 48 h in LB supplemented with 100 μg/mL ampicillin and harvested by centrifugation. The cells were resuspended in low-salt buffer [20 mM Tris (pH 8.0) and 5 mM $MgCl_2$] and lysed by sonication. The lysate was clarified by centrifugation, and the supernatant was applied to a DEAE-Sepharose Fast Flow (GE Healthcare) column in low-salt buffer. The protein was eluted with a linear gradient of NaCl (from 0.5 to 1 M) in low-salt buffer. Fractions containing MLE were identified by SDS−PAGE analysis and applied to a Phenyl Sepharose 6 Fast Flow column in 0.5 M $(NH_4)_2SO_4$ for further purification. The protein was eluted with a linear gradient of $(NH_4)_2SO_4$ (from 0.5 to 0 M), and the purest fractions were pooled and dialyzed against 20 mM Tris (pH 8.0) and 5 mM $MgCl_2$.

Cloning, Expression, and Purification of MLE-1 from P. fluorescens (gi 70731221). The PCR product from *P. fluorescens* genomic DNA (ATCC) was cloned into a TOPO vector (Invitrogen) with a noncleavable C-terminal His_6 tag. The clone was transformed into BL21-RIL cells (CodonPlus). Overnight cultures were grown from glycerol stocks in HY medium (Medicilon Inc.) and transferred to 2 L shake flasks containing 1 L of HY medium with kanamycin and chloramphenicol. Cells were grown to an OD_{595} of 1.0 at 37 °C; IPTG was added to induce expression of the MLE. After 18 h at 22 °C, the cells were harvested by centrifugation; the pellets were stored at −80 °C. Cells were lysed by sonication in 20 mM Tris (pH 8.0), 0.5 M NaCl, 25 mM imidazole, and 0.1% Tween 20. The protein was purified by sequential nickel-affinity and size-exclusion chromatographies.

Spectrophotometric Assay for MLE Activity. MLE activity was measured at 25 °C by quantitating the decrease in absorbance at 260 nm due to the consumption of *cis*,*cis*-muconate ($\varepsilon_{260} = 16900\ M^{-1}\ cm^{-1}$) in 50 mM Tris-HCl (pH 7.5) and 0.1 mM $MnCl_2$.

Polarimetric Assay for Detection of (4S)-Muconolactone Configurations. The change in optical rotation due to muconolactone production was followed at Hg 405 nm using a JASCO P-1010 polarimeter at 25 °C until equilibrium was established. The reaction mixture contained 10 mM *cis*,*cis*-muconate and 0.1 μM MLE in 50 mM Tris-HCl (pH 7.5) and 0.1 mM $MnCl_2$.

Stereochemical Studies. ^{1}H NMR spectra (500 MHz) were recorded in a D_2O-containing buffer at 37 °C. The reaction mixtures contained 10 mM *cis*,*cis*-muconate, 50 mM sodium phosphate buffer (pD 6.0), 5 mM $MgCl_2$, and 1 μM MLE.

Crystallization and Data Collection. Five crystal forms (Table 1) were grown by the hanging drop vapor diffusion method at room temperature: (1) MLE from *M. smegmatis* (MLE-MS) and Mg^{2+}, (2) MLE-MS, Mg^{2+}, and *cis*,*cis*-muconate (CCM) in a monoclinic space group, (3) MLE-MS, Mg^{2+}, and CCM in a triclinic space group, (4) MLE from *P. fluorescens* (MLE-PF) and Mg^{2+}, and (5) MLE-PF, Mg^{2+}, and CCM. Crystallization utilized the following conditions.

(1) For MLE-MS and Mg^{2+} (MLE-MS·Mg^{2+}), the protein (33.7 mg/mL) was in 20 mM Tris-HCl (pH 8.0) and 5 mM $MgCl_2$; the precipitant consisted of 10% 2-propanol, 0.1 M MES (pH 6.0), and 0.2 M $Ca(OAc)_2$.

(2) For crystallization of MLE-MS, Mg^{2+}, and CCM (MLE-MS·Mg^{2+}·CCM) in tetragonal space group *I*422, the protein (33.7 mg/mL) was in 20 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, and 20 mM CCM; the precipitant consisted of 1.0 M $(NH_4)_2SO_4$, 0.1 M HEPES (pH 7.0), and 0.5% PEG 8000.

(3) For crystallization of MLE-MS·Mg^{2+}·CCM in monoclinic space group *C*2, the conditions were identical to those used for the previous sample, except that the CCM concentration was 40 mM.

(4) For MLE-PF and Mg^{2+} (MLE-PF·Mg^{2+}), the protein (5.0 mg/mL) was in 20 mM HEPES (pH 7.5), 150 mM NaCl, 10 mM $MgCl_2$, and 10% glycerol; the precipitant consisted of 20% PEG 1000, 0.1 M cacodylate (pH 6.5), and 0.2 M $MgCl_2$.

(5) For MLE-PF, Mg^{2+}, and CCM (MLE-PF·Mg^{2+}·CCM), the protein (4.6 mg/mL) was in 20 mM HEPES (pH 7.5), 150 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, and 80 mM CCM; the precipitant consisted of 2.0 M $(NH_4)_2SO_4$, 0.1 M MES (pH 6.0), and 5% 2-propanol.

The crystals were transferred to a cryoprotectant solution composed of their mother liquid and 20% glycerol and then flash-cooled in a nitrogen stream. X-ray diffraction data were collected at NSLS beamline X4A (Brookhaven National Laboratory) on an ADSC CCD detector. Diffraction intensities were integrated and scaled with DENZO and SCALEPACK (*24*). The data collection statistics are given in Table 1.

Structure Determination and Model Refinement. The structure of MLE-MS·Mg^{2+} was determined by molecular replacement with the fully automated molecular replacement pipeline BALBES (*25*), using only diffraction data and the amino acid sequence. Iterative cycles of manual rebuilding with TOM (*26*), refinement with CNS (*27*), and automatic rebuilding with ARP (*28*) were performed. Residues 17−27 in the 20s loop have no density and were excluded from the final refinement model.

For the structure of the tetragonal MLE-MS·Mg^{2+}·CCM complex, the MLE-MS·Mg^{2+} structure was used as the starting point. Iterative cycles of manual rebuilding with TOM, refinement with CNS, and automatic rebuilding with ARP with subsequent inclusion of water molecules were performed. The structure of the monoclinic MLE-MS·Mg^{2+}·CCM complex was determined and partially refined with BALBES. Subsequent iterative cycles of manual rebuilding with TOM, refinement with CNS, and automatic rebuilding with ARP were performed. Both structures of MLE-MS·Mg^{2+}·CCM contain residues 1−366. The 20s loop, Mg^{2+} ion, and CCM have well-defined density in the single polypeptide chain in the tetragonal crystal form and in all four copies of the polypeptide chain comprising the asymmetric unit for the monoclinic crystal form.

The structures of MLE-PF·Mg^{2+} and MLE-PF·Mg^{2+}·CCM were determined and refined by analogous methods. The final atomic model of MLE-PF·Mg^{2+} includes residues 4−375 with the exception of residues 22−32 in the 20s loop. The final atomic model of MLE-PF·Mg^{2+}·CCM includes residues 4−374; the 20s loop, Mg^{2+} ion, and CCM are well-defined in the structure.

Crystallographic refinement statistics are also provided in Table 1.

Molecular Graphics Images. The images in Figure 1A and Figure 5A,C,D were constructed using Chimera (*29*).

RESULTS AND DISCUSSION

As described in the introductory section, several subfamilies of monophyletic OSBSs and dipeptide epimerases as well as two phylogenetically distinct families of NSARs can be identified in the MLE subgroup (Figure 2) (*9, 15*). Also shown in Figure 2 are two families of MLEs: one (MLE-1) that includes the paradigmatic MLE from *P. putida* and a second (MLE-2) that includes a protein from *Arthrobacter* sp. BA-5-17 (gi 42627739) that was recently shown to be an MLE by enzymatic assays (*30*). We extended the MLE functional assignment to other members of the "new" family[2] using operon context, i.e., proximal to genes encoding catechol 1,2-dioxygenase, muconolactone δ-isomerase, and/or enol-lactone hydrolase.

Table 2: Kinetic Constants for MLEs

MLE	k_{cat} (s^{-1})	K_m (M)	k_{cat}/K_m (M^{-1} s^{-1})
M. smegmatis	78 ± 6	$(5.6 \pm 1) \times 10^{-4}$	1.4×10^5
P. putida	11 ± 0.5	$(5.4 \pm 0.5) \times 10^{-4}$	2.0×10^4
P. fluorescens	110 ± 2	$(1.6 \pm 0.2) \times 10^{-5}$	7.0×10^6

Kinetic and Stereochemical Characterization of a New MLE (MLE-2). We isolated a new MLE-2 from *M. smegmatis* (gi 118470554, 31% identical in sequence with MLE-1 from *P. putida*, gi 157838298, and occupying a different branch in the phylogenetic tree in Figure 2); the kinetic constants of the MLEs from *M. smegmatis* and *P. putida* for cycloisomerization of *cis*,*cis*-muconate are compared in Table 2.

We determined the stereochemical course of the reaction catalyzed by the MLE-2 from *M. smegmatis* by (1) measuring the optical rotation of the muconolactone product and (2) performing the reaction in D_2O so that the location of the solvent-derived hydrogen in the muconolactone products could be determined by ^1H NMR spectroscopy (at neutral pH, the equilibrium favors the muconolactone). Like MLE-1 from *P. putida* (*4, 5*), MLE-2 from *M. smegmatis* catalyzes formation of (+)-(4*S*)-muconolactone (data not shown); however, unlike the paradigmatic MLE, the solvent-derived deuterium is incorporated into the 5-pro-*R* hydrogen, not the 5-pro-*S* hydrogen (Figure 3B). These observations establish that cycloisomerization catalyzed by MLE-2 proceeds via an *anti*-stereochemical course, in contrast to the *syn*-stereochemical course for MLE-1 from *P. putida*.

Structures of Product-Liganded MLEs. We obtained crystals and determined structures for the *anti*-MLE from *M. smegmatis* both in the absence and in the presence of *cis*,*cis*-muconate.

Because crystals of a substrate/product-liganded complex of the MLE from *P. putida* could not be determined, we isolated other MLEs from the *syn*-family (MLE-1)[3] and subjected these to crystallization trials in the absence and presence of *cis*,*cis*-muconate. Crystals were obtained and structures determined for MLE-1 from *P. fluorescens* both in the absence and in the presence of *cis*,*cis*-muconate (gi 70731221, 76% identical in sequence with MLE-1 from *P. putida* and in the same branch of the tree and 30% identical in sequence with MLE-2 from *M. smegmatis*). We determined the kinetic constants (Table 2) and stereochemical course (Figure 3C) for the reaction catalyzed by the MLE from *P. fluorescens*. This enzyme catalyzes *syn*-cycloisomerization (with incorporation of solvent deuterium in the 5-pro-*S* hydrogen) like the paradigmatic MLE from *P. putida*.

As expected on the basis of sequence similarity (Figure 4), the MLEs from *M. smegmatis* and *P. fluorescens* are structurally similar (Figure 5A; rmsd for 273 atom pairs of 1.24 Å), with the barrel domains containing the expected

[3] gi 23308936, *Corynebacterium glutamicum*; gi 152970427, *Klebsiella pneumoniae* MGH 78758; gi 70731221, *P. fluorescens* PF-5.

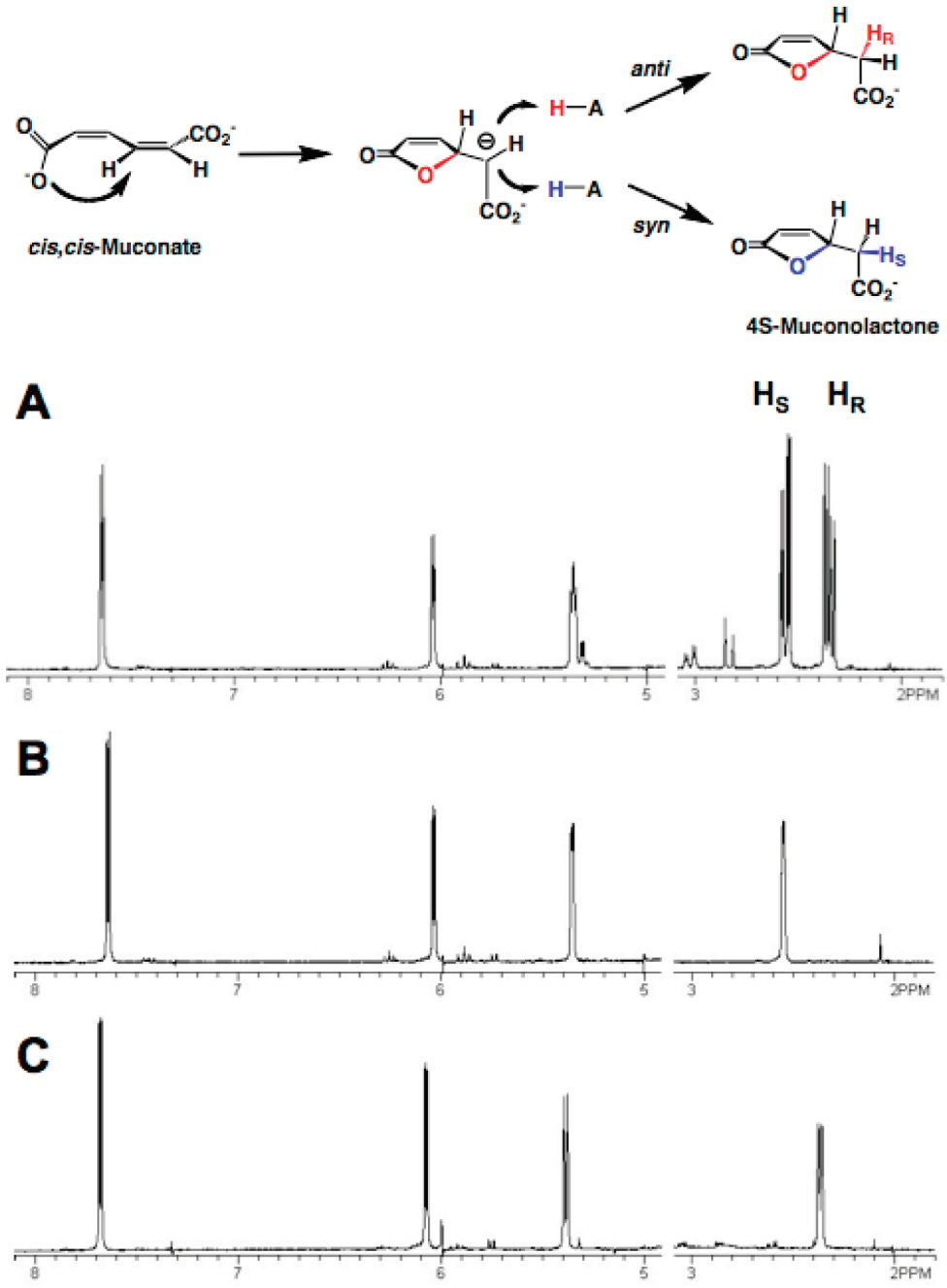

FIGURE 3: Stereochemical course of the MLE-catalyzed reaction. The facial relationships of the carboxylate group and active site acid/base catalyst in *anti*- and *syn*-cycloisomerization reactions are shown at the top. (A) ^{1}H NMR spectrum of (4*S*)-muconolactone obtained in H_2O. (B) ^{1}H NMR spectrum of (4*S*)-muconolactone obtained in D_2O with the MLE from *M. smegmatis*. (C) ^{1}H NMR spectrum of (4*S*)-muconolactone obtained in D_2O with the MLE from *P. fluorescens*.

carboxylate ligands for the Mg^{2+} at the ends of the third (Asp), fourth (Glu), and fifth (Asp) β-strands and Lys residues at the ends of the second and sixth β-strands (Figure 5C,D). In both structures, the (4*S*)-muconolactone product is sequestered from solvent by the capping domain [electron densities for the (4*S*)-muconolactone products are shown in Figure 5B]. The sequence-divergent 20s loops contain the same number of residues, whereas the sequence-divergent 50s loops differ in length by two residues (Figure 4). Thus, the structures of the active site cavities are not conserved between the two families.

In both MLEs, the Lys at the end of the second β-strand is proximal to C5 of the (4*S*)-muconolactone product (Figure 5C,D). Lys 171 in the *syn*-MLE is spatially proximal to both the 5-pro-*S* hydrogen and the lactone oxygen. Lys 162 in the *anti*-MLE is spatially proximal to the 5-pro-*R* hydrogen and opposite the lactone oxygen derived from the distal carboxylate group. These relationships are those expected from the stereochemical courses and establish that the Lys at the end of the second β-strand serves as the acid/base catalyst in both families. In both structures, the Lys at the end of the sixth β-strand (Lys 275 in the *syn*-MLE and Lys 266 in the *anti*-MLE) is proximal to the carboxylate group of the muconolactone product that is coordinated to the Mg^{2+} and presumably assists in stabilization of the enolate anion intermediate.

The lactone rings (derived from the distal carboxylate group of the substrate) are located in different positions in

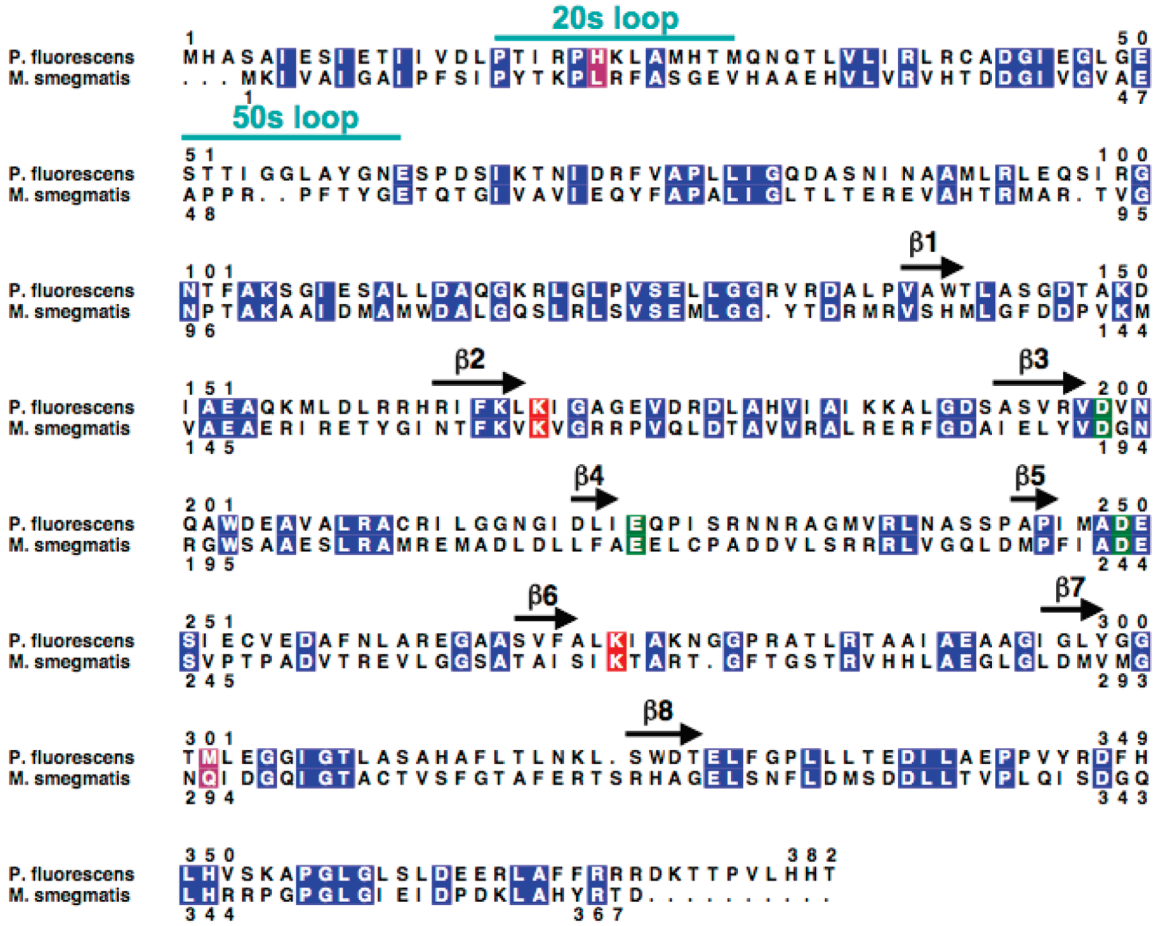

FIGURE 4: Sequence alignment of the *syn*- and *anti*-MLEs from *P. fluorescens* and *M. smegmatis*, respectively. The positions of the 20s and 50s loops in the capping domain and the β-strands in the $(\beta/\alpha)_7\beta$-barrel domain are indicated. The conserved residues are highlighted in blue, the ligands for Mg^{2+} in green, the Lys residues at the end of the second and sixth β-strands in red, and the positions of the hydrogen bond donors to the lactone carbonyl of (4*S*)-muconolactone in magenta (the structurally homologous residues for these hydrogen bond donors also are highlighted to emphasize the differing identities and positions of the specificity determinants).

the active sites of the two MLEs. In the *syn*-MLE, the lactone carbonyl oxygen is hydrogen-bonded to His 24 in the 20s loop that is conserved in the MLE-1 family; in the *anti*-MLE, the lactone carbonyl oxygen is hydrogen-bonded to Gln 294 in the loop following the seventh β-strand that is conserved in the MLE-2 family. These interactions enforce different rotamers between the Mg^{2+}-coordinated carboxylate group and proximal double bond *cis*,*cis*-muconate substrate, although in both enzymes the distal carboxylate group attacks the *re*-face of the proximal double bond to produce the same (4*S*)-muconolactone enolate anion intermediate. The conserved Lys at the end of the second β-strand is located on opposite faces of the proximal double bond and delivers a proton to opposite faces of the conserved enolate anion intermediate, producing the observed opposite stereochemical courses for the reactions.

Divergent Evolution of the MLE Function. Despite the differing stereochemical courses, the mechanisms of the *syn*- and *anti*-cycloisomerization reactions are expected to be the same: (1) addition of the distal carboxylate to the same face of the proximal double bond to form the enolate intermediate that is stabilized by coordination to the essential Mg^{2+} as well as the proximity to the Lys at the end of the sixth β-strand and (2) protonation of the intermediate, albeit on opposite faces, by the conserved Lys acid catalyst at the end of the second β-strand as the result of the differing binding geometries for the *cis*,*cis*-muconate substrate.

Unlike the 1,1-proton transfer reactions catalyzed by AEEs and NSARs and the substrate structure-enforced *syn*-dehydration reaction catalyzed by the OSBSs, the cycloisomerization reactions catalyzed by the divergent MLEs occur via distinct stereochemical pathways that involve opposite orientations of the carboxylate nucleophile vis-à-vis the acid catalyst. This difference in catalytic strategy is of no metabolic consequence because the same (4*S*)-muconolactone product is produced by both families.

Pseudoconvergent Evolution of the MLE Function in the Enolase Superfamily? Presumably, the progenitors for the *syn*- and *anti*-MLEs differed in the shapes and polarities of their active sites so that the evolved MLEs would enforce *cis*,*cis*-muconate to bind as the distinct rotamers that determine the stereochemical courses of the *syn*- and *anti*-cycloisomerization reactions. Phylogenetic and genomic context relationships provide support for the hypothesis that the progenitors for the *syn*- and *anti*-MLEs may have had different catalytic functions. With regard to the phylogenetic tree in Figure 2, although the MLEs share the same active site motif (metal ion ligands and Lys acidic/electrophilic catalysts), they are separated by a group of uncharacterized proteins that likely do not catalyze the MLE reaction as suggested by operon/genome context. The genomes that encode this intervening proteins lack homologues of catechol 1,2-dioxygenase and muconolactone δ-isomerase, suggesting that these organisms do not have a β-ketoadipate pathway

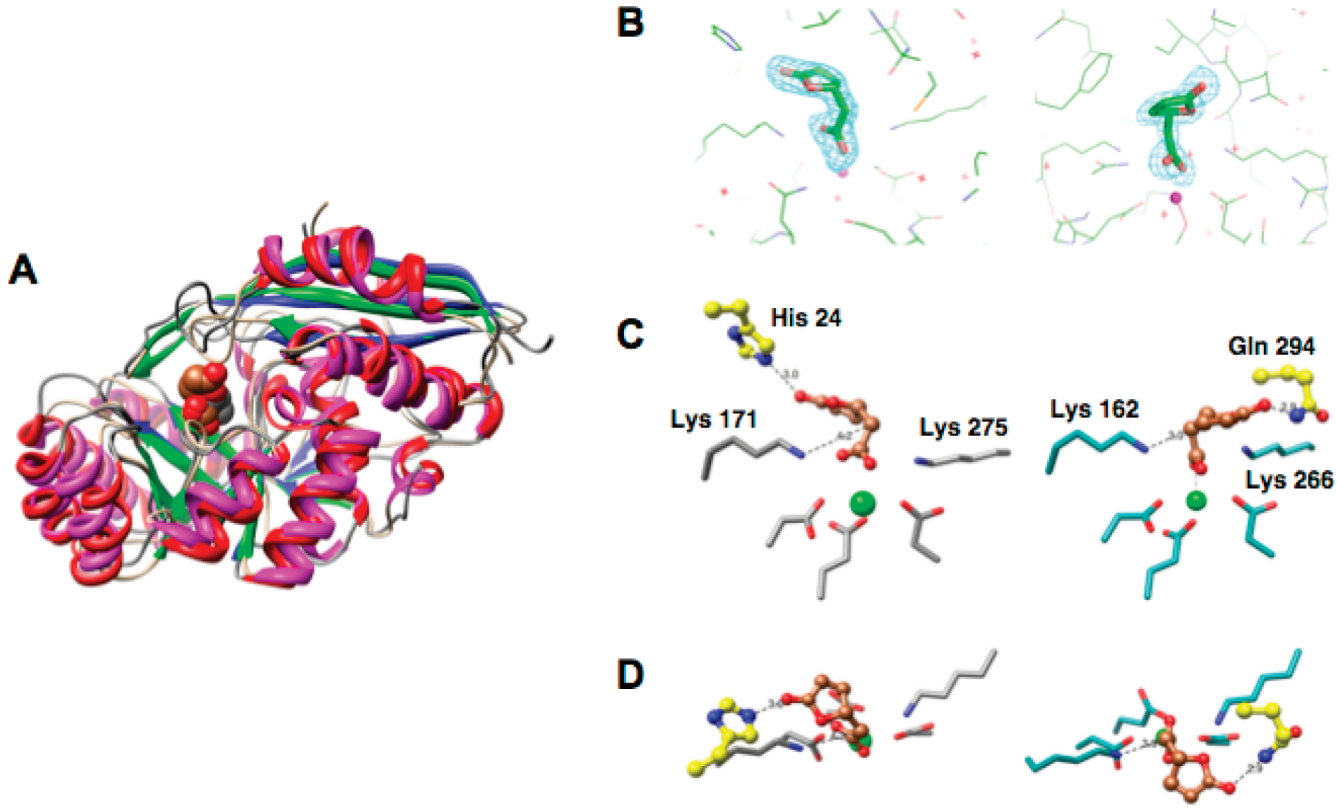

FIGURE 5: (A) Superposition of the polypeptides of the MLEs from *P. fluorescens* and *M. smegmatis*. The backbone of the MLE from *P. fluorescens* is colored gray; the α-helices are colored red, and the *β*-strands are colored blue. The backbone of the MLE from *M. smegmatis* is colored tan; the α-helices are colored magenta, and the *β*-strands are colored green. The (4*S*)-muconolactone ligands are represented by the space-filling structures. (B) Electron density for the (4*S*)-muconolactone ligand in (left) the active site of the MLE from *P. fluorescens* and (right) the active site of the MLE from *M. smegmatis*. (C and D) Left, side, and top views of the active site of the MLE from *P. fluorescens*. Right, side, and top views of the active site of the MLE from *M. smegmatis*. The (4*S*)-muconolactone ligands are colored brown.

and, therefore, do not require the MLE function. The functions of these proteins are of interest; e.g., perhaps they catalyze a promiscuous MLE reaction, and are under investigation.

If this hypothesis is correct, the stereochemically distinct families of MLEs would provide the second example of pseudoconvergent evolution of function within the superfamily; i.e., the *syn*- and *anti*-MLEs evolved from different ancestors within the MLE subgroup. Of course, these intermediate progenitors would have evolved divergently from the common progenitor of the MLE subgroup. We previously identified two evolutionarily distinct families of NSARs, also in the MLE subgroup, one more closely related to the OSBS family (*15*−*17*) and the second more closely related to the dipeptide epimerases (*11*), suggesting that Nature can evolve the same overall reaction from different starting points, perhaps providing a useful lesson for in vitro enzyme engineering (*31*).

Support for this interpretation is provided by experiments we previously described. In vitro design followed by directed evolution was used to change enzymatic functions in the MLE subgroup by changing specificity-determining residues; i.e., we introduced the OSBS function into the AEE structural scaffold (*12*−*14*). Now, with the discovery of stereochemically distinct families of homologous MLEs, we conclude that substitution for specificity-determining residues is a strategy that has, in fact, been used in the natural evolution of new functions.

REFERENCES

1. Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. (1990) Mandelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologous. *Nature 347*, 692–694.
2. Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., and Gerlt, J. A. (1996) The enolase superfamily: A general strategy for enzyme-catalyzed abstraction of the α-protons of carboxylic acids. *Biochemistry 35*, 16489–16501.
3. Gerlt, J. A., Babbitt, P. C., and Rayment, I. (2005) Divergent evolution in the enolase superfamily: The interplay of mechanism and specificity. *Arch. Biochem. Biophys. 433*, 59–70.
4. Avigad, G., and England, S. (1984) *Fed. Proc. 28*, 345 (Abstract 486).
5. Chari, R. V. J., Whitman, C. P., Kozarich, J. W., Ngai, K.-L., and Ornston, L. N. (1987) Absolute stereochemical coure of the 3-carboxylmuconate cycloisomerases from *Pseudomonas putida* and *Acinetobacter calcoaceticus*: Analysis and implications. *J. Am. Chem. Soc. 109*, 5514–5519.
6. Hasson, M. S., Schlichting, I., Moulai, J., Taylor, K., Barrett, W., Kenyon, G. L., Babbitt, P. C., Gerlt, J. A., Petsko, G. A., and Ringe, D. (1998) Evolution of an enzyme active site: The structure of a new crystal form of muconate lactonizing enzyme compared with mandelate racemase and enolase. *Proc. Natl. Acad. Sci. U.S.A. 95*, 10396–10401.
7. Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) Evolution of enzymatic activity in the enolase superfamily: Structure of o-succinylbenzoate synthase from *Escherichia coli* in complex with Mg^{2+} and o-succinylbenzoate. *Biochemistry 39*, 10662–10676.
8. Schmidt, D. M., Hubbard, B. K., and Gerlt, J. A. (2001) Evolution of Enzymatic Activities in the Enolase Superfamily: Functional

Assignment of Unknown Proteins in *Bacillus subtilis* and *Escherichia coli* as L-Ala-D/L-Glu Epimerases. *Biochemistry 40*, 15707–15715.

9. Kalyanaraman, C., Imker, H. J., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Babbitt, P. C., Almo, S. C., Gerlt, J. A., and Jacobson, M. P. (2008) Discovery of a dipeptide epimerase enzymatic function guided by homology modeling and virtual screening. *Structure 16*, 1668–1677.
10. Sakai, A., Xiang, D. F., Xu, C., Song, L., Yew, W. S., Raushel, F. M., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: N-Succinylamino acid racemase and a new pathway for the irreversible conversion of D- to L-amino acids. *Biochemistry 45*, 4455–4462.
11. Song, L., Kalyanaraman, C., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Brown, S., Imker, H. J., Babbitt, P. C., Almo, S. C., Jacobson, M. P., and Gerlt, J. A. (2007) Prediction and assignment of function for a divergent N-succinyl amino acid racemase. *Nat. Chem. Biol. 3*, 486–491.
12. Schmidt, D. M., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary potential of $(\beta/\alpha)_8$-barrels: Functional promiscuity produced by single substitutions in the enolase superfamily. *Biochemistry 42*, 8387–8393.
13. Vick, J. E., Schmidt, D. M., and Gerlt, J. A. (2005) Evolutionary Potential of $(\beta/\alpha)_8$-Barrels: In Vitro Enhancement of a "New" Reaction in the Enolase Superfamily. *Biochemistry 44*, 11722–11729.
14. Vick, J. E., and Gerlt, J. A. (2007) Evolutionary potential of $(\beta/\alpha)_8$-barrels: Stepwise evolution of a "new" reaction in the enolase superfamily. *Biochemistry 46*, 14589–14597.
15. Glasner, M. E., Fayazmanesh, N., Chiang, R. A., Sakai, A., Jacobson, M. P., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of structure and function in the o-succinylbenzoate synthase/N-acylamino acid racemase family of the enolase superfamily. *J. Mol. Biol. 360*, 228–250.
16. Taylor Ringia, E. A., Garrett, J. B., Thoden, J. B., Holden, H. M., Rayment, I., and Gerlt, J. A. (2004) Evolution of enzymatic activity in the enolase superfamily: Functional studies of the promiscuous o-succinylbenzoate synthase from *Amycolatopsis*. *Biochemistry 43*, 224–229.
17. Thoden, J. B., Taylor Ringia, E. A., Garrett, J. B., Gerlt, J. A., Holden, H. M., and Rayment, I. (2004) Evolution of Enzymatic Activity in the Enolase Superfamily: Structural Studies of the Promiscuous o-Succinylbenzoate Synthase from *Amycolatopsis*. *Biochemistry 43*, 5716–5727.
18. Galperin, M. Y., Walker, D. R., and Koonin, E. V. (1998) Analogous enzymes: Independent inventions in enzyme evolution. *Genome Res. 8*, 779–790.
19. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res. 13*, 2498–2504.
20. Edgar, R. C. (2004) MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res. 32*, 1792–1797.
21. Ronquist, F., and Huelsenbeck, J. P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics 19*, 1572–1574.
22. Altekar, G., Dwarkadas, S., Huelsenbeck, J. P., and Ronquist, F. (2004) Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics 20*, 407–415.
23. Whelan, S., and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol. 18*, 691–699.
24. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., and Simon, M. I., Eds.) pp 307−326, Academic Press, New York.
25. Long, F., Vagin, A. A., Young, P., and Murshudov, G. N. (2008) BALBES: A molecular-replacement pipeline. *Acta Crystallogr. D64*, 125–132.
26. Jones, A. T. (1985) Interactive computer graphics: FRODO. *Methods Enzymol. 115*, 157–171.
27. Brunger, 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. D54*, 905–921.
28. Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models. *Acta Crystallogr. D49*, 129–147.
29. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 1605–1612.
30. Murakami, S., Kohsaka, C., Okuno, T., Takenaka, S., and Aoki, K. (2004) Purification, characterization, and gene cloning of cis,cis-muconate cycloisomerase from benzamide-assimilating *Arthrobacter* sp. BA-5-17. *FEMS Microbiol. Lett. 231*, 119–124.
31. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies. *Curr. Opin. Chem. Biol. 10*, 492–497.

BI802277H