Biochemistry

 Volume 42, Number 28 July 22, 2003

Accelerated Publications

Evolutionary Potential of $(\beta/\alpha)_8$-Barrels: Functional Promiscuity Produced by Single Substitutions in the Enolase Superfamily†

Dawn M. Z. Schmidt,‡ Emily C. Mundorff,§ Michael Dojka,§ Ericka Bermudez,§ Jon E. Ness,§,|| Sridhar Govindarajan,§,|| Patricia C. Babbitt,⊥ Jeremy Minshull,§,|| and John A. Gerlt*,‡

Departments of Biochemistry and Chemistry, 419 Roger Adams Laboratory, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801, Maxygen, Inc., 515 Galveston Drive, Redwood City, California 94063, and Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, University of California, San Francisco, California 94143

Received May 12, 2003; Revised Manuscript Received May 29, 2003

ABSTRACT: The members of the mechanistically diverse, $(\beta/\alpha)_8$-barrel fold-containing enolase superfamily evolved from a common progenitor but catalyze different reactions using a conserved partial reaction. The molecular pathway for natural divergent evolution of function in the superfamily is unknown. We have identified single-site mutants of the $(\beta/\alpha)_8$-barrel domains in both the L-Ala-D/L-Glu epimerase from *Escherichia coli* (AEE) and the muconate lactonizing enzyme II from *Pseudomonas* sp. P51 (MLE II) that catalyze the *o*-succinylbenzoate synthase (OSBS) reaction as well as the wild-type reaction. These enzymes are members of the MLE subgroup of the superfamily, share conserved lysines on opposite sides of their active sites, but catalyze acid- and base-mediated reactions with different mechanisms. A comparison of the structures of AEE and the OSBS from *E. coli* was used to design the D297G mutant of AEE; the E323G mutant of MLE II was isolated from directed evolution experiments. Although neither wild-type enzyme catalyzes the OSBS reaction, both mutants complement an *E. coli* OSBS auxotroph and have measurable levels of OSBS activity. The analogous mutations in the D297G mutant of AEE and the E323G mutant of MLE II are each located at the end of the eighth β-strand of the $(\beta/\alpha)_8$-barrel and alter the ability of AEE and MLE II to bind the substrate of the OSBS reaction. *The substitutions relax the substrate specificity, thereby allowing catalysis of the mechanistically diverse OSBS reaction with the assistance of the active site lysines.* The generation of functionally promiscuous and mechanistically diverse enzymes via single-amino acid substitutions likely mimics the natural divergent evolution of enzymatic activities and also highlights the utility of the $(\beta/\alpha)_8$-barrel as a scaffold for new function.

The number of enzymes far exceeds the number of protein folds, with estimates of the number of protein folds at approximately 1000 (*1*). The $(\beta/\alpha)_8$-barrel, or TIM barrel, fold is the most common structure in the Protein Data Bank (*2*), accounting for at least 10% of all enzymes of known structure (*3*, *4*). This fold provides the scaffold for diverse enzymatic functions; the current SCOP database (1.61) lists 25 functionally distinct $(\beta/\alpha)_8$-barrel-containing superfamilies. The common location of the active sites in $(\beta/\alpha)_8$-barrels at the C-terminal ends of the β-strands suggests that many enzymes sharing this fold are related by divergent evolution from a limited number of progenitors (*4*, *5*).

† This research was supported by Grants GM-52594 (to J.A.G.) and GM-60595 (to P.C.B.) from the National Institutes of Health.

* To whom correspondence should be addressed: Department of Biochemistry, University of Illinois at Urbana-Champaign, 600 S. Mathews Ave., Urbana, IL 61801. Phone: (217) 244-7414. Fax: (217) 265-0385. E-mail: j-gerlt@uiuc.edu.

‡ University of Illinois at Urbana-Champaign.

§ Maxygen, Inc.

|| Present address: DNA 2.0, 1455 Adams Dr., Menlo Park, CA.

⊥ University of California.

10.1021/bi034769a CCC: $25.00
Published on Web 06/24/2003

Divergent evolution of enzyme function is commonly explained by gene duplication followed by mutational events that allow the protein encoded by the copy to acquire a new function (*6*). In one variation of this process, the gene duplication event is preceded by a period of "gene sharing" during which the enzyme is functionally promiscuous, i.e., catalyzes multiple reactions (*7*); the additional reaction(s) may or may not be physiologically important (*8*). After gene duplication, the promiscuous reaction catalyzed by the copy can provide a selective advantage if improved activity can be accessed by one or two additional mutations (*9*). In agreement with this view, naturally promiscuous enzymes that catalyze different chemical reactions have been discovered (*8*−*10*). In addition, enzymes catalyzing the same chemical reaction but with promiscuous substrate specificities occur as intermediates in directed evolution experiments (*11*, *12*).

In another variation, the progenitor catalyzes a single reaction, but an alternate chemical reaction is achieved by a limited number of mutations following gene duplication. Functional interconversion of homologous enzymes that catalyze different chemical reactions has been accomplished *in vitro* with as few as four to eight substitutions (*13*, *14*). However, the natural acquisition of a new reaction with this number of substitutions is unlikely; during random accumulation of the required mutations, deleterious mutations are also expected to occur. In other words, the probability of evolving a new enzymatic function that requires specific multiple mutations is low.

To explore divergent evolution of $(\beta/\alpha)_8$-barrels, we have investigated whether single-site mutants of enzymes possessing the $(\beta/\alpha)_8$-barrel fold are sufficient for generation of a new chemical reaction. Such mutants would confirm the evolutionary potential of the $(\beta/\alpha)_8$-barrel fold and support the divergent evolution of mechanistically diverse superfamilies, including the enolase superfamily.

The enolase superfamily is a mechanistically diverse enzyme superfamily whose homologous members catalyze different overall reactions but retain a common catalytic strategy (*15*): abstraction of a proton from the carbon adjacent to the carboxylate group (α-proton) of the substrate to form a dianionic enediolate intermediate stabilized by an essential Mg^{2+} (*16*). The conserved location for the active site is at the interface between the C-terminal end of the $(\beta/\alpha)_8$-barrel domain where the catalytic residues are located and an overlying "capping domain" formed by segments from the N- and C-termini that influences substrate specificity.

Although levels of pairwise sequence identity relating heterofunctional members are usually less than 30%, key active site residues are conserved, allowing the superfamily to be divided into three subgroups (*16*). The isofunctional enolase subgroup is characterized by a conserved general basic Lys at the end of the sixth β-strand of the $(\beta/\alpha)_8$-barrel. The functionally diverse members of the mandelate racemase subgroup share a general basic His-Asp dyad in which the His at the end of the seventh β-strand is spatially homologous to the conserved Lys of the enolase subgroup. The functionally diverse members of the muconate lactonizing enzyme (MLE)[1] subgroup share acidic/basic Lys residues at the ends of the second and sixth β-strands. Members of all three subgroups possess conserved Mg^{2+}-binding ligands at the ends of the third, fourth, and fifth β-strands. We have used

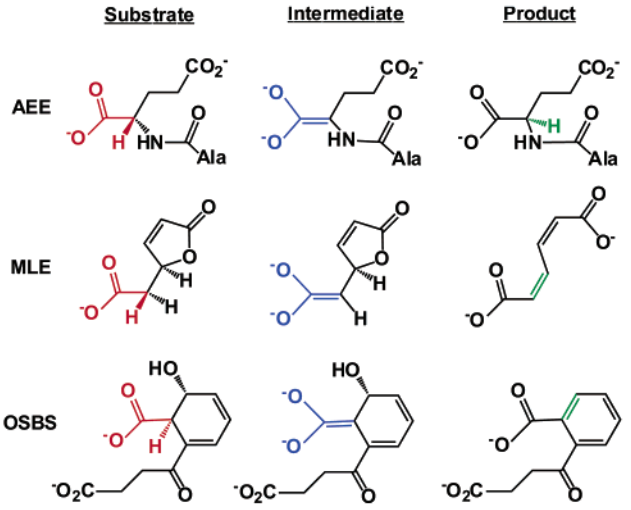

FIGURE 1: Reactions catalyzed by members of the MLE subgroup of the enolase superfamily: 1,1-proton transfer (AEE), cycloisomerization (MLE II), and β-elimination/dehydration (OSBS).

our understanding of these structure−function paradigms to assign function to unknown members discovered in genomes (*17*, *18*).

The members of the MLE subgroup catalyze three different chemical reactions: 1,1-proton transfer [L-Ala-D/L-Glu epimerase (AEE)], cycloisomerization (MLE), and β-elimination [*o*-succinylbenzoate synthase (OSBS)] (Figure 1). For our experiments aimed at investigating the evolutionary potential of the $(\beta/\alpha)_8$-barrel, we selected the OSBS reaction as our target. OSBS is part of the pathway for menaquinone biosynthesis and is required for anaerobic growth by *Escherichia coli* and other microbes, thereby providing a metabolic selection for evolved OSBS activity by complementation of a strain of *E. coli* lacking a functional gene encoding OSBS, *menC*. As progenitors for *in vitro* evolution we selected the AEE from *E. coli* (*18*) and the MLE II from *Pseudomonas* sp. P51 that catalyzes the cycloisomerization of 2-chloro-*cis*,*cis*-muconate and, with decreased kinetic constants, the cycloisomerization of *cis*,*cis*-muconate (the MLE I reaction shown in Figure 1) (*19*). *Neither progenitor catalyzes the OSBS reaction.* Although the substrates, products, and overall reactions for these three enzymes are markedly different, their active sites share conserved geometries (Figure 2). We have identified *single substitutions* of the $(\beta/\alpha)_8$-barrel domains in both AEE and MLE II that allow catalysis of the OSBS reaction as well as the starting reaction. This generation of functional promiscuity *via* single substitutions demonstrates a remarkably simple pathway for evolution of new enzymatic reactions and confirms the inherent functional capability of the $(\beta/\alpha)_8$-barrel fold for both natural and "unnatural" protein engineering.

MATERIALS AND METHODS

Construction of Insertional Deletions in E. coli and Site-Directed Mutagenesis. The method of Datsenko and Wanner

[1] Abbreviations: AEE, L-Ala-D/L-Glu epimerase; MLE, muconate lactonizing enzyme; OSB, *o*-succinylbenzoate; OSBS, *o*-succinylbenzoate synthase; ProFAR, *N*′-[(5′-phosphoribosyl)formimino]-5-aminoimidazole-4-carboxamide ribonucleotide; PRA, phosphoribosylanthranilate; SHCHC, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate.

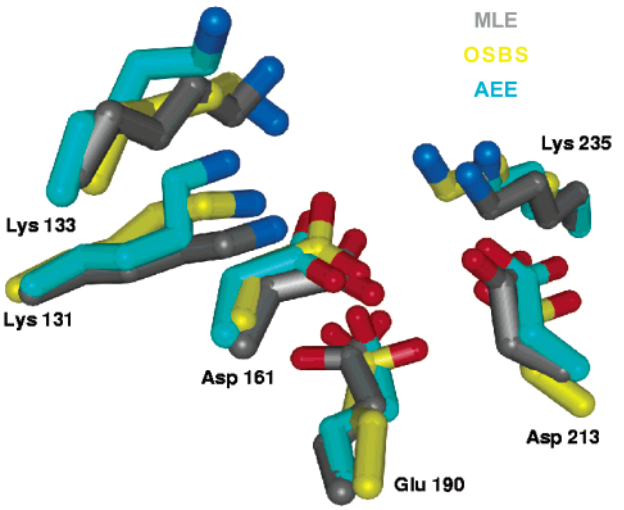

FIGURE 2: Superposition of the active sites of MLE II (*29*) (gray), OSBS from *E. coli* (*21*) (yellow), and AEE from *E. coli* (*22*) (cyan). The numbering is that of OSBS from *E. coli*. Lys131 and Lys133 are at the end of the second β-strand. Asp161 is at the end of the third β-strand. Glu190 is at the end of the fourth β-strand. Asp213 is at the end of the fifth β-strand. Lys235 is at the end of the sixth β-strand. The superpositions were generated with the MINRMS algorithm and visualized with the CHIMERA visualization package provided by the Computer Graphics Laboratory, University of California, San Franicisco, CA.

(*20*) was used to construct insertional deletions of genes *menC* (encoding OSBS) and *ycjG* (encoding AEE) in *E. coli* strain BW25113. A 500 bp region in the middle of the approximately 1100 bp *menC* gene was replaced with a fragment encoding kanamycin resistance (*menC::kan)*. A 400 bp region in the middle of the approximately 1000 bp *ycjG* gene was replaced with a fragment encoding chloramphenicol resistance (*ycjG::cam*). The location of the antibiotic resistance cassette in each knockout was verified by DNA sequencing of the junctions between the cassette and the flanking regions of the *menC* and *ycjG* genes. Site-directed mutants were created using the overlap extension PCR method.

Random Mutagenesis and Selection of MLE Mutants. The wild-type gene for MLE II from *Pseudomonas* sp. P51 was amplified by PCR, fragmented, reassembled, and rescued as previously described (*21*). The resulting shufflants were cloned into the *Eco*RI and *Pst*I sites of the broad host range vector pMMB66EH (ATCC 37620) (*22*) and transformed into electrocompetent *E. coli* DH10B (Stratagene). The plasmid was prepared from 10 000 independent colonies; 0.2 μg of this DNA mixture was used to transform *E. coli menC::kan*, and transformed cells were spread onto minimal agar media with trimethylamine *N*-oxide as the terminal electron acceptor and glycerol as the carbon source (*23*), as well as 1 mM IPTG and 50 μg/mL carbenicillin, and incubated anaerobically at 37 °C. Plasmid DNA prepared from transformants that grew anaerobically was retransformed into the *menC::kan* strain, and the growth phenotype was confirmed. To rule out spontaneous mutation in *E. coli* as the source of complementation, plasmid pMMB66EH expressing unmutagenized MLEII was similarly pooled, transformed into the *menC::kan* strain, and tested for complementation.

Protein Expression and Purification. Proteins assayed for OSBS activity were purified from the *menC::kan* strain of *E. coli*; proteins assayed for AEE activity were purified from the *ycjG::cam* strain. This was done to prevent contamination of OSBS-assayed enzymes by the chromosomally encoded OSBS or AEE-assayed enzymes by the chromosomally encoded AEE. In either case, all genes encoding wild-type and mutant proteins were cloned into the *Nde*I and *Bam*HI sites of pDMS-1a (a derivative of pKK223-3 with a modified multiple cloning site), and the recombinant plasmids were transformed into the *menC::kan* or *ycjG::cam* strain of *E. coli* as appropriate. Cultures for expression were grown in 2 L of LB at 37 °C and induced with 0.5 mM IPTG. Cells were harvested by centrifugation, resuspended in resuspension buffer [10 mM Tris-HCl (pH 7.5 for MLE II variants and pH 8.5 for AEE variants) and 5 mM $MgCl_2$], and lysed by sonication; after centrifugation, the lysate was applied to a DEAE-Sepharose FF column. The column was washed with 2.5 column volumes of resuspension buffer, and the protein was eluted with a 1600 mL linear gradient (from 0 to 50%) of 1 M NaCl in resuspension buffer and collected in 20 mL fractions. Fractions containing the desired MLE II or AEE variant were pooled, concentrated, and dialyzed against fresh resuspension buffer. The dialyzed protein was then purified further over a Source 15Q column, eluted with a 240 mL linear gradient (from 0 to 50%) of 1 M NaCl in resuspension buffer, and collected in 8 mL fractions. Fractions containing pure protein were pooled, concentrated, and dialyzed against resuspension buffer containing 50 mM NaCl before use in activity assays.

Circular Dichroism Experiments. The secondary structures of lysine mutants of the D297G mutant of AEE and of the E323G mutant of MLE II were compared with those of the progenitors using circular dichroism (CD) spectra recorded with a Jasco J720 spectropolarimeter at the Laboratory of Fluorescence Dynamics at the University of Illinois at Urbana-Champaign. The protein concentrations were adjusted to give a final absorbance of 0.4 at 280 nm [in a 1 cm path length cuvette in a Perkin-Elmer Lambda 2S spectrophotometer in 50 mM Tris (pH 8.5)]. CD measurements were taken at room temperature using a 0.1 cm path length cuvette. Spectra were acquired from 250 to 200 nm at a rate of 50 nm/min, with a step resolution of 0.5 nm, a response time of 4.0 s, and a bandwidth of 1.0 nm. For each protein, a baseline scan (of buffer) was subtracted from the average of four scans to give the final averaged scan.

Enzymatic Activity Assays. OSBS activity was measured as previously described (*8*) in 50 mM Tris (pH 8.0) and 0.1 mM $MnCl_2$ for the E323G mutant of MLE II and 50 mM Tris (pH 7.0) and 0.1 mM $MnCl_2$ for the D297G mutant of AEE because of differences in their pH optima. MLE activity was measured by following the decrease in absorbance at 260 nm due to the consumption of *cis*,*cis*-muconate (ϵ_{260} = 16 900 M^{-1} cm^{-1}) in 50 mM Tris (pH 7.5) and 0.1 mM $MnCl_2$. AEE activity was measured as described previously at pH 8.5 (*18*). Values for K_M and k_{cat} were determined by varying the concentration of substrate at a constant enzyme concentration. Assays for OSBS activity using [α-^{2}H]-SHCHC (*24*) as a substrate to determine substrate kinetic isotope effects were performed in a fashion identical to that for assays using a protiated substrate. The values of the kinetic constants were determined with the program CLELAND (*25*).

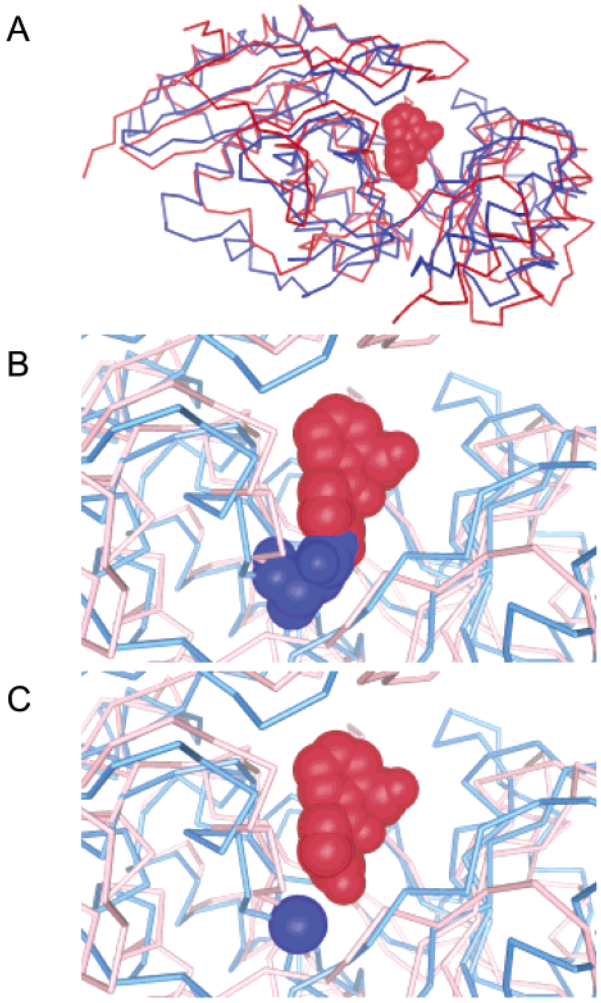

FIGURE 3: (A) Superposition of the structures of the AEE from *E. coli* (blue) and the OSBS from *E. coli*, the latter with the OSB product bound in the active site. (B) Superposition showing the overlap of Asp297 in the active site of AEE (blue) with the OSB product in the active site of OSBS (red). (C) Predicted effect of the D297G mutant of AEE (blue) on removal of the structural overlap.

RESULTS

Rational Design of AEE D297G. We superimposed the structures of the unliganded AEE (*26*) and the OSB-liganded OSBS (*27*) from *E. coli*, whose sequences are 30% identical (Figure 3A). Focusing on residues at the C-terminal ends of the β-strands in the $(\beta/\alpha)_8$-barrel domain, we observed residue differences at three of the eight β-strands. Most notable is the substitution of a Gly at the end of the eighth β-strand in OSBS (Gly288) with an Asp in AEE (Asp297). In our superposition, Asp297 in AEE overlaps the succinyl moiety of the OSB product in the active site of OSBS (Figure 3B). We constructed the D297G mutant of AEE with the expectation that this change would allow the SHCHC substrate to bind and undergo dehydration (Figure 3C). The gene encoding the wild-type AEE does not complement the *menC::kan* OSBS auxotroph (Figure 4), but the gene encoding the D297G mutant of AEE does, albeit at a growth rate slower than that of wild-type *E. coli*.

Directed Evolution of OSBS Activity in MLE II. Independently and in parallel with our design of OSBS activity in AEE, we used directed evolution to identify a mutant of MLE II from *Pseudomonas* sp. P54 that complements the OSBS auxotroph. The sequences of OSBS from *E. coli* and MLE II are 24% identical. The gene encoding MLE II was mutagenized by fragmentation and PCR reassembly. The progenitor did not complement the auxotroph (Figure 4); of three mutants that complemented it, one contained only a single mutation, E323G (the other two were multiple mutants that contained E323G as well as additional mutations). In this case, the E323G44 mutant of MLE II rapidly complemented the knockout (Figure 4), with a growth curve more similar to that of wild-type *E. coli*. Glu323 is located at the end of the eighth β-strand, so the E323G substitution is analogous to the D297G substitution designed in AEE.

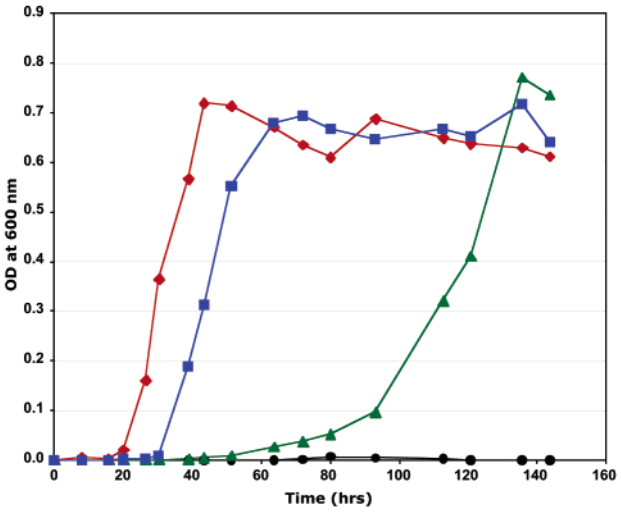

FIGURE 4: Anaerobic growth of various strains of *E. coli*: wild-type BW25113 (◆), the *menC::kan* mutant of BW25113 (●), the *menC::kan* mutant transformed with a plasmid encoding the D297G mutant of AEE (▲), and the *menC::kan* mutant transformed with a plasmid encoding the E323G mutant of MLE II (■). The data for the *menC::kan* mutant transformed with plasmids encoding wild-type AEE and wild-type MLE II can be superimposed with those shown for the *menC::kan* mutant and are omitted for clarity.

Quantitation of Enzymatic Activities. The progenitor enzymes and mutants were expressed in the *menC::kan* strain of *E. coli*, purified, and assayed for both OSBS activity and progenitor activity (Table 1). The wild-type AEE and MLE II have no detectable OSBS activity, but both mutants showed significant activity relative to that of the uncatalyzed reaction (*24*). The differences in the rates of anaerobic growth during complementation (Figure 4) qualitatively correlate with the values of k_{cat} for the OSBS activities of the mutants, suggesting that the intracellular concentrations of the SHCHC substrate are equivalent so menaquinone synthesis is limited by the values of k_{cat} (*29*). The rate acceleration (ratio of k_{cat} to k_{uncat}) is 10^6 for the D297G mutant of AEE and 10^9 for the E323G mutant of MLE II. The rate acceleration is 10^{10} for the natural OSBS from *E. coli* (*24*). The proficiency [ratio of k_{cat}/K_M to the uncatalyzed rate (*28*)] is 13 kcal/mol for the D297G mutant of AEE, 16 kcal/mol for the E323G mutant of MLE II, and 20 kcal/mol for the natural OSBS (Figure 5). Both mutants are functionally promiscuous (*9*), with the ability to catalyze the OSBS reaction accompanied by a reduction in the level of the progenitor reaction. In addition, because the members of the MLE subgroup catalyze three different reactions (1,1-proton transfer, dehydration, and cycloisomerization, Figure 1), we also assessed the ability

Table 1: Kinetic Parameters for OSBS, MLE, and AEE Activities[a]

enzyme	k_{cat} (s^{-1})	K_M (M)	k_{cat}/K_M (M^{-1} s^{-1})
	OSBS Activity		
E. coli OSBS[b]	24 ± 0.8[c]	$(8 \pm 1.7) \times 10^{-6}$	3.1×10^6
wild-type MLE II	–	–	$<1.5 \times 10^{-3}$ [d]
E323G MLE II	1.5 ± 0.1[c]	$(8.1 \pm 0.4) \times 10^{-4}$	1.9×10^3
E323G/K165A MLE II	–	–	$<1.5 \times 10^{-3}$ [d]
E323G/K269A MLE II	–	–	6.0×10^{-2} [e]
wild-type AEE	–	–	$< 5.2 \times 10^{-3}$ [d]
D297G AEE	$(2.5 \pm 0.3) \times 10^{-3}$ [c]	$(2.0 \pm 0.1) \times 10^{-4}$	12.5
D297G/K151A AEE	–	–	0.11[e]
D297G/K247A AEE	–	–	5.5×10^{-2} [f]
	MLE Activity		
wild-type MLE II	10.8 ± 0.5	$(5.4 \pm 0.5) \times 10^{-4}$	2.0×10^4
E323G MLE II	2.3 ± 0.2	$(1.7 \pm 0.2) \times 10^{-3}$ [g]	1.3×10^3
K165A MLE II	–	–	7.7[f]
K269A MLE II	–	–	0.25[f]
E323G/K165A MLE II	–	–	2.5×10^{-2} [f]
E323G/K269A MLE II	–	–	0.12[f]
wild-type AEE	–	–	$<5.0 \times 10^{-3}$ [d]
D297G AEE	–	–	$<6.7 \times 10^{-3}$ [d]
	AEE Activity		
wild-type *E. coli* AEE[e]	10 ± 0.4	$(1.3 \pm 0.3) \times 10^{-4}$	7.7×10^4
D297G AEE	0.043 ± 0.005	$(4.4 \pm 0.3) \times 10^{-3}$	9.8
K151A AEE	–	–	<0.019[d]
K247A AEE	$(2.1 \pm 0.1) \times 10^{-3}$	$(5.1 \pm 1.0) \times 10^{-5}$	42
D297G/K151A AEE	$(8.5 \pm 1.0) \times 10^{-4}$	$(1.8 \pm 0.4) \times 10^{-3}$	0.47
D297G/K247A AEE	$(1.1 \pm 0.1) \times 10^{-3}$	$(2.7 \pm 0.8) \times 10^{-4}$	0.41
wild-type MLE II	–	–	<0.031[d]
E323G MLE II	–	–	<0.031[d]

[a] Assay conditions described in Materials and Methods. [b] E. Taylor Ringia, unpublished results. [c] For comparison, $k_{non} = 1.6 \times 10^{-10}$ s^{-1}. [d] Lower limit of detection for activity. [e] From ref *8*. [f] Insufficient activity to determine values of k_{cat} and K_M. [g] MLE assay limited to substrate concentrations of ≤1 mM.

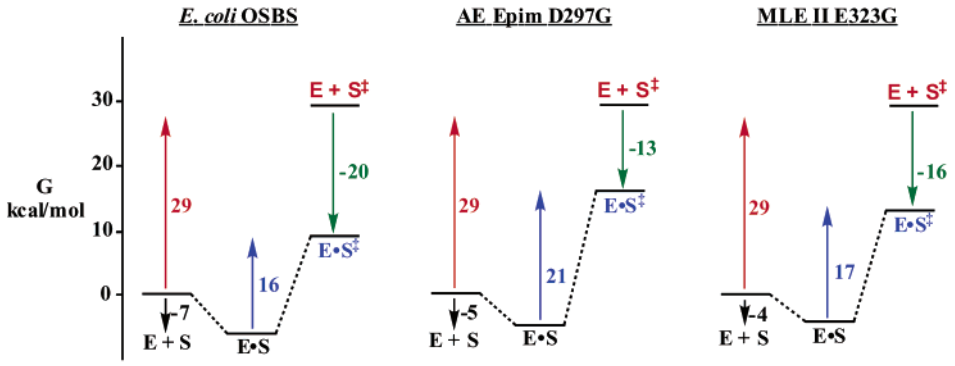

FIGURE 5: Reaction coordinates for the reactions catalyzed by the OSBS from *E. coli* (left panel), the D297G mutant of AEE (center panel), and the E323G mutant of MLE II (right panel). The activation free energy associated with the uncatalyzed OSBS reaction is shown in red; the activation free energies associated with the catalyzed OSBS reactions are shown in blue, and the free energies associated with the proficiencies are shown in green.

of both promiscuous mutants to catalyze the third reaction. The D297G mutant of AEE had no detectable MLE activity, and the E323G mutant of MLE II had no detectable AEE activity (Table 1).

Mutation of Catalytic Residues. As members of the MLE subgroup, the active sites of OSBS, AEE, and MLE II contain Lys residues on opposite faces at the ends of the second and sixth β-strands (Lys133 and Lys235 in OSBS, Lys151 and Lys247 in AEE, and Lys165 and Lys269 in MLE II); separate substitution of these with Ala in the wild-type proteins significantly impairs the natural reaction, consistent with their participation as acid/base catalysts (Table 1 and E. A. Taylor Ringia and J. A. Gerlt, unpublished results). The analogous Ala substitutions were introduced into the promiscuous D297G mutant of AEE and the E323G mutant of MLE II, and the purified proteins were assayed for the progenitor and promiscuous activities (Table 1). Circular dichroism spectra for all mutants were identical to those of the progenitors (data not shown). Consistent with the results obtained for the natural functions, each Ala substitution compromises both the natural and promiscuous activities.

Substrate Kinetic Isotope Effects. Both promiscuous mutants were assayed with [α-^{2}H]SHCHC as a substrate for the OSBS reaction. The kinetic isotope effects on k_{cat} (5.3 ± 1.1 for the D297G mutant of AEE and 4.4 ± 0.7 for the E323G mutant of MLE II) are similar to that measured for the natural OSBS (6.1 ± 0.4) (*24*). Taken together, these results and those for the active site lysine mutants suggest

that the mechanisms of the promiscuous evolved OSBS mutants are similar to that of the natural OSBS-catalyzed reaction, despite the reduced kinetic constants.

DISCUSSION

In an effort to elucidate the pathway for divergent evolution of enzymatic function, particularly of $(\beta/\alpha)_8$-barrels, we have isolated functionally promiscuous, single-amino acid mutants of two members of the enolase superfamily. The D297G mutant of AEE and the E323G mutant of MLE II catalyze the dehydration of SHCHC to OSB (the reaction catalyzed by a third enolase superfamily member, OSBS), as well as their progenitor reactions, 1,1-proton transfer and cycloisomerization, respectively. The amino acid changes for each mutant are analogous in sequence and structure; an acidic residue at the end of the eighth $\beta-$strand of the $(\beta/\alpha)_8$-barrel is changed to a glycine, presumably accommodating the succinyl side chain of the substrate and product of the OSBS reaction.

The quantitative comparison of the reaction coordinates for these reactions presented in Figure 5 succinctly illustrates the evolutionary potential for catalytic diversity in the enolase superfamily. The rate accelerations (ratios of k_{cat} to k_{uncat}) and proficiencies (ratios of k_{cat}/K_M to k_{uncat}) for both the D297G mutant of AEE and the E323G mutant of MLE II are remarkably similar to those that characterize the reaction catalyzed by the natural OSBS. If natural divergent evolution of mechanistic diversity is to be accomplished [established by the existence of the enolase and other mechanistically diverse superfamilies (*15*)], only a limited number of substitutions can be required for selective advantage; otherwise, the probability of accumulating the necessary mutations would be too low. Our results establish that single substitutions are sufficient to achieve levels of new activities that allow growth advantage; with time, additional substitutions will "perfect" the energetics of the reaction coordinate of the evolving activity (*29*). However, without the remarkable catalytic activity provided by the first substitution, evolution of new enzymatic functions by divergent evolution would be improbable, if not impossible.

Single-amino acid substitutions in other $(\beta/\alpha)_8$-barrel-containing enzymes that result in new activities have been reported. HisA (ProFAR isomerase) and TrpF (PRA isomerase) whose sequences are ~10% identical catalyze Amadori rearrangements in the biosynthetic pathways for histidine and tryptophan, respectively; the enzymes are specific for their substrate and do not catalyze the other reaction. Using directed evolution, a point mutant of HisA from *Thermotoga maritima* (D127V) was identified that catalyzed the TrpF reaction but not the HisA reaction (a mutant with four substitutions was promiscuous for both reactions) (*30*). In more recent work, dihydrodipicolinate synthase (DHDPS) activity was rationally enhanced in the *N*-acetylneuraminate lyase (NAL) scaffold via a single substitution, L142R (*31*). NAL and DHDPS belong to the same superfamily of $(\beta/\alpha)_8$-barrels, share <24% sequence identity, and catalyze the aldol condensation of pyruvate with different aldehydes. NAL is naturally promiscuous and catalyzes the DHDPS reaction at a low level; L142R catalyzes the DHDPS reaction with a 19-fold increased value of k_{cat}/K_M.

In both of these, as well as in other (*32*), experiments, a change in function is the result of a change in substrate specificity, with the mechanism of the reaction remaining the same. The *in vitro* evolution experiments presented here are distinct and noteworthy because the single substitutions in MLE II and AEE modified both the substrate specificity *and* the mechanism of the catalyzed reaction.

The active sites of MLE, AEE, and OSBS are remarkably conserved with respect to the positions of the (1) the ligands at the ends of the third, fourth, and fifth β-strands for the essential Mg^{2+} that stabilizes the enediolate intermediate and (2) the Lys catalysts on opposite faces of the active site at the ends of the second and sixth β-strands (Figure 2). The homologous active sites appear to be "hard-wired" for acid/base-catalyzed reactions on both faces of the substrates and the Mg^{2+}-stabilized enediolate intermediates.

The residue at the end of the eighth β-strand in the $(\beta/\alpha)_8$-barrel domain varies across the functionally diverse, structurally characterized members of the superfamily (*26*, *27*, *33*−*41*). The identity of this residue cannot be correlated with either the shared formation of the stabilized enediolate intermediate or divergent partitioning of the intermediate to distinct products. Instead, this residue appears to direct binding of the various carboxylate anion substrates in distinct geometries appropriate for the different reactions catalyzed by the acid/base catalysts. As a result, the identity of the reaction is determined by *both* the structure and the conformation of the carboxylate anion bound to the essential Mg^{2+}. We conclude that a mutation that allows binding of an alternative substrate in a productive geometry is sufficient for generation of a mechanistically distinct promiscuous reaction and, therefore, provides a pathway for the functional evolution.

The reactions catalyzed by members of the enolase superfamily share initial base-catalyzed abstraction of the α-proton from a carboxylate substrate but diverge in the nature of the subsequent (usually acid-catalyzed) reaction that yields the reaction product. In contrast, the members of the recently discovered, $(\beta/\alpha)_8$-barrel fold-containing, orotidine 5′-monophosphate decarboxylase (OMPDC) suprafamily utilize conserved active site residues in mechanistically *distinct* reactions; i.e., neither the structures of the substrates nor the mechanisms of the reactions are conserved (*42*). We are investigating whether single substitutions in this suprafamily can direct a change from a metal-independent reaction that avoids a vinyl anion intermediate (the OMPDC-catalyzed reaction) to a metal-dependent reaction in which an enediolate anion is stabilized (e.g., 3-keto-L-gulonate 6-phosphate decarboxylase). If so, the diversity of reactions associated with SCOP's 25 functionally distinct $(\beta/\alpha)_8$-barrel superfamilies may be explained by divergent evolution from a single progenitor.

Conclusions. Divergent evolution of function within the enolase superfamily and other superfamilies in which the active site structure "hard-wires" chemistry (*15*, *43*) likely begins with single substitutions that relax substrate specificity and generate functional (mechanistic) promiscuity. A better understanding of those structural features that determine substrate specificity will not only facilitate prediction of the functions of unknown members of mechanistically diverse superfamilies discovered in genome projects but also facilitate the design of catalysts for "unnatural" reactions.

ACKNOWLEDGMENT

We thank E. Taylor Ringia for providing [α-^{2}H]SHCHC and M. Lies for providing *cis,cis*-muconate. Figure 3 was produced using the Chimera package from the Computer Graphics Laboratory, University of California, San Francisco, CA (supported by NIH Grant P41 RR-01081). We dedicate this paper to the memory of Michael Dojka, 1969−2002.

REFERENCES

1. Wolf, Y. I., Grishin, N. V., and Koonin, E. V. (2000) *J. Mol. Biol. 299*, 897−905.
2. Wierenga, R. K. (2001) *FEBS Lett. 492*, 193−198.
3. Gerlt, J. A. (2000) *Nat. Struct. Biol. 7*, 171−173.
4. Nagano, N., Orengo, C. A., and Thornton, J. M. (2002) *J. Mol. Biol. 321*, 741−765.
5. Farber, G. K., and Petsko, G. A. (1990) *Trends Biochem. Sci. 15*, 228−234.
6. Jensen, R. A. (1976) *Annu. Rev. Microbiol. 30*, 409−425.
7. Hughes, A. L. (1994) *Proc. R. Soc. London, Ser. B 256*, 119−124.
8. Palmer, D. R., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) *Biochemistry 38*, 4252−4258.
9. O'Brien, P. J., and Herschlag, D. (1999) *Chem. Biol. 6*, R91−R105.
10. O'Brien, P. J., and Herschlag, D. (2001) *Biochemistry 40*, 5691−5699.
11. Raillard, S., Krebber, A., Chen, Y., Ness, J. E., Bermudez, E., Trinidad, R., Fullem, R., Davis, C., Welch, M., Seffernick, J., Wackett, L. P., Stemmer, W. P., and Minshull, J. (2001) *Chem. Biol. 8*, 891−898.
12. Matsumura, I., and Ellington, A. D. (2001) *J. Mol. Biol. 305*, 331−339.
13. Broun, P., Shanklin, J., Whittle, E., and Somerville, C. (1998) *Science 282*, 1315−1317.
14. Xiang, H., Luo, L., Taylor, K. L., and Dunaway-Mariano, D. (1999) *Biochemistry 38*, 7638−7652.
15. Gerlt, J. A., and Babbitt, P. C. (2001) *Annu. Rev. Biochem. 70*, 209−246.
16. 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) *Biochemistry 35*, 16489−16501.
17. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L., and Gerlt, J. A. (1995) *Science 267*, 1159−1161.
18. Schmidt, D. M., Hubbard, B. K., and Gerlt, J. A. (2001) *Biochemistry 40*, 15707−15715.
19. Vollmer, M. D., Schell, U., Seibert, V., Lakner, S., and Schlomann, M. (1999) *Appl. Microbiol. Biotechnol. 51*, 598−605.
20. Datsenko, K. A., and Wanner, B. L. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 6640−6645.
21. Stemmer, W. P. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10747−10751.
22. Furste, J. P., Pansegrau, W., Frank, R., Blocker, H., Scholz, P., Bagdasarian, M., and Lanka, E. (1986) *Gene 48*, 119−131.
23. Spencer, M. E., and Guest, J. R. (1973) *J. Bacteriol. 114*, 563−570.
24. Taylor, E. A., Palmer, D. R., and Gerlt, J. A. (2001) *J. Am. Chem. Soc. 123*, 5824−5825.
25. Cleland, W. W. (1979) *Methods Enzymol. 63*, 103−138.
26. Gulick, A. M., Schmidt, D. M., Gerlt, J. A., and Rayment, I. (2001) *Biochemistry 40*, 15716−15724.
27. Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) *Biochemistry 39*, 10662−10676.
28. Radzicka, A., and Wolfenden, R. (1995) *Science 267*, 90−93.
29. Albery, W. J., and Knowles, J. R. (1976) *Biochemistry 15*, 5631−5640.
30. Jurgens, C., Strom, A., Wegener, D., Hettwer, S., Wilmanns, M., and Sterner, R. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 9925−9930.
31. Joerger, A. C., Mayer, S., and Fersht, A. R. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100*, 5694−5699.
32. Rothman, S. C., and Kirsch, J. F. (2003) *J. Mol. Biol. 327*, 593−608.
33. Gulick, A. M., Palmer, D. R., Babbitt, P. C., Gerlt, J. A., and Rayment, I. (1998) *Biochemistry 37*, 14358−14368.
34. Helin, S., Kahn, P. C., Guha, B. L., Mallows, D. G., and Goldman, A. (1995) *J. Mol. Biol. 254*, 918−941.
35. 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) *Proc. Natl. Acad. Sci. U.S.A. 95*, 10396−10401.
36. Kleywegt, G. J., Hoier, H., and Jones, A. T. (1996) *Acta Crystallogr. D52*, 858−863.
37. Landro, J. A., Gerlt, J. A., Kozarich, J. W., Koo, C. W., Shah, V. J., Kenyon, G. L., Neidhart, D. J., Fujita, S., and Petsko, G. A. (1994) *Biochemistry 33*, 635−643.
38. Wieczorek, S. W., Kalivoda, K. A., Clifton, J. G., Ringe, D., Petsko, G. A., and Gerlt, J. A. (1999) *J. Am. Chem. Soc. 121*, 4540−4541.
39. Larsen, T. M., Wedekind, J. E., Rayment, I., and Reed, G. H. (1996) *Biochemistry 35*, 4349−4358.
40. Levy, C. W., Buckley, P. A., Sedelnikova, S., Kato, Y., Asano, Y., Rice, D. W., and Baker, P. J. (2002) *Structure 10*, 105−113.
41. Asuncion, M., Blankenfeldt, W., Barlow, J. N., Gani, D., and Naismith, J. H. (2002) *J. Biol. Chem. 277*, 8306−8311.
42. Wise, E., Yew, W. S., Babbitt, P. C., Gerlt, J. A., and Rayment, I. (2002) *Biochemistry 41*, 3861−3869.
43. Holm, L., and Sander, C. (1997) *Proteins 28*, 72−82.

BI034769A