Evolutionary Potential of $(\beta/\alpha)_8$-Barrels: In Vitro Enhancement of a "New" Reaction in the Enolase Superfamily†

Jacob E. Vick, Dawn M. Z. Schmidt, and John A. Gerlt*

Departments of Biochemistry and Chemistry, University of Illinois at Urbana−Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

Received May 24, 2005; Revised Manuscript Received July 12, 2005

ABSTRACT: The repertoire of reactions in the mechanistically diverse enolase superfamily is the result of divergent evolution that conserved enolization of a carboxylate anion substrate but allowed different overall reactions using different substrates. Details of the pathways for the natural evolutionary process are unknown, but the events reasonably involve (1) incremental increases in the level of the "new" reaction that would provide a selective advantage and (2) an accompanying loss of the "old" reaction catalyzed by the progenitor. In an effort to better understand the molecular processes of divergent evolution, the D297G mutant of the L-Ala-D/L-Glu epimerase (AEE) from *Escherichia coli* was designed so that it could bind the substrate for the *o*-succinylbenzoate synthase (OSBS) reaction and, as a result, catalyze that reaction [Schmidt, D. M. Z., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) *Biochemistry 42*, 8387−8393]. The AEE progenitor did not catalyze the OSBS reaction, but the D297G mutant catalyzed a low level of the OSBS reaction (k_{cat}, 0.013 s^{-1}; K_m, 1.8 mM; k_{cat}/K_m, 7.4 M^{-1} s^{-1}) that was sufficient to permit anaerobic growth by an OSBS-deficient strain of *E. coli*; the level of the progenitor's natural AEE reaction was significantly diminished. Using random mutagenesis and an anaerobic metabolic selection, we now have identified the I19F substitution as an additional mutation that enhances both growth of the OSBS-deficient strain and the kinetic constants for the OSBS reaction (k_{cat}, 0.031 s^{-1}; K_m, 0.34 mM; k_{cat}/K_m, 90 M^{-1} s^{-1}). Several other substitutions for Ile 19 also enhanced the level of the OSBS reaction. All of the substitutions substantially decreased the level of the AEE reaction from that possessed by the D297G progenitor. The changes in the kinetic constants for both the OSBS and AEE reactions are attributed to a readjustment of substrate specificity so that the substrate for the OSBS reaction is more productively presented to the conserved acid/base catalysts in the active site. These observations support our hypothesis that evolution of "new" functions in the enolase superfamily can occur simply by changes in specificity-determining residues.

The enolase superfamily of mechanistically diverse enzymes continues as a paradigm for understanding how homologous proteins that share a conserved structure can catalyze different reactions (*1*−*3*). Irrespective of the identity of the overall reaction, each is initated by Mg^{2+}-assisted, general base-catalyzed enolization of a carboxylate anion substrate; the stabilized enolate anion intermediate is usually directed to a specific product by a general acid-catalyzed reaction. The members of the superfamily share a two-domain structure: (1) a $(\beta/\alpha)_7\beta$-barrel (modified TIM-barrel) domain containing conserved catalytic groups separately located at the C-terminal ends of the β-strands and (2) an N-terminal $\alpha + \beta$ capping domain containing residues that determine the shape and polarity of the active site cavity and, therefore, substrate specificity. In the capping domain the specificity-determining residues are located in two loops, designated the 20s and 50s loops; the 20s loop can be disordered in the absence of an active site ligand but closes to sequester substrate and active site residues from solvent.

The process of natural divergent evolution to produce the repertoire of enzymes in the enolase superfamily likely resulted from duplication of a gene encoding a progenitor member of the superfamily; the gene-copy would then be available for evolution of a "new" function that provides a selective advantage in a new metabolic niche (*4*−*7*). Random mutational events in the duplicated gene could occur, some of which improve the ability of the organism to survive because the encoded protein catalyzes a greater level of the "new" function. Subsequent mutational events would either further enhance the catalytic activity and provide additional selective advantage or prove deleterious. Thus, evolution of the "new" function is a competition between the opposing phenotypic outcomes of random mutational events. As such, acquisition of a sufficient level of the "new" function to allow stable selective advantage is expected to occur after a limited number of mutations.

A priori, one might expect a low probability that a single mutational event in an "inactive" progenitor would be able

† This research was supported by Grant GM-52594 (to J.A.G.) from the National Institutes of Health.

* To whom correspondence should be addressed. Mailing address: Department of Biochemistry, University of Illinois, 600 S. Mathews Avenue, Urbana, IL 61801. Phone: (217) 244-7414. Fax: (217) 244-6538. E-mail: j-gerlt@uiuc.edu.

10.1021/bi050963g CCC: $30.25
Published on Web 08/09/2005

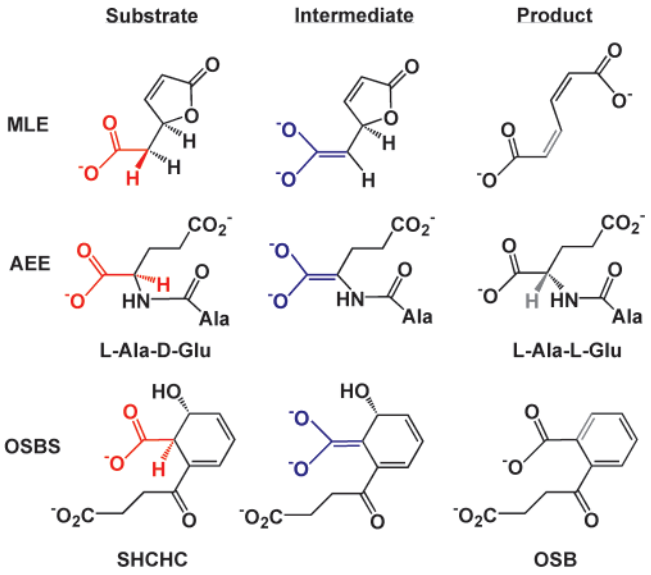

FIGURE 1: Reactions catalyzed by members of the MLE subgroup of the enolase superfamily: cycloisomerization (MLE II), 1,1-proton transfer (AEE), and dehydration (OSBS).

to provide sufficient activity for selective advantage. With this expectation, the process of divergent evolution is often envisioned to involve a progenitor that is promiscuous, i.e., it already catalyzes the "new" function at a level that would provide at least a minimal selective advantage. However, we recently reported that a single mutation in either of two inactive progenitors in the enolase superfamily was, in fact, sufficient to generate a significant level of a "new" activity (*8*).

The members of the muconate lactonizing enzyme (MLE)[1] subgroup of the enolase superfamily catalyze three different chemical reactions: cycloisomerization (MLE), 1,1-proton transfer (L-Ala-D/L-Glu epimerase, AEE), and β-elimination (*o*-succinylbenzoate synthase, OSBS) (Figure 1). The enzymes that catalyze these reactions share conserved acidic/basic Lys residues at the C-terminal ends of the second and sixth β-strands and Mg^{2+}-binding ligands at the ends of the third (Asp), fourth (Glu), and fifth (Asp) β-strands of the $(\beta/\alpha)_7\beta$-barrel domain. Despite the different reactions that are catalyzed, each is initiated by abstraction of the α-proton of a carboxylate anion substrate by the Lys residue at the end of the second β-strand to generate a stabilized enolate anion. In the MLE reaction, subsequent reversible, vinylogous elimination of the carboxylate group need not be general acid-catalyzed. In the AEE reaction, epimerization in the D- to L-direction is achieved by protonation of the intermediate by the conjugate acid of the Lys at the end of the sixth β-strand; in the L- to D-direction, the roles of the Lys residues are reversed (*9*). And, in the OSBS reaction, the conjugate acid of the Lys at the end of the second β-strand catalyzes departure of the hydroxide leaving group from the 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) substrate in the *syn*-dehydration reaction, with the Lys at the end of the sixth β-strand stabilizing the anionic intermediate (*10*, *11*). However, MLE, AEE, and OSBS contain different residues at the ends of the eighth β-strand in their barrel domains and in the 20s and 50s loops of their capping domains that form the solvent-sequestered active site cavity, thereby conferring distinct substrate specificities and, consequently, the identities of the different reactions.

OSBS catalyzes an intermediate step in the biosynthesis of menaquinone, an electron transfer cofactor required for anaerobic growth of *Escherichia coli* (*12*). The sequences of orthologous OSBSs are highly diverged, often sharing $<10\%$ pairwise identity (*16*). Therefore, with minimal global requirements for conserved residues, we hypothesized that another member of the MLE subgroup might be able to serve as the progenitor in in vitro experiments to generate an efficient OSBS with a minimal number of mutations. The sequence of mutations could mimic events in natural evolution of "new" functions in the enolase superfamily.

A designed Asp 297 to Gly (D297G) substitution in the AEE from *E. coli* and a selected Glu 323 to Gly (E323G) substitution in MLE II from *Pseudomonas sp.* P51 are each sufficient to generate a level of OSBS activity that allows growth of an OSBS-deficient strain of *E. coli* under anaerobic conditions (*8*). Neither progenitor catalyzes detectable levels of the OSBS reaction nor complements the OSBS-deficient strain. In the case of the designed D297G substitution of AEE, the rate acceleration for the OSBS reaction produced by the single substitution would be a remarkable 10^8, if the wild-type AEE progenitor for which no OSBS activity can be measured is truly unable to catalyze the reaction (*13*); for the E323G substitution of MLE II, the rate acceleration could be an even more remarkable 10^{10}. The rate acceleration associated with the reaction catalyzed by the natural OSBS from *E. coli* is 2×10^{11}.

The positions of the D297G and E323G substitutions are structurally homologous, each located at the end of the eighth β-strand of the $(\beta/\alpha)_7\beta$-barrel domain of the progenitor. In AEE, Asp 297, the second Asp in an Asp-x-Asp motif that is conserved in all orthologues, is expected to participate in an electrostatic interaction with the α-ammonium group of the dipeptide substrate based on the structure of the substrate-liganded AEE from *Bacillus subtilis* (*11*). In MLE II, the role of Glu 323 is unknown because no liganded structure is available for any MLE. But, many natural OSBSs have a Gly residue at this position, thereby contributing to an electrostatically neutral pocket in which the succinyl moiety of the SHCHC substrate binds (*14*). In the case of the designed D297G substitution in AEE, the replacement of the carboxymethyl side chain of Asp with a proton in Gly was predicted to alleviate unfavorable electrostatic and steric interactions that would exclude SHCHC from the active site. Presumably, the same reasoning would explain the effect of the selected E323G substitution in MLE II. If either of these were to be used as progenitors in natural evolution, this mutation would be required as the first step in the evolution of an OSBS.

We now are exploring whether we can identify additional substitutions that allow further selective advantage for the OSBS reaction, presumably by improving the complementarity of the active site cavity for the SHCHC substrate. We expect that additional favorable substitutions would be the result of improved substrate binding, i.e., a likely decrease in the value of the K_m, although accompanying alterations in the geometry of the bound substrate vis-à-vis the Lys

[1] Abbreviations: AEE, L-Ala-D/L-Glu epimerase; MLE, muconate lactonizing enzyme; OSB, *o*-succinylbenzoate; OSBS, *o*-succinylbenzoate synthase; SHCHC, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate.

residues at the ends of the second and sixth β-strands could accelerate dehydration of the bound substrate, i.e., an increase in the value of the k_{cat}.

The D297G mutant of the AEE was selected as the starting point for further evolution of the OSBS reaction, because the slow growth rate provided to an OSBS-deficient strain provides a large dynamic range for improvement in metabolic selection. In error-prone PCR followed by selection, we identified a second substution, I19F, resulting from a single base change, that significantly enhanced the growth rate of the OSBS-deficient strain of *E. coli*. In exploration of the phenotypes of other substitutions at this position, including all of the single-base changes that would be expected in natural divergent evolution as well as several unnatural two- and three-base changes, we identified several other substitutions for Ile 19 that also allowed enhanced growth rates. Characterization of the various purified proteins revealed that those that allowed enhanced growth rates had improved kinetic constants for the OSBS reaction as well as significantly diminished kinetic constants for the progenitor AEE reaction. The location of these substitutions in the 20s loop of the specificity-determining capping domain provides additional support for our previous conclusion that evolution of function in the enolase superfamily can occur by alteration of residues that determine substrate specificity.

MATERIALS AND METHODS

Construction of menC/ycjG::kan/cam Double Knockout Strain of E. coli. A mutant strain of *E. coli* BW25113 in which the genes encoding AEE (*ycj*G) and OSBS (*menC*) were insertionally inactivated was constructed from previously described strains that contain the component disrupted genes, *ycj*G::cam and *men*C::kan (*8*). The *men*C::kan strain was used to produce a phage P1 lysate that was used to transduce the *ycj*G::cam strain. The transduction product was identified by selection on LB plates containing 50 μg/mL kanamycin and 34 μg/mL chloramphenicol; the desired junctions were verified by DNA sequence analysis.

Construction of the Error-Prone Library. An error-prone library was generated with the "GeneMorph" kit (Stratagene, La Jolla, CA) using the gene encoding the D297G mutant of AEE cloned in the pET17b expression vector as template. A mutation rate of 1−3 bp was selected, using a 55 pg template per 50 μL reaction as specified in the manufacturer's instructions. Eleven equivalent PCR reactions (50 μL each) contained the following components: 5 μL of 10x buffer, 20 pmol of dNTPs, 20 pmol each of amplifying primers flanking the gene, 1 μL of Mutazyme DNA polymerase, and DNA template. The thermal cycler program was programmed as follows: 40 cycles of 30 s at 95 °C, 30 s at 50 °C, and 1.5 min at 72 °C, followed by 10 min at 72 °C. The products were separately digested with *Bam*HI and *Nde*I, gel-purified, combined, and concentrated to 50 μL to yield the error-prone library of PCR products. A total of 30 μL of this library was ligated in 3 μL aliquots into pDMS-1a digested with *Bam*HI and *Nde*I; the ligation mixtures were electroporated into *E. coli* XL1Blue cells and plated on a total of 39 LB/AMP plates. Approximately 110 000 colonies were obtained; these were scraped from the plates into 500 mL of LB/AMP, and the culture was grown overnight at 37 °C. Plasmids were isolated from this culture to yield the error-prone library of plasmids.

Construction of Site-Directed Mutants. Site-directed substitutions for residue 19 in the D297G mutant of AEE were constructed using the QuickChange protocol (Stratagene, La Jolla, CA); the gene was located in the pDMS1a vector previously described. The following primers were used, where NNN represents the codon for residue 19 and was varied depending on the identity of the mutant to be constructed: sense primer (5′-CCTTACATACCCCGNNNGTGTTTGCCCGGGGAAG-3′); and antisense primer (5′-GCGACTTCCCCGGGCAACACNNNCGGGGTATGTAAGGGCCA-3′). All mutations were verified by DNA sequence analysis.

Anaerobic Selection for OSBS Activity. The error-prone library was transformed into the *men*C::kan strain of *E. coli*. The transformed cells were spread on agar plates prepared from minimal medium using glycerol as carbon source, trimethylamine *N*-oxide as electron acceptor (*15*) and containing 100 μg/mL carbenicillin, 50 μg/mL kanamycin, and 1 mM IPTG. The plates were incubated at 37 °C in an anaerobic chamber; the atmosphere was 85% N_2, 10% CO_2, and 5% H_2.

Anaerobic growths of liquid cultures utilized the same minimal medium and supplements and were contained in completely filled 10 mL screw-top vials sealed with paraffin. The cultures were incubated at 37 °C; growth was monitored by the optical density at 600 nm using a Spectronic 20 spectrophotometer.

Purification of Proteins. All variants of AEE were expressed in the *men*C::kan/*ycj*G::cam strain of *E. coli* using the pDMS-1a expression vector to prevent contamination by chromosomally encoded wild-type enzymes. The purification procedure previously described was used for all variants (*8*).

Enzymatic Activity Assays. OSBS activity was determined as previously described (*16*) with the exception that the assays were performed in 50 mM sodium HEPES, pH 7.0, containing 0.1 mM $MnCl_2$, and the SHCHC substrate was neutralized and stored in 50 mM HEPES, pH 7. AEE assays were performed as previously described (*9*). The values of k_{cat} and K_M were determined using the program CLELAND (*17*). When saturation with substrate could not be achieved, the value of k_{cat}/K_M was determined by dividing the slope of the plot of velocity vs substrate by the concentration of enzyme.

RESULTS AND DISCUSSION

We have described the design of the D297G mutation of the AEE from *E. coli* so that it could catalyze the OSBS reaction (*8*). The D297G substitution was constructed to avoid unfavorable ectrostatic interactions and/or steric overlap between the carboxymethyl side chain of Asp 297 and the succinyl moiety of the SHCHC substrate for the OSBS reaction. The D297G mutant catalyzed low levels of the OSBS reaction (k_{cat}, 0.013 s^{-1}; K_m, 1.8 mM; k_{cat}/K_m, 7.4 M^{-1} sec^{-1}; Table 1). This level of activity is sufficient for slow complementation of a *men*C::kan strain that is unable to grow anaerobically because it lacks the ability to synthesize menaquinone. The mutant also catalyzes the progenitor's AEE reaction, albeit with significantly less favorable kinetic constants than the wild-type AEE progenitor (Table 1).

Although a three-dimensional structure is not available for the D297G mutant, the gain of the OSBS function associated

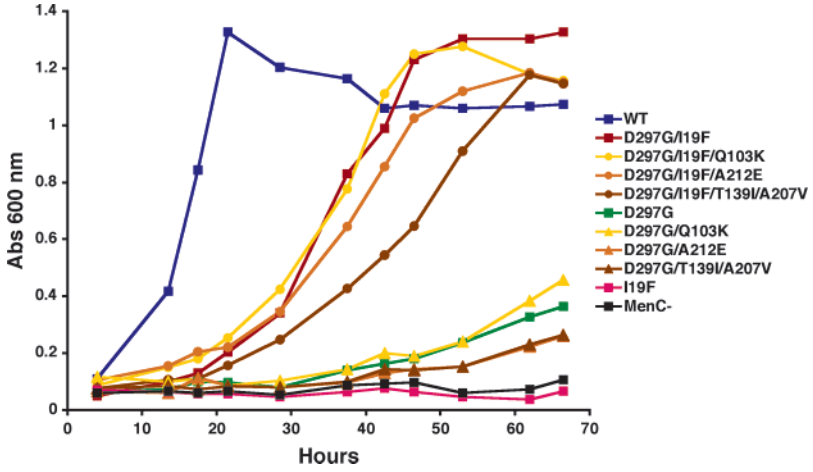

FIGURE 2: Anaerobic growth of various strains of *E. coli*: in squares, the data for WT, wild-type BW25113; MenC-, the *menC::kan* mutant of BW25113; D297G, the engineered mutant of AEE (*8*); I19F/D297G, a double mutant; and I19F, a single mutant. Data represented by circles are those for the three mutants that were identified by error-prone PCR followed by anaerobic selection; the data represented by triangles are these mutants in which the Phe substitution at residue 19 has been replaced with the wild-type Ile.

Table 1: Kinetic Parameters fr OSBS and AEE Activities

enzyme	k_{cat} (s^{-1})	K_M (M)	k_{cat}/K_M (M^{-1} s^{-1})
OSBS Activity[a]			
OSBS (*E. coli*)[b]	28	1.6×10^{-5}	1.8×10^{6}
AEE (wild-type)[c]			$<5.2 \times 10^{-3}$ [b]
D297G	$(1.3 \pm 0.6) \times 10^{-2}$	$(1.8 \pm 0.2) \times 10^{-3}$	7.4
single base pair changes			
I19F/D297G	$(3.1 \pm 0.9) \times 10^{-2}$	$(3.4 \pm 0.27) \times 10^{-4}$	90
I19T/D297G	$(2.0 \pm 0.05) \times 10^{-2}$	$(1.9 \pm 0.22) \times 10^{-4}$	110
I19S/D297G	$(6.1 \pm 0.9) \times 10^{-3}$	$(1.0 \pm 0.13) \times 10^{-4}$	60
I19N/D297G	$(1.9 \pm 0.09) \times 10^{-2}$	$(6.4 \pm 0.89) \times 10^{-4}$	29
I19V/D297G	$(9.5 \pm 0.8) \times 10^{-3}$	$(1.4 \pm 0.32) \times 10^{-3}$	7.0
I19L/D297G	$(3.2 \pm 0.2) \times 10^{-3}$	$(1.9 \pm 0.23) \times 10^{-3}$	1.7
I19M/D297G			
two base pair changes			
I19Y/D297G	0.45 ± 0.01	$(3.2 \pm 0.25) \times 10^{-3}$	140
I19C/D297G			43[d]
I19A/D297G	$(2.1 \pm 0.07) \times 10^{-2}$	$(8.9 \pm 0.80) \times 10^{-4}$	23
I19H/D297G	$(1.1 \pm 0.07) \times 10^{-2}$	$(1.1 \pm 0.20) \times 10^{-3}$	10.5
three base pair changes			
I19W/D297G	0.25 ± 0.02	$(7.4 \pm 1.5) \times 10^{-4}$	340
AEE Activity[a]			
AEE (*E. coli*)[c]	10 ± 0.4	$(1.3 \pm 0.3) \times 10^{-4}$	7.7×10^{4}
D297G[c]	$(4.3 \pm 0.05) \times 10^{-2}$	$(4.4 \pm 0.3) \times 10^{-3}$	9.8
D297G/I19F			9.7×10^{-2} [d]

[a] Assay conditions described in Materials and Methods. [b] From ref *13*. [c] From ref *8*. [d] Unable to saturate due to assay concentration constraints.

with the D297G substitution must be the result of access of the SHCHC substrate for the OSBS reaction to the divalent metal ion and active site Lys residues that are appropriately positioned to catalyze not only the 1,1-proton transfer reaction of the AEE progenitor but also the *syn*-dehydration in the OSBS reaction. Because the specificity determinants in AEE should be optimized for the AEE reaction, not the OSBS reaction, we hypothesized that the kinetic constants for the OSBS reaction catalyzed by the D297G mutant might be improved by additional substitutions.

Although we expected that substitutions in the 20s and/or 50s loops in the capping domain would enhance binding of the substrate for the OSBS reaction, the only structure for the AEE from *E. coli* is that of the unliganded wild-type protein in which the 20s loop is disordered. Thus, we turned to error-prone PCR coupled with selection to search for a second substitution that enhances the OSBS activity of the designed D297G substitution.

Selection and Identification of Mutants from the Error-Prone Library. An error-prone library of the gene encoding the D297G mutation of AEE was generated using conditions that were expected to introduce between one and three base changes in the gene. After transformation into a strain of *E. coli* that cannot grow anaerobically because of a disruption in the *men*C gene encoding OSBS and inspection of colony sizes, several colonies were identified that allowed more rapid growth than that supported by the template that encoded only the D297G substitution. The AEE-derived genes in the plasmids were sequenced, and three unique sequences were obtained, all of which contained the D297G mutation of the template: I19F/Q103L/D297G (two base changes relative to the D297G template), I19F/A212E/D297G (two base changes), and I19F/T139I/A207V/D297G (three base changes); no silent base changes accompanied these mutations. Cells transformed with these plasmids were grown anaerobically in liquid medium; in Figure 2 the growth curves of these (yellow, orange, and brown circles) are compared with those of a wild-type strain of *E. coli* (blue squares), the OSBS-deficient *men*C::kan strain (black squares), and the OSBS-deficient strain transformed with the D297G template (green squares). All three of the multiple mutants generated by error-prone PCR allowed faster growth rates than the D297G single mutant.

All three mutants contained the I19F substitution, a single base change (ATT → TTT). The following anaerobic growth experiments established the requirement for *both* the I19F and D297G substitutions for the more rapid growth:

(1) The I19F mutation was removed from each of the three multiple mutants by site-directed mutagenesis; the growth rates were reduced such that they approximated the D297G mutant (Figure 2; yellow, orange, and brown triangles).

(2) The I19F substitution in the absence of D297G was unable to support anaerobic growth (Figure 2; fuschia squares).

(3) The I19F and D297G substitutions together were sufficient to reproduce the rapid growth rates observed for the multiple mutants (Figure 2; red squares).

As a result, we hypothesized that the I19F substitution alone is responsible for the improved kinetic constants for the OSBS reaction relative to those previously reported for the D297G substitution.

Effect of I19F on OSBS Activity. The I19F/D297G double mutant was isolated from the *men*C::kan/*ycj*G::cam strain to prevent any contamination with wild-type enzymatic activities and assayed for both OSBS and AEE activity. The values of the kinetic constants are compared with those obtained for wild-type and the D297G mutant in Table 1. The increased rate of anaerobic growth is associated with "improved" kinetic constants for the OSBS reaction: a 2.3-fold increase in k_{cat}, a 5.3-fold decrease in K_m, and a 12.2-fold increase in k_{cat}/K_m compared to the D297G mutant.

Is I19F a Unique Solution to Improved OSBS Activity? The I19F substitution is the result of a single base change (ATT → TTT). Although our error-prone library should have been large enough to detect other substitutions for Ile 19 resulting from single base changes that also would have allowed more rapid growth than D297G under anaerobic conditions, we constructed other substitutions for Ile 19 in the D297G mutant. These include the six remaining one-base mutants [Asn (AAT), Leu (CTT), Met (ATG), Ser (AGT), Thr (ACT), and Val (GTT)] that would be possible from mutational events in natural evolution. We also constructed four two-base mutants [Ala (GCT), Cys (TGT), His (CAT), and Tyr (TAT)] and one three-base mutant [Trp (TGG)] to explore the effect of hydrophobic substitutions for Ile 19; these mutants would not be expected in our error-prone library.

The mutants were transformed into the OSBS-deficient strain and subjected to anaerobic selection. The growth curves that were obtained from the selected I19F substitution as well as the constructed six one-base mutants are displayed in Figure 3A, along with a wild-type strain (blue squares), the OSBS-deficient *men*C::kan strain (black squares), the OSBS-deficient strain transformed with the D297G template (green squares), and the OSBS-deficient strain transformed with the D297G/I19F mutant (red squares). Unexpectedly, given the polar character of these side chains, both I19S and I19T significantly increased the growth rate; the hydrophobic I19V substitution modestly increased the growth rate. However, the effects of the three constructed substitutions were less than that produced by the selected I19F substitution. The I19N substitution had no distinguishable effect on growth rate relative to the D297G progenitor; the I19L and I19M substitutions were deleterious. Thus, four of the seven possible substitutions for Ile 19 resulting from one-base change, Phe, Ser, Thr, and Val, provide selective advantage to the D297G substitution.

Given the effects of the Ser and Thr substitutions, the phenotypes of the I19A and I19C substitutions that involve two-base changes were of interest. Both substitutions allowed growth rates comparable to those observed for Ser and Thr (Figure 3B). These likely would be inaccessible by natural evolution.

The I19H and I19Y two-base change substitutions and the I19W three-base change substitution were constructed to fully interrogate the effects of other aromatic side chains at this position. Remarkably, I19W resulted in the most significant increase in growth rate observed for any mutant, including the selected I19F (Figure 3B). The I19Y substitution had a similar phenotype to I19F; the I19H substitution provided no increase in growth rate.

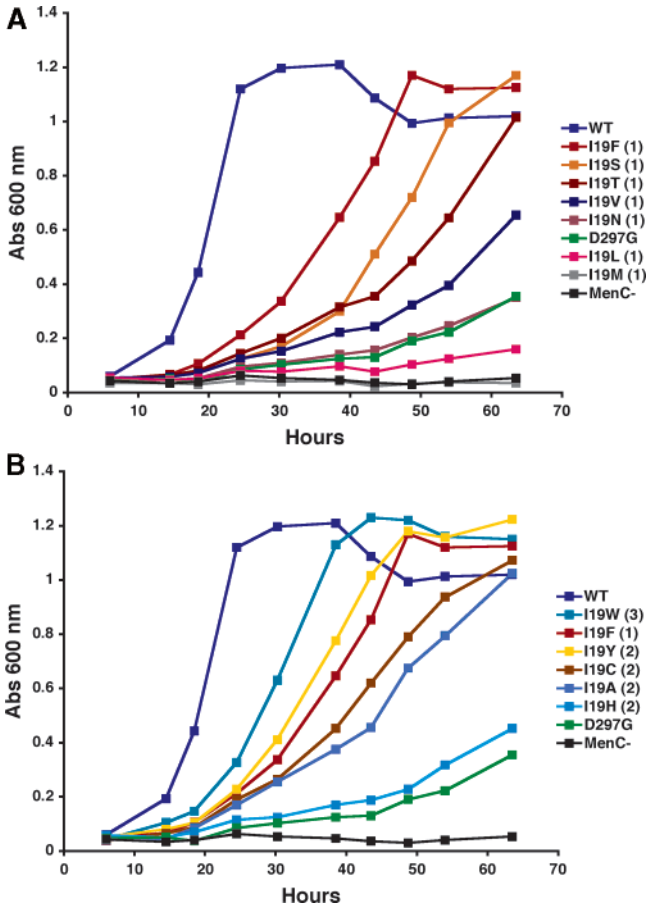

FIGURE 3: Anaerobic growth of various strains of *E. coli*: WT, wild-type BW25113; MenC-, the *menC::kan* mutant of BW25113; and D297G, the engineered mutant of AEE (*8*). (A) The double mutants that contain D297G as well as one-base changes for Ile 19. (B) The double mutants that contain D297G as well as two- and three-base changes for Ile 19; the numbers in parentheses designate the number of base pair changes.

In summary, eight of the twelve substitutions for Ile 19 increased the growth rate relative to that for the D297G progenitor.

Effect of Constructed Mutants on OSBS Activity. The various constructed mutants were isolated from the *men*C::kan/*ycj*G::cam strain and assayed for both OSBS and AEE activity; the values of the kinetic constants are displayed in Table 1. [The I19M mutant could not be isolated as a homogeneous sample using the procedures used for the mutants in Table 1; recall that the complementation phenotype of this substitution was indistinguishable from that of the *men*C::kan strain.] In five cases where the growth rate is increased, the value for the K_m is decreased at least 2-fold from that measured for the D297G progenitor, and the value of k_{cat} is only modestly changed (either decreased or increased by a factor of less than 2- to 3-fold); the resulting values of k_{cat}/K_m are increased from 3- to 12-fold (I19F, I19T, I19S, I19N, and I19A). I19F, I19T, I19S, and I19N are accessible by natural mutational events (one-base change).

In contrast, the increased growth rates of two other mutants can be associated with significant increases in k_{cat}. The value of the K_m for I19Y is increased 2-fold (from 1.8 mM to 3.4 mM), but the value of k_{cat} is increased 35-fold (from 0.013

s^{-1} to 0.45 s^{-1}); the value of k_{cat}/K_m is increased 19-fold. The value of the K_m for I19W is decreased (from 1.8 mM to 0.7 mM), but the value of k_{cat} is increased 20-fold (from 0.013 s^{-1} to 0.25 s^{-1}); the value of k_{cat}/K_m is increased 45-fold. Although these substitutions are not accessible by natural mutational events (one-base change), their effects on catalysis are impressive.

We conclude that increased growth rates are associated with improvements in catalysis, thereby establishing the utility of the selection procedure to identify either alternative substitutions at other locations or additional substitutions at other locations that improve the kinetic constants for the OSBS reaction.

Structural Basis for Enhancement of OSBS Activity. A crystal structure is not yet available for substrate-liganded wild-type AEE, the designed D297G mutant, or any of the double mutants described in this manuscript. Nevertheless, the acquisition of the "new" OSBS function by the D297G substitution can be explained by "new" access of the SHCHC substrate to the active site (*8*). If properly oriented, the SHCHC molecule could coordinate to the essential divalent metal ion, its α-proton could be presented to Lys 151 at the end of the second β-strand for abstraction, and subsequent vinylogous β-elimination of hydroxide could be catalyzed by the conjugate acid of the same Lys residue.

How can the enhanced activities associated with substitutions for Ile 19 be explained? Ile 19 is located in the 20s loop of the capping domain. The 20s loops of the AEE's from *E. coli* and *B. subtilis* share the same number of amino acid residues (between the bold, black residues in Figure 4A), although the sequences are not strictly conserved. In the structure of the AEE from *B. subtilis* (*11*) (Figure 4B), Arg 24 (highlighted in magenta in Figure 4A) forms an ionic interaction with the carboxylate of the Glu moiety of the substrate (Figure 4B). In the functional orthologues of the AEE from *E. coli*, Arg 21 (highlighted in green in Figure 4A) is strictly conserved, so we expect that it participates in an ionic interaction with the carboxylate group of the Asp/Glu moiety of the substrate. The Arg residues do not occupy homologous positions in the sequences of the loops, so the precise structural strategies by which the AEE's determine their substrate specificities are not conserved. Nonetheless, we expect that the structures of the 20s loops, including the spatial orientations of the side chains, will be similar. Ile 19 in the AEE from *E. coli* and Thr 21 in the AEE from *B. subtlis* (highlighted in red in Figure 4A) occupy the same sequence position with the 20s loops. The γ-methyl group of Thr 21 residue points into the active site cavity (Figure 4B); therefore, we predict that Ile 19 will form part of the surface of the active site cavity in the AEE from *E. coli*. Accordingly, substitutions for Ile 19 are predicted to alter the shape of the active site cavity, with those substitutions that enhance catalytic activity allowing a more productive binding of the SHCHC substrate for the OSBS reaction.

Thus, we expect that the I19F substitution alters the shape of the active site cavity and, perhaps, provides favorable $\pi-\pi$ stacking interactions that increase the affinity of the cavity for the SHCHC substrate. This interpretation also likely applies to the improved substrate affinity displayed by the I19W and I19Y substitutions, assuming that the value of K_m is a measure of substrate binding.

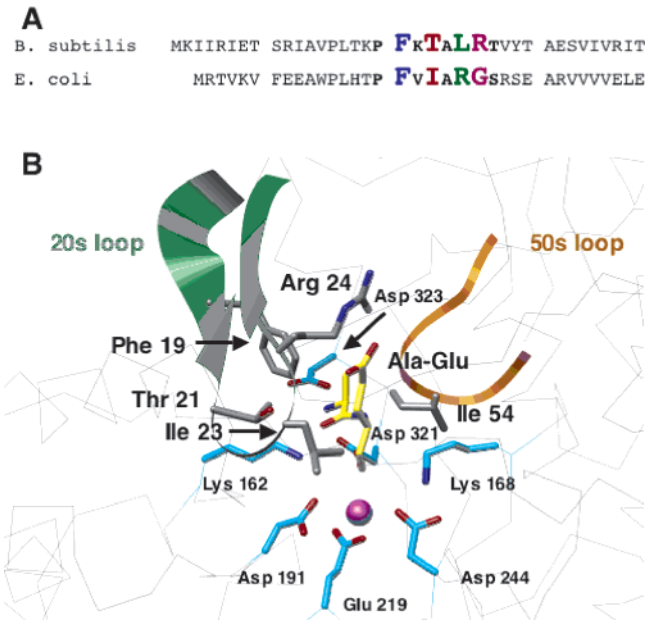

FIGURE 4: (A) The sequences of the 20s loops in the AEE's from *E. coli* and *B. subtilis*. (B) The active site of the AEE from *B. subtilis* highlighting the positions of the residues in the 20s and 50s loop relative to the Ala-Glu substrate (shaded in yellow). The side chains of Phe 19, Thr 21, Ile 23, Arg 24, and Ile 54 (shaded in gray) are pointed into the active site cavity as partial determinants of shape and polarity. The acidic/basic Lys residues at the ends of the second (Lys 162) and sixth (Lys 268) β-strands, the Mg^{2+}-binding ligands at ends of the third (Asp 191), fourth (Glu 219), and fifth (Asp 244) β-strands, and the α-ammonium group binding residues at the end of the eighth β-strand (Asp 321 and 323) of the $(\beta/\alpha)_8$-barrel domain are shaded in cyan.

Curiously, the I19T and I19S substitutions produced the highest affinity for the SHCHC substrate, suggesting the importance of the β-hydroxyl group in substrate binding. Perhaps these smaller side chains allow a reorientation of the spatially proximal side chains in the 20s and 50s loops that form the active site cavity, including Phe 17 and Arg 21 located on the 20s loop and Tyr 48 and Arg 50 located on the 50s loop, thereby increasing the affinity for the substrate.

Effect of Substitutions on AEE Activity. We previously reported that the D297G substitution causes a nearly 10^4-fold reduction in the value of k_{cat}/K_m for the AEE reaction (Table 1), with a significant decrease in k_{cat} and an increase in K_m contributing to this composite effect. Based on the substrate-liganded structure of the AEE from *B. subtilis*, Asp 297 is expected to participate in an electrostatic interaction with the α-ammonium group of the substrate, so the absence of a carboxylate positioned at the end of the eighth β-strand is expected to significantly decrease the affinity for the substrate for the AEE reaction.

The selected I19F substitution and the constructed substitutions for Ile 19 that improved the OSBS activity were also assayed for AEE activity (Table 1). The I19F substitution decreased k_{cat}/K_m for the AEE reaction by an additional 10^2-fold, although the exact effects on k_{cat} and K_m cannot be quantitated because saturation with the L-Ala-D-Glu substrate could not be achieved. Evidently, the bulkier phenyl side chain of Phe 19 interferes with the binding of the L-Ala-D-Glu substrate for the AEE reaction as it improves the binding of the SHCHC substrate for the OSBS reaction. We were

unable to quantitate the kinetic constants for AEE reactions catalyzed by the remaining substitutions because these did not produce rates in our coupled assay that could be distinguished from background. Thus, these substitutions had even larger deleterious effects on the progenitor's AEE reaction.

Implications for Divergent Evolution. We, and others, have speculated on the importance of functional promiscuity in the evolution of "new" functions (*16*, *18*−*23*). The ability of the progenitor to catalyze even a low level of a "new" activity is expected to facilitate the evolution by providing an immediate, recognizable selective advantage. In one potential natural strategy, duplication of the gene encoding the promiscuous enzyme would allow optimization of the promiscuous activity associated with the copy, presumably with loss of the natural function of the progenitor, while the original function is "protected" by the original gene.

However, Tawfik and others have proposed that gene duplication need not be the initial event in the pathway for optimization of the promiscuous activity (*7*, *24*, *25*). Instead, mutational events in the original gene prior to gene duplication might allow the encoded protein to catalyze both the original reaction and enhanced levels of the new function that would provide selective advantage. This pathway would require that the original function be immune to those mutations that enhance the promiscuous function required for selective advantage.

In support of this proposal, Tawfik et al. recently described the directed evolution of a promiscuous esterase activity possessed by human carbonic anhydrase II (hCAII) (*26*). The hydration of carbon dioxide and the hydrolysis of acyl esters occur by similar mechanisms, i.e., attack of Zn^{2+}-coordinated/activated water on a "carbonyl" group of the substrate to form an anionic species that is stabilized by coordination to the same active site Zn^{2+}. The hydration of carbon dioxide occurs at the diffusion-controlled limit, $k_{cat}/K_m = 2.8 \times 10^7$ M^{-1} s^{-1}. The hydrolysis of aryl esters is slower: the value of the k_{cat}/K_m for the activated 4-nitrophenyl acetate (NPA) is 2050 M^{-1} s^{-1}; the value for the less activated and sterically more demanding 2-naphthyl acetate (NA) is 25 M^{-1} s^{-1}. Tawfik and co-workers used directed evolution to identify A65V and T200A substitutions that preferentially enhance the hydrolysis of NA (k_{cat}/K_m for NPA = 8600 M^{-1} s^{-1}; k_{cat}/K_m for NA = 620 M^{-1} s^{-1}) by making the active site more accessible to the larger NA. However, these substitutions had only a 2-fold effect on the ability of the enzyme to hydrate the smaller carbon dioxide substrate for the natural hydration reaction ($k_{cat}/K_m = 1.3 \times 10^7$ M^{-1} s^{-1}).

Whether these experiments provide general support for the proposal that a native function is necessarily robust and immune to mutational events that enhance promiscuity is a matter of debate. Given the mechanistic similarity between the one-step hydration of carbonic anhydrase and hydrolysis of an aryl ester that involves nucleophlic attack followed by leaving group departure, the retention of substantial activity for the "simpler" processing of the smaller substrate is not surprising.

In the case of evolution of the OSBS function from the AEE function in the enolase superfamily, both are initiated by a conserved enolization reaction, but neither the structure of the substrate nor the identity of the general acid-catalyzed reaction that accomplishes product formation is conserved. Also, in the example we are studying, the AEE progenitor does not catalyze a detectable level of the OSBS reaction. Nevertheless, with the single D297G substitution, the mutant is capable of catalyzing a sufficient level of the OSBS reaction to allow growth of an OSBS-deficient strain. If the AEE progenitor truly has *no* ability to catalyze the OSBS reaction, the gain of function achieved with the D297G substitution is an impressive 10^8 rate acceleration. As reported previously, this gain of OSBS function is accompanied by a nearly 10^4-fold decrease in k_{cat}/K_m for the AEE reaction (a decrease from 7.7×10^4 M^{-1} s^{-1} to 9.7 M^{-1} s^{-1}). Now, with newly identified second substitutions at residue 19, the values of k_{cat}/K_m for the AEE reaction decrease by additional factors of $\geq 10^2$ (to 0.097 M^{-1} s^{-1} for I19F and even more for the remaining substitutions). So, in our studies the natural function is *not* robust and is quickly lost as the new function is generated (D297G) and enhanced (D297G plus several substitutions for Ile 19).

Based on our observations we conclude that the susceptibility of the natural reaction to mutational events that enhance functional promiscuity depends on the requirements for both the natural and promiscuous reactions, including the sizes and structural complexities of the substrates.

CONCLUSIONS

Our design of the D297G mutant of AEE that catalyzes the OSBS reaction is persuasive evidence that the active sites of members of the MLE subgroup of the enolase superfamily are "hard-wired" for acid/base chemistry, with the identity of the catalyzed reaction determined by the specificity determinants (*8*). In support of that view, we now have identified substitutions for Ile 19 that both provide a growth advantage under selective conditions and improve the values of the kinetic constants for the OSBS reaction. Presumably, these improvements result from additional changes in substrate specificity that further enhance binding of the substrate for the OSBS reaction. We conclude that divergent evolution of a "new" function by in vitro approaches is possible with a surprisingly limited number of substitutions. In our conversion of the AEE from *E. coli* to an OSBS, the progenitor is not functionally promiscuous, and the reaction catalyzed by the progenitor is rapidly lost as the new function evolves. These observations presumably are relevant to the process of natural divergent evolution of "new" functions in the enolase superfamily.

ACKNOWLEDGMENT

We thank John Rakus for his expert assistance in the construction of the site-directed substitutions for Ile 19.

REFERENCES

1. 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 alpha-protons of carboxylic acids, *Biochemistry 35*, 16489−501.
2. Gerlt, J. A., and Babbitt, P. C. (2001) Divergent Evolution Of Enzymatic Function: Mechanistically Diverse Superfamilies and Functionally Distinct Suprafamilies, *Annu. Rev. Biochem. 70*, 209−46.
3. Gerlt, J. A., and Raushel, F. M. (2003) Evolution of function in (beta/alpha)(8)-barrel enzymes, *Curr. Opin. Chem. Biol. 7*, 252−64.

4. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function, *Annu. Rev. Microbiol. 30*, 409−25.
5. Ohno, S. (1970) *Evolution by Gene Duplication*, Springer-Verlag, New York.
6. Hughes, A. L. (1994) The evolution of functionally novel proteins after gene duplication, *Proc. R. Soc. London, Ser. B 256*, 119−24.
7. Lynch, M., and Katju, V. (2004) The altered evolutionary trajectories of gene duplicates, *Trends Genet. 20*, 544−9.
8. 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*, 8387−93.
9. 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*, 15707−15.
10. Klenchin, V. A., Taylor Ringia, E. A., Gerlt, J. A., and Rayment, I. (2003) Evolution of enzymatic activity in the enolase superfamily: structural and mutagenic studies of the mechanism of the reaction catalyzed by o-succinylbenzoate synthase from *Escherichia coli*, *Biochemistry 42*, 14427−33.
11. Klenchin, V. A., Schmidt, D. M., Gerlt, J. A., and Rayment, I. (2004) Evolution of enzymatic activities in the enolase superfamily: structure of a substrate-liganded complex of the L-Ala-D/L-Glu epimerase from *Bacillus subtilis*, *Biochemistry 43*, 10370−8.
12. Meganathan, R. (1996) in *Escherichia coli and Salmonella* (Neidhart, F. C., Curtiss, R., Ingraham, J. L., LIn, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M., and Umbarger, H. E., Eds.) pp 642−56, ASM Press, Washington, DC.
13. Taylor, E. A., Palmer, D. R., and Gerlt, J. A. (2001) The Lesser "Burden Borne" by o-Succinylbenzoate Synthase: An "Easy" Reaction Involving a Carboxylate Carbon Acid, *J. Am. Chem. Soc. 123*, 5824−5.
14. 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(II) and o-succinylbenzoate, *Biochemistry 39*, 10662−76.
15. Spencer, M. E., and Guest, J. R. (1973) Isolation and properties of fumarate reductase mutants of *Escherichia coli*, *J. Bacteriol. 114*, 563−70.
16. Palmer, D. R., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) Unexpected divergence of enzyme function and sequence: "N-acylamino acid racemase" is o-succinylbenzoate synthase, *Biochemistry 38*, 4252−8.
17. Cleland, W. W. (1979) Statistical analysis of enzyme kinetic data, *Methods Enzymol. 63*, 103−38.
18. O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities, *Chem. Biol. 6*, R91−R105.
19. James, L. C., and Tawfik, D. S. (2001) Catalytic and binding polyreactivities shared by two unrelated proteins: The potential role of promiscuity in enzyme evolution, *Protein Sci. 10*, 2600−7.
20. Copley, S. D. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity, *Curr. Opin. Chem. Biol. 7*, 265−72.
21. Kazlauskas, R. J. (2005) Enhancing catalytic promiscuity for biocatalysis, *Curr. Opin. Chem. Biol. 9*, 195−201.
22. Bornscheuer, U. T., and Kazlauskas, R. J. (2004) Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways, *Angew. Chem., Int. Ed. 43*, 6032−40.
23. Yew, W. S., Akana, J., Wise, E. L., Rayment, I., and Gerlt, J. A. (2005) Evolution of enzymatic activities in the orotidine 5′-monophosphate decarboxylase suprafamily: enhancing the promiscuous D-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase, *Biochemistry 44*, 1807−15.
24. Aharoni, A., Gaidukov, L., Khersonsky, O., McQ Gould, S., Roodveldt, C., and Tawfik, D. S. (2005) The 'evolvability' of promiscuous protein functions, *Nat. Genet. 37*, 73−6.
25. Kondrashov, F. A. (2005) In search of the limits of evolution, *Nat. Genet. 37*, 9−10.
26. Gould, S. M., and Tawfik, D. S. (2005) Directed evolution of the promiscuous esterase activity of carbonic anhydrase II, *Biochemistry 44*, 5444−52.

BI050963G