BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Role of an Active Site Loop in the Promiscuous Activities of *Amycolatopsis* sp. T-1-60 NSAR/OSBS

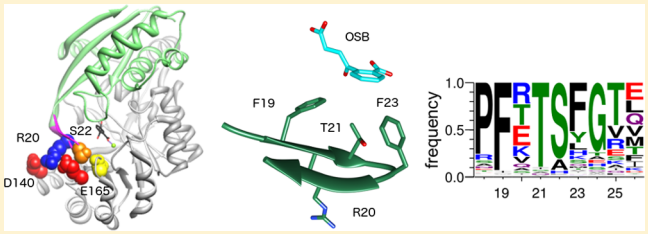
Andrew W. McMillan,[†] Mariana S. Lopez,[†] Mingzhao Zhu,[‡] Benjamin C. Morse,[†] In-Cheol Yeo,[†] Jaleesia Amos,[†] Ken Hull,[‡] Daniel Romo,[‡,§] and Margaret E. Glasner*[,†]

[†]Department of Biochemistry and Biophysics, Texas A&M University, 2128 TAMU, College Station, Texas 77843-2128, United States

[‡]Natural Products LINCHPIN Laboratory and [§]Department of Chemistry, Texas A&M University, 3255 TAMU, College Station, Texas 77840, United States

Supporting Information

ABSTRACT: The *o*-succinylbenzoate synthase (OSBS) family is part of the functionally diverse enolase superfamily. Many proteins in one branch of the OSBS family catalyze both OSBS and *N*-succinylamino acid racemization in the same active site. In some promiscuous NSAR/OSBS enzymes, NSAR activity is biologically significant in addition to or instead of OSBS activity. Identifying important residues for each reaction could provide insight into how proteins evolve new functions. We have made a series of mutations in *Amycolatopsis* sp. T-1-60 NSAR/OSBS in an active site loop, referred to as the 20s loop. This loop affects substrate specificity in many members of the enolase superfamily but is poorly conserved within the OSBS family. Deletion of this loop decreased OSBS and NSAR catalytic efficiency by 4500-fold and 25,000-fold, respectively, showing that it is essential. Most point mutations had small effects, changing the efficiency of both NSAR and OSBS activities <10-fold compared to that of the wild type. An exception was F19A, which reduced k_{cat}/K_M^{OSBS} 200-fold and k_{cat}/K_M^{NSAR} 120-fold. Mutating the surface residue R20E, which can form a salt bridge to help close the 20s loop over the active site, had a more modest effect, decreasing k_{cat}/K_M of OSBS and NSAR reactions 32- and 8-fold, respectively. Several mutations increased K_M of the NSAR reaction more than that of the OSBS reaction. Thus, both activities require the 20s loop, but differences in how mutations affect OSBS and NSAR activities suggest that some substitutions in this loop made a small contribution to the evolution of NSAR activity, although additional mutations were probably required.



Many enzymes have flexible loops near the active site which play important roles in their functions. Upon binding a substrate, many active site loops become more ordered or undergo conformational changes to form a cap over the active site.[1−11] These changes can allow the loop to directly interact with substrates[12−15] or isolate the active site to alter the chemical environment by changing the dielectric constant.[16−18] This can lead to stronger binding between charges[16,19,20] or altered pK_a values within the active site.[2,21] Because such loops contain less regular structure than α-helices or β-strands, mutations in conformationally dynamic active site loops are expected to be less detrimental than in rigidly structured regions, providing a route to evolve new functions both in nature and by protein engineering.[22−24]

We are investigating the possible role of a flexible active site loop in the evolution of new functions in the *o*-succinylbenzoate synthase (OSBS) family. All experimentally characterized members of this family are able to catalyze a dehydration reaction that forms *o*-succinylbenzoate (OSB), a required step in menaquinone biosynthesis (Figure 1A).[25−29] In addition, some members of one subfamily catalyze both OSBS and the racemization of *N*-succinylamino acids using the same catalytic residues in the active site (Figure 1B).[26,27,30] In one species, *Geobacillus kaustophilus*, *N*-succinylamino acid racemase (NSAR) activity is used in a pathway to convert D-amino acids to L-amino acids.[27] Its NSAR/OSBS is bifunctional, because it also requires OSBS activity for menaquinone synthesis. In many other species that encode NSAR/OSBS subfamily enzymes, genome context indicates that NSAR (or another) activity is the protein's biological function, because menaquinone biosynthesis genes are missing.[31] In this article, we analyzed the role of an active site loop in the NSAR/OSBS from *Amycolatopsis* sp. T-1-60, which catalyzes OSBS and racemization of *N*-succinyl-L-phenylglycine with similar efficiencies.[30] Although the genome of this species has not been sequenced, NSAR is probably the biological function of this enzyme, based on the genome context of a related species, *Amycolatopsis azurea*. *A. azurea*, which is an Actinobacteria, encodes the menaquinone synthesis pathway.[32] This organism not only encodes an OSBS from the Actinobacteria OSBS subfamily, but it also encodes an NSAR/OSBS that shares 90%

Received: May 14, 2014
Revised: June 19, 2014
Published: June 23, 2014

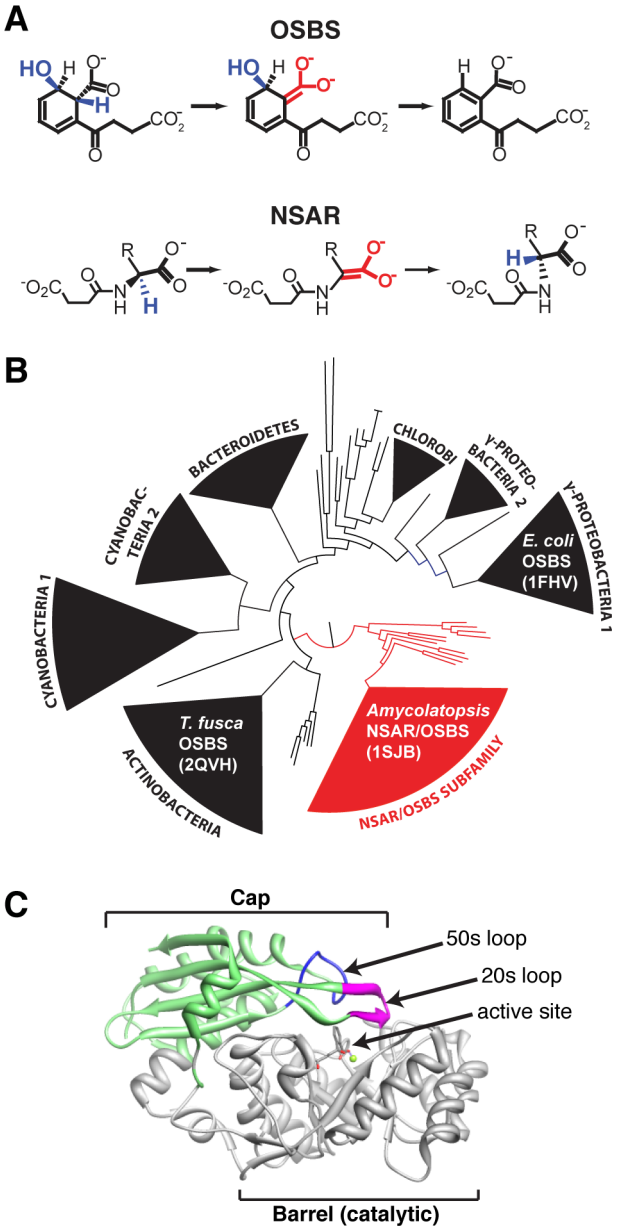


Figure 1. (A) *o*-Succinylbenzoate synthase (OSBS) and *N*-succinylamino acid racemase (NSAR) reactions. Structural similarities of the intermediates are red; blue atoms are lost or rearranged in the reactions. R = hydrophobic amino acid side chain. (B) Phylogenetic tree of the OSBS family.[38] Wedges represent clusters of sequences that share at least 40% sequence identity. Species and PDB codes of crystal structures discussed in the text are shown in the wedges. The NSAR/OSBS subfamily is colored red. (C) Structure of *Amycolatopsis* NSAR/OSBS (PDB: 1SJB Chain A) showing the location of the 20s loop (magenta).[33] The 50s loop is blue, the rest of the capping domain is green, and the barrel domain is gray. OSB and a magnesium ion are shown in the active site. Structural representations shown in all figures were prepared using UCSF Chimera.[62]

amino acid sequence identity with the *Amycolatopsis* sp. T-1-60 NSAR/OSBS.

The OSBS family is part of the enolase superfamily. Like other members of the enolase superfamily, the active site is at the interface between a capping domain and a $(\beta/\alpha)_7\beta$-barrel domain (Figure 1C). The barrel domain contains the catalytic residues that are conserved in all enolase superfamily members. These consist of three acidic residues that bind a magnesium ion needed to stabilize the enolate anion intermediate (D189, E214, and D239 of *Amycolatopsis* NSAR/OSBS), a lysine which serves as the catalytic base for the OSBS reaction and either a base for D-to-L racemization or acid for L-to-D racemization (K163 in *Amycolatopsis* NSAR/OSBS), and a second lysine (K263 in *Amycolatopsis* NSAR/OSBS) that stabilizes the transition state of the OSBS reaction or serves as the second acid or base for racemization.[30,33]

The capping domain includes two loops that close over the active site, known as the 20s loop and 50s loop, because of the approximate position of these loops in the protein sequences of most enolase superfamily members. The 50s loop usually has a well-defined structure rather than being flexible, and it forms part of the interface in homoligomers of the NSAR/OSBS subfamily and many other enolase superfamily members.[29,34−36]

In contrast, the 20s loop is often disordered in crystal structures of enolase superfamily enzymes in the absence of bound substrate or product analogues.[29,34−37] The position of the 20s loop also differs among members of the OSBS family, even when OSB is bound (Figure 2).[28,33,37] Compared to the 20s loop of OSBS enzymes from other subfamilies, the 20s loop of *Amycolatopsis* NSAR/OSBS extends further over the active site, allowing more residues to interact with the substrate or with the barrel domain. The two other OSBS enzymes whose OSB-bound structures have been determined are from different subfamilies that lack NSAR activity. In the *Thermobifida fusca* OSBS from the Actinobacteria subfamily, the 20s loop contains a small deletion, and in the *Escherichia coli* OSBS from the γ-Proteobacteria 1 subfamily, the 20s loop is somewhat twisted compared to the NSAR/OSBS loop.[28,37] In both cases, the 20s loop does not contact the barrel domain as it does in the NSAR/OSBS subfamily. Instead, in both *E. coli* and *T. fusca* OSB synthases, an arginine from the 20s loop reaches across the active site to interact with the 50s loop, although the position of this arginine is not conserved in the two subfamilies.[28,38] These results suggest that the 20s loop could help determine the difference in specificity between the promiscuous NSAR/OSBS subfamily and other OSBS subfamilies.

The 20s loop has long been proposed to play a role in determining substrate specificity in the enolase superfamily, with much of the evidence coming from studies of the dipeptide epimerase (DE) family, another member of the enolase superfamily.[39] The specificity of *E. coli* L-Ala-D/L-Glu epimerase could be changed by mutations in the 20s loop combined with another mutation in the barrel domain.[14,40] As further evidence that this loop determines specificity, Lukk et al.[13] showed that hydrophobic contacts, hydrogen bonds, or salt bridges between the 20s loop and the ligand determine substrate specificity for different dipeptides. However, Bourque and Bearne found that the 20s loop of mandelate racemase can tolerate a wide variety of hydrophobic substitutions, suggesting that these contacts are mostly nonspecific.[15] Likewise, point mutations that disrupt the interaction between the ligand and the 20s loop in *E. coli* and *T. fusca* OSB synthases had only a small effect.[28,38] These results indicate that the role of the 20s loop varies in the enolase superfamily.

In this work, we examined the role of the 20s loop in *Amycolatopsis* NSAR/OSBS by measuring the effects of a series of mutations on OSBS and NSAR activities. We discovered that *Amycolatopsis* NSAR/OSBS is much less tolerant of mutations

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434−4444

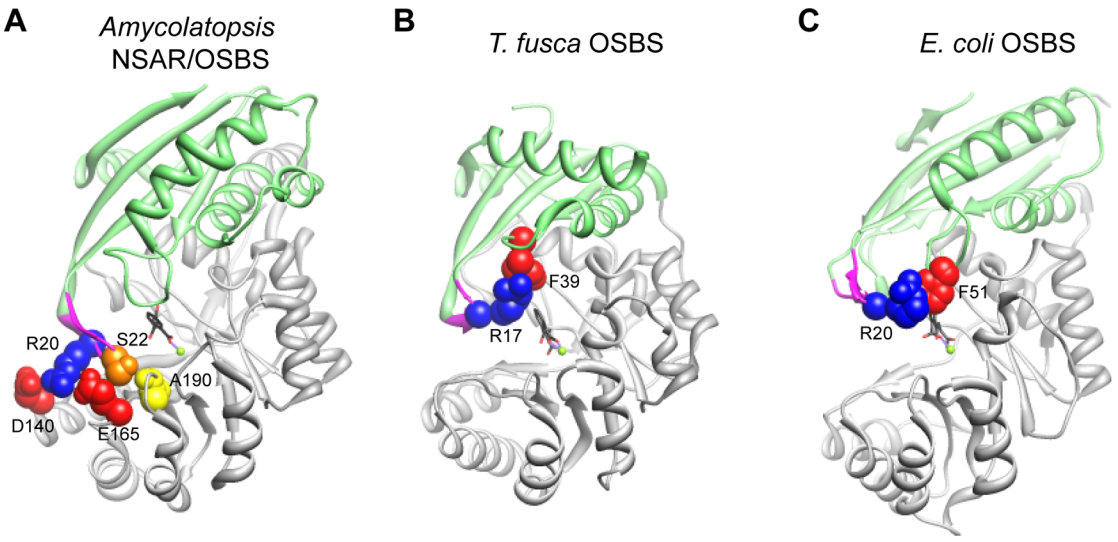


Figure 2. Interaction of the 20s loop with other regions of *Amycolatopsis* NSAR/OSBS, *E. coli* OSBS, and *T. fusca* OSBS. The 20s loop is magenta, the capping domain is green, the barrel domain is gray, and OSB is black. (A) *Amycolatopsis* NSAR/OSBS. R20 is blue, S22 is orange, D140 and E165 are red, and A190 is yellow. (B) *T. fusca* OSBS. R17 is blue, and F39 is red. (C) *E. coli* OSBS. R20 is blue, and F51 is red.

in its 20s loop than homologous OSBS enzymes from *E. coli* and *T. fusca*. The largest effect of a single mutation was the F19A mutation, which decreased OSBS efficiency by approximately 200-fold and NSAR efficiency by approximately 120-fold. This residue directly contacts OSB and *N*-succinylphenylglycine, as well as several other residues in the active site.[33] Another mutation, R20E, reduced OSBS efficiency 32-fold and NSAR efficiency 8-fold, apparently by disrupting a salt bridge between the 20s loop and the barrel domain, rather than through direct interactions with the substrate. Most of the mutations that have modest effects increase the K_M but do not directly contact the substrate, suggesting that the conformation or flexibility of the 20s loop helps determine substrate binding affinity. The K_M for the NSAR reaction (K_M^{NSAR}) was more sensitive to several 20s loop mutations than K_M^{OSBS}. This suggests that amino acid substitutions in the 20s loop could affect specificity differences between promiscuous NSAR/OSBS enzymes and nonpromiscuous OSB synthases. Thus, mutations in the 20s loop could have contributed to the evolution of NSAR activity, although they were probably not sufficient given the relatively small effects of mutations that differentially impact NSAR and OSBS activities.

MATERIALS AND METHODS

Mutagenesis. Site-directed mutagenesis was performed using the QuikChange mutagenesis protocol and two-stage polymerase chain reaction (PCR).[41] Template DNA came from the *Amycolatopsis* sp. T-1-60 NSAR/OSBS gene (GI 49259021) cloned into a pET17b vector (a gift from J.A Gerlt, University of Illinois, Urbana, IL). Mutations were confirmed by sequencing in each direction (Eton Bioscience, Inc.). Primers and PCR conditions are included in Supporting Information.

Protein Expression and Purification. Proteins were expressed in *E. coli* strain BW25113 (*menC::kan*, DE3).[38] Expression in *menC*$^-$ cells ensured that OSBS from the host cell would not copurify with the protein expressed from the plasmid. Cultures were grown for 48 h at 30 °C without induction in Luria−Bertani broth supplemented with carbenicillin and kanamycin, then harvested by centrifugation.

The cell pellet was resuspended in 20 mM Tris (pH 8.0), 5 mM $MgCl_2$, 0.4 mM phenylmethylsulfonyl fluoride (PMSF), and 10 μg/mL DNase I. Resuspended cells were lysed by sonication and centrifuged. The supernatant was filtered, then applied to a 20 mL HiPrep 16/10 DEAE FF column (GE Healthcare). The protein was eluted with a buffer containing 20 mM Tris (pH 8.0), 5 mM $MgCl_2$, and 500 mM NaCl using a linear gradient from 30% elution buffer to 65% elution buffer over 10 column volumes. Fractions containing *Amycolatopsis* NSAR/OSBS variants were identified by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) and combined with $(NH_4)_2SO_4$ at a final concentration of 0.4 M. This sample was then applied to three 5 mL HiTrap Phenyl FF (low sub) columns (GE Healthcare) attached in tandem and equilibrated in a buffer containing 20 mM Tris (pH 8.0), 5 mM $MgCl_2$, and 0.4 M $(NH_4)_2SO_4$. The protein was eluted with 20 mM Tris (pH 8.0) and 5 mM $MgCl_2$ using a linear gradient from 0% elution buffer to 100% elution buffer over 15 column volumes. Fractions containing the protein were identified by SDS−PAGE, and OSBS activity was confirmed before pooling the fractions. If no activity was detected, the same fractions that had activity when purifying the wild-type protein were pooled. Pooled fractions were exchanged into storage buffer (20 mM Tris (pH 8.0) and 5 mM $MgCl_2$) and concentrated using an Amicon Ultra-15 centrifuge filter with a 30 kDa molecular weight cutoff. After supplementing the storage buffer with a final concentration of 25% glycerol, the proteins were stored at 4 °C for immediate use or at −80 °C for longer terms.

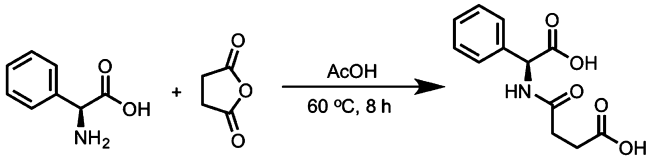
OSBS Activity Assay. 2-Succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) was synthesized from chorismate and α-ketoglutarate using the enzymes from the menaquinone synthesis pathway as described previously.[38] Wild-type and mutant enzymes were assayed using a SpectraMax Plus384 plate reader (Molecular Devices) with varying concentrations of SHCHC in 50 mM Tris (pH 8.0) and 0.1 mM $MnCl_2$ at 25 °C. Reactions for the series of substrate concentrations were monitored by the loss of absorbance of SHCHC at 310 nm ($\Delta\varepsilon$ = −2400 M^{-1} cm^{-1}).[26,30] Initial rates were determined by fitting the linear portion of the curve in Excel (Microsoft) and fitting to the

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434−4444

Michaelis–Menten equation using Kaleidagraph (Synergy Software).

Synthesis of *N*-Succinyl-L-phenylglycine. The synthesis was modified from previously described methods.[27,30] All reactions were carried out under a nitrogen atmosphere in flame-dried glassware. All solvents and commercial reagents were used as received. The following is a general procedure used for all succinyl acid substrates. A mixture of L-phenylglycine (4.99 g, 33.0 mmol, 1 equiv) and succinic anhydride (3.99 g, 39.9 mmol, 1.2 equiv) in 25 mL of acetic acid was heated at 60 °C under N_2 for 8 h (Scheme 1). After

Scheme 1



concentrating *in vacuo* to remove acetic acid, the crude product was purified by flash chromatography (CH_2Cl_2/MeOH = 10:1) with 60 Å silica gel (230–400 mesh) as a stationary phase using an isocratic solvent system to give the product as semisolid (4.2 g, 85%). ^{1}H NMR (500 MHz, DMSO-d_6) δ 8.62 (d, J = 7.6 Hz, 1H), 7.39–7.32 (m, 5H), 5.33 (d, J = 7.6 Hz, 1H), 2.50–2.39 (m, 4H). ^{13}C NMR (125 MHz, DMSO-d_6) δ 0173.9, 1172.0, 2171.0, 137.3, 128.6 (2C), 1128.0, 127.7 (2C), 56.3, 29.7, 29.1. HRMS (ESI$^-$): calculated mass for $C_{12}H_{12}NO_5$ ([M − H]$^-$) is 250.0715. Actual mass found was 250.0706.

^{1}H and ^{13}C NMR spectra (Figure S1, Supporting Information) were recorded at 500 and 125 MHz, respectively. ^{1}H NMR chemical shifts are reported as δ values in ppm relative to DMSO-d_6 (2.50 ppm), coupling constants (J) are reported in Hertz (Hz), and multiplicity follows convention. DMSO-d_6 served as an internal standard (39.52 ppm) for ^{13}C spectra. Mass spectra were obtained at the Laboratory for Biological Mass Spectrometry (Texas A&M University). Thin layer chromatography (TLC) was performed using glass-backed silica gel 60F_{254}.

Specific rotation of *N*-succinyl-L-phenylglycine at 405 nm was found to be 6.54 deg M^{-1} cm^{-1}, as determined by fitting three independent serial dilutions to a straight line.

NSAR Activity Assay. NSAR activity was assayed as previously described but with slight modifications.[27,30] The change in optical rotation of *N*-succinyl-L-phenylglycine was monitored at 405 nm in 200 mM Tris (pH 8.0) and 0.1 mM $MnCl_2$ at 25 °C using a P-2000 polarimeter (Jasco). Measurements were taken using a 10 s integration time, reading every 30 s using a cell with a 5 cm path length. Calculation of initial rates and kinetic parameters was done in the same way as that for the OSBS assay.

NSAR/OSBS Protein Sequence Analysis. NSAR/OSBS subfamily members were identified by using BLAST to search the NCBI Nonredundant protein sequence database for matches to the *Bacillus subtilis* strain 168 OSBS (GI 16080130). Using *B. subtilis* OSBS (a nonpromiscuous member of the NSAR/OSBS subfamily), rather than *Amycolatopsis* NSAR/OSBS, provided better coverage of the NSAR/OSBS subfamily because of it being closer to root sequences that would be otherwise missed. The 1000 best matches were downloaded and separated into clusters sharing at least 40% amino acid sequence identity using CD-HIT.[42,43] Only sequences that were clustered with previously identified NSAR/OSBS proteins were retained.[31] These clusters were aligned separately and combined into a single alignment using MUSCLE's profile option.[44] The alignment was manually refined by comparing it to a structural alignment that included six members of the NSAR/OSBS subfamily (PDB entries 1SJB, 2ZC8, 1XPY, 1WUE, 1WUF, and 3QLD).[29,33,45,46] This set of 872 sequences was filtered to sequences sharing less than 90% sequence identity for further analysis. A cluster containing sequences from the Streptosporangineae suborder was also excluded from analysis in this article due to a large insertion in the 20s loop, producing a final set of 371 sequences. Sequence logos were produced using the WebLogo 3 server.[47,48] Sequence logos for the muconate lactonizing enzyme family and *E. coli* dipeptide epimerase subfamily were based on the sequence alignment from ref 49.

RESULTS

Effect of Mutations on OSBS and NSAR Activity. To establish the importance of the 20s loop for the two reactions of *Amycolatopsis* NSAR/OSBS, we deleted the entire loop. Deleting residues 18 to 26 decreased the efficiency of OSBS and NSAR activities by 4500-fold and 25,000-fold, respectively. The efficiency of the 20s loop-deletion mutant is still well above the rate of noncatalyzed reactions (k_{non}), with catalytic proficiencies ($k_{cat}/K_M/k_{non}$) of 1.1 × 10^{11} M^{-1} for OSBS and 5.7 × 10^{12} M^{-1} for NSAR.[50,51] (Catalytic proficiency for the NSAR reaction was estimated using k_{non} for the mandelate racemase reaction, which is catalyzed by another member of the enolase superfamily and uses a similar reaction mechanism.) Retention of some activity was expected because deleting the 20s loop does not disrupt the catalytic residues. However, retention of <0.1% of wild-type OSBS activity is not sufficient *in vivo* because expressing the 20s loop-deletion mutant in *menC*$^-$ *E. coli* results in very slow growth rates compared to that of the wild type under anaerobic conditions, in which OSBS activity is required for menaquinone synthesis (data not shown). The decrease in OSBS efficiency due to deleting the 20s loop is 25 times greater than the decrease caused by equivalent deletions in *E. coli* and *T. fusca* OSBSs.[28,38] *Amycolatopsis* NSAR/OSBS is more sensitive to mutations in the 20s loop than these representatives from other OSBS subfamilies, suggesting that different subfamilies use this loop in different ways to carry out the same reaction.

To determine which residues in *Amycolatopsis* NSAR/OSBS are responsible for this loss of activity, amino acids in the 20s loop were mutated individually and in combination. The effects of these mutations are shown in Tables 1 and 2 and Figure 3.

The first set of mutations were F19A, T21A, and F23A (Figure 4A). The side chains of these residues face the active site and could make contact with the substrate, influencing substrate binding or catalysis directly. F19 and T21 are also near the side chain of K163, which could help position K163 correctly to abstract a proton in the OSBS reaction or serve as the proton donor or acceptor for the racemase reaction. In *Amycolatopsis* NSAR/OSBS, F19A had the largest effect of any of the single point mutations, decreasing the catalytic efficiencies of OSBS activity by approximately 200-fold and NSAR activity by approximately 120-fold. This could be because it packs against S135 and F323, two other residues that also contact the substrate of both reactions (Figure 4B).

The T21A mutation had similar effects on OSBS and NSAR activities, decreasing catalytic efficiency approximately 4-fold.

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434–4444

Table 1. OSBS Activity Kinetic Parameters

mutation	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	relative k_{cat}/K_m
WT	46 ± 5	550 ± 120	(8.3 ± 2.0) × 10^4	1.00
Δ20s	n.d.[a]	n.d.[a]	18 ± 2	0.000 22
F19A	n.d.[a]	n.d.[a]	(4.1 ± 0.2) × 10^2	0.0049
T21A	7.2 ± 0.9	290 ± 100	(2.4 ± 0.9) × 10^4	0.29
F23A[b]	19 ± 2	600 ± 100	(3.4 ± 0.9) × 10^4	0.40
F19A/T21A	n.d.[a]	n.d.[a]	(9.0 ± 0.1) × 10^2	0.011
F19A/F23A	n.d.[a]	n.d.[a]	39 ± 4	0.000 47
S22R	33 ± 2	130 ± 20	(2.6 ± 0.5) × 10^5	3.09
R20E	n.d.[a]	n.d.[a]	(2.6 ± 0.2) × 10^3	0.031
D140R	14 ± 0.3	450 ± 30	(3.0 ± 0.2) × 10^4	0.36
E165R[b]	38 ± 8	220 ± 30	(1.7 ± 0.4) × 10^5	2.02
R20E/D140R	n.d.[a]	n.d.[a]	(4.3 ± 0.1) × 10^4	0.52
R20E/E165R	n.d.[a]	n.d.[a]	(4.7 ± 0.1) × 10^4	0.56
D140R/ E165R	63 ± 2	390 ± 30	(1.6 ± 0.1) × 10^5	1.95
R20E/ D140R/ E165R	115 ± 5	580 ± 50	(2.0 ± 0.2) × 10^5	2.36
G24A	21 ± 0.9	290 ± 40	(7.1 ± 1) × 10^5	0.86
P18A	52 ± 1.2	450 ± 30	(1.2 ± 0.1) × 10^5	1.38

[a]Not determined because substrate saturation could not be achieved. [b]Because of low expression, the protein copurified with another band and was only ~50% pure according to SDS−PAGE. As a result, k_{cat} could be somewhat higher than that reported.

Like F19, T21 contacts K163, and it also contacts F19, F23, and N191 (Figure 4C). Thus, it could have a small effect on the position of the catalytic base and other active site residues. The F23A mutation reduced OSBS efficiency 2.5-fold. However, it significantly increased both k_{cat}^{NSAR} and K_M^{NSAR}, with little effect on k_{cat}/K_M^{NSAR}. The simplest explanation for the effect of F23A is that product release becomes less rate-limiting in the NSAR reaction. Although the rate-limiting step of the NSAR reaction has not been definitively determined, the rate of deuterium exchange of the α-proton that is abstracted to initiate the reaction is faster than k_{cat}, suggesting that a step

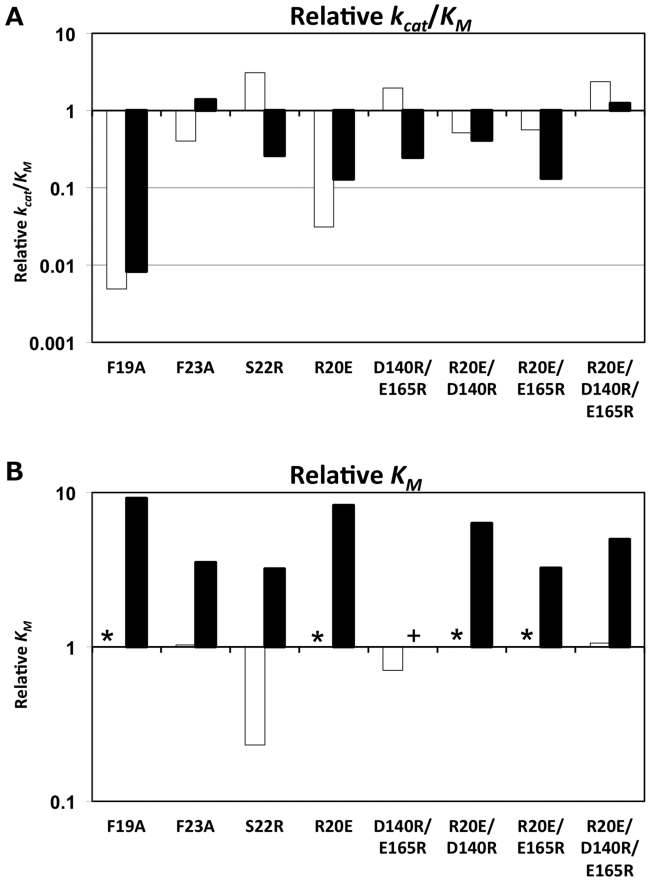


Figure 3. Comparison of kinetic parameters of selected mutants relative to those of the wild type. In both panels, the OSBS reaction parameter is plotted as open bars, and the NSAR reaction parameter is plotted as solid bars. (A) Relative k_{cat}/K_M. (B) Relative K_M. Stars indicate mutations which have unknown K_M^{OSBS} values due to not achieving substrate saturation (minimum value is >1.5 times wild-type K_M^{OSBS}). Plus sign indicates a mutation which has an unknown K_M^{NSAR} value due to not achieving substrate saturation (minimum value is >10 times wild-type K_M^{NSAR}).

Table 2. NSAR Activity Kinetic Parameters[a]

mutation	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (M^{-1} s^{-1})	relative k_{cat}/K_m
WT	42 ± 2	1000 ± 100	(4.3 ± 0.6) × 10^4	1
Δ20s	n.d.[b]	n.d.[b]	1.7 ± 0.2	0.000040
F19A	3.2 ± 0.2	9100 ± 1500	(3.5 ± 0.7) × 10^2	0.0083
T21A	18 ± 1	1500 ± 200	(1.2 ± 0.2) × 10^4	0.28
F23A[c]	210 ± 9	3500 ± 600	(5.9 ± 0.1) × 10^4	1.39
F19A/T21A[c]	n.d.[b]	n.d.[b]	(2.5 ± 0.1) × 10^2	0.0059
F19A/F23A	1.8 ± 0.1	7800 ± 800	(2.3 ± 0.3) × 10^2	0.0054
S22R	35 ± 1	3200 ± 300	(1.1 ± 0.1) × 10^4	0.26
R20E	45 ± 1	8100 ± 1400	(5.5 ± 1.0) × 10^3	0.13
D140R	60 ± 5	4900 ± 1000	(1.2 ± 0.3) × 10^4	0.29
E165R[c]	186 ± 14	4000 ± 900	(4.7 ± 1.0) × 10^4	1.11
R20E/D140R	109 ± 6	6200 ± 900	(1.8 ± 0.3) × 10^4	0.41
R20E/E165R	18 ± 0.6	3200 ± 300	(5.6 ± 0.6) × 10^3	0.13
D140R/E165R	n.d.[b]	n.d.[b]	(1.1 ± 0.2) × 10^4	0.25
R20E/D140R/E165R	240 ± 20	4400 ± 900	(5.4 ± 2.0) × 10^4	1.27
G24A	11 ± 0.8	1000 ± 300	(1.1 ± 0.8) × 10^4	0.26
P18A	73 ± 15	3095 ± 1600	(2.4 ± 1.3) × 10^4	0.56

[a]*N*-Succinyl-L-phenylglycine was used as the substrate. [b]Not determined because substrate saturation could not be achieved. [c]Because of low expression, the protein copurified with another band and was only ~50% pure according to SDS−PAGE. As a result, k_{cat} could be somewhat higher than that reported.

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434−4444

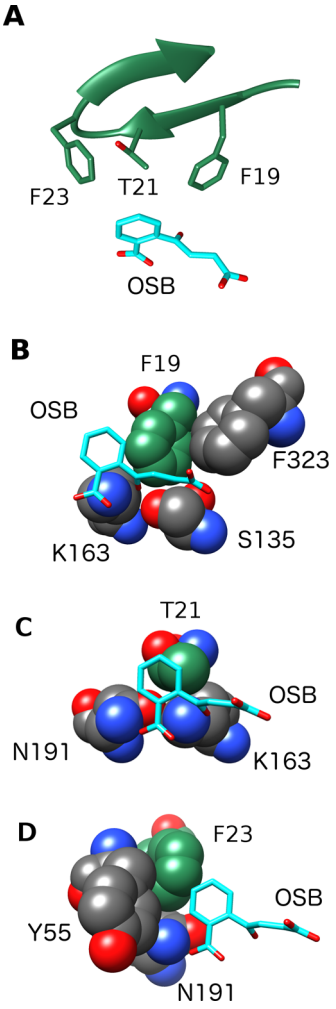


Figure 4. Active site residues in the 20s loop of *Amycolatopsis* NSAR/OSBS.[33] (A) Ribbon representation of 20s loop residues (green). F19, T21, and F23 side chains of *Amycolatopsis* NSAR/OSBS (green) in relation to OSB (cyan) (PDB code 1SJB). (B) Space filling model of F19 (green) with residues that are within 5 Å of both F19 and OSB (gray). OSB is colored as described in panel A. (C) Similar to panel B, but showing residues near T21. (D) Similar to panel B, but showing residues near F23.

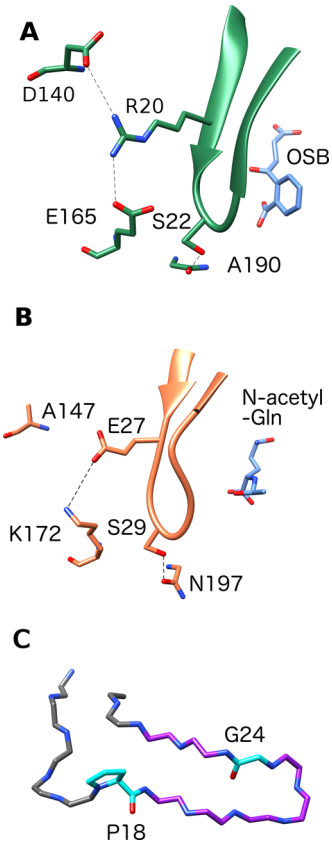


Figure 5. 20s loop residues that do not contact the bound substrate or product. (A) R20 of *Amycolatopsis* NSAR/OSBS forms salt bridges with D140 and E165, and S22 forms a hydrogen bond to the backbone of A190. OSB is shown in light blue. (B) Homologous residues from *D. radiodurans* NSAR/OSBS (PDB: 1XPY chain C) make interactions similar to those shown in A.[45] (C) Backbone atoms of *Amycolatopsis* NSAR/OSBS residues 14 through 28. 20s loop residues 18 to 26 are colored purple except for P18 and G24, which are colored cyan.

after proton abstraction, such as product release, is at least partially rate-limiting.[30] F23 is near the tip of the 20s loop, where it interacts with the substrate, N191, and Y55 (Figure 4D). Eliminating these interactions could hinder the 20s loop from closing, which could reduce substrate and product affinity and increase the rate of product release. The double mutants F19A/T21A and F19A/F23A had effects on NSAR activity similar to that of the F19A mutation alone. In contrast, F19A/F23A was about 10-fold less efficient at carrying out the OSBS reaction than F19A alone. F19A/T21A showed a small restoration of OSBS activity compared to that of the F19A single mutant, but it is not clear whether this 2-fold increase in efficiency is significant.

In addition to the active site side chain interactions discussed above, two surface-exposed residues in the *Amycolatopsis* NSAR/OSBS 20s loop contact the barrel domain. It is plausible that these surface residues could play an important role in OSBS or NSAR activity by closing the 20s loop. The first of these, R20, is equidistant from E165 and D140 (Figures 2A and 5A). Another NSAR/OSBS from *Deinococcus radiodurans* also has a similar interaction with the homologous positions E27 and K172, but it has an alanine at position 147 (Figure 5B). This potential interaction was disrupted by R20E, D140R, and E165R mutations. R20E reduced k_{cat}/K_M^{OSBS} 30-fold and k_{cat}/K_M^{NSAR} 8-fold. Mutating D140 or E165 individually had a smaller effect. This cannot be easily explained by the fact that one partner of the salt bridge to R20 remains since the activity of the double mutant D140R/E165R was not additive (D140R/E165R has OSBS activity similar to that of E165R and NSAR activity similar to that of D140R). Restoring the salt bridge with the double mutation R20E/D140R partially restored OSBS and NSAR efficiencies, but the K_M was still significantly higher than that of the wild type. R20E/E165R also partially restored OSBS activity, but it did not restore NSAR activity. However, the triple mutant R20E/D140R/E165R restored the efficiency of both reactions back to wild-type levels. This suggests that the salt bridge interactions between the 20s loop and the barrel domain contribute to catalytic efficiency. However, the positions of positive and negative charges also appear to be important, especially for NSAR activity, because K_M^{NSAR} of R20E/D140R/E165R was still quite a bit higher than that of the wild type and because the effects of some mutations were not as expected (E165R had little effect on efficiency, and D140R/E165R was not additive and had different effects on OSBS and NSAR activity).

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434−4444

S22 also contacts the barrel domain via a hydrogen bond with the backbone carbonyl of A190 (Figure 5A). Serine is highly conserved at that position in the NSAR/OSBS subfamily, but it is often arginine or other amino acids in the rest of the OSBS family. In *D. radiodurans* NSAR/OSBS, this serine only makes a hydrogen bond in one of the monomers of the crystal structure and is positioned slightly differently in the other monomer. We made the S22R mutation because alanine is already present at this position in some NSAR/OSBS enzymes, and arginine occurs at this position in many other OSBS subfamilies. This mutation had little effect on the k_{cat} of either reaction, but K_M^{OSBS} decreased while K_M^{NSAR} increased, resulting in opposite effects on catalytic efficiency of the these two reactions.

The last two positions we examined were P18 and G24, which are conserved in most of the NSAR/OSBS subfamily and could be important for positioning or ordering the 20s loop. P18 is located near where the peptide backbone bends almost 90 deg to begin forming the 20s loop, and the rigid nature of the proline side chain could help determine this change of direction (Figure 5C). However, P18A had little effect on OSBS or NSAR efficiency. G24 is located where the 20s loop forms a turn, and the torsion angle needed for this turn requires a glycine in this position. However, G24A had near wild-type OSBS activity, and the NSAR activity decreased 4-fold due to a lower k_{cat}^{NSAR}. Thus, P18 and G24 make a very small contribution to positioning the 20s loop for catalysis, and that small effect appears to only be important for NSAR activity.

Conservation of the 20s Loop in the NSAR/OSBS Subfamily. Previous sequence analysis showed that the 20s loop is highly divergent within the OSBS family, contrary to expectations if the 20s loop determines substrate specificity (Figure 6A).[31,38] However, within different subfamilies there is more conservation, either because different residues in the 20s loop make important interactions in each subfamily or because the sequences share a more recent common ancestor (Figure 6B–D).[38] We previously showed that the most important residues in the 20s loops of *E. coli* and *T. fusca* OSB synthases are arginine residues that interact with the adjacent 50s loop to partially close the active site (Figure 2).[28] Not only is the sequence of the 20s loop in the NSAR/OSBS subfamily quite different, but the loop plays a different structural role, completely closing the active site when a ligand is bound.[33]

The 20s loop extends from position 18–26. In this study, positions 25 and 26 were not mutated because they are more variable and also more distant from bound OSB in the structure of *Amycolatopsis* NSAR/OSBS.[33] A mutation at F19, which is conserved in most of the NSAR/OSBS subfamily, had the largest effect among the single mutations, whereas mutating F23, which is variable, had little impact on catalytic efficiency. The third position (R20 in *Amycolatopsis* NSAR/OSBS) is often a charged residue. *Amycolatopsis* NSAR/OSBS and *D. radiodurans* NSAR/OSBS have oppositely charged amino acids at this position, both of which form a salt bridge with a homologous residue in the barrel domain (E165 of *Amycolatopsis* NSAR/OSBS or K172 of *D. radiodurans* NSAR/OSBS). In fact, 41% of the NSAR/OSBS subfamily could form at least one salt bridge between this position in the 20s loop and either position in the barrel domain. It will be interesting to subdivide the NSAR/OSBS subfamily according to their biological function (OSBS, NSAR, or bifunctional) to determine if some of the sequence variation reflects special-

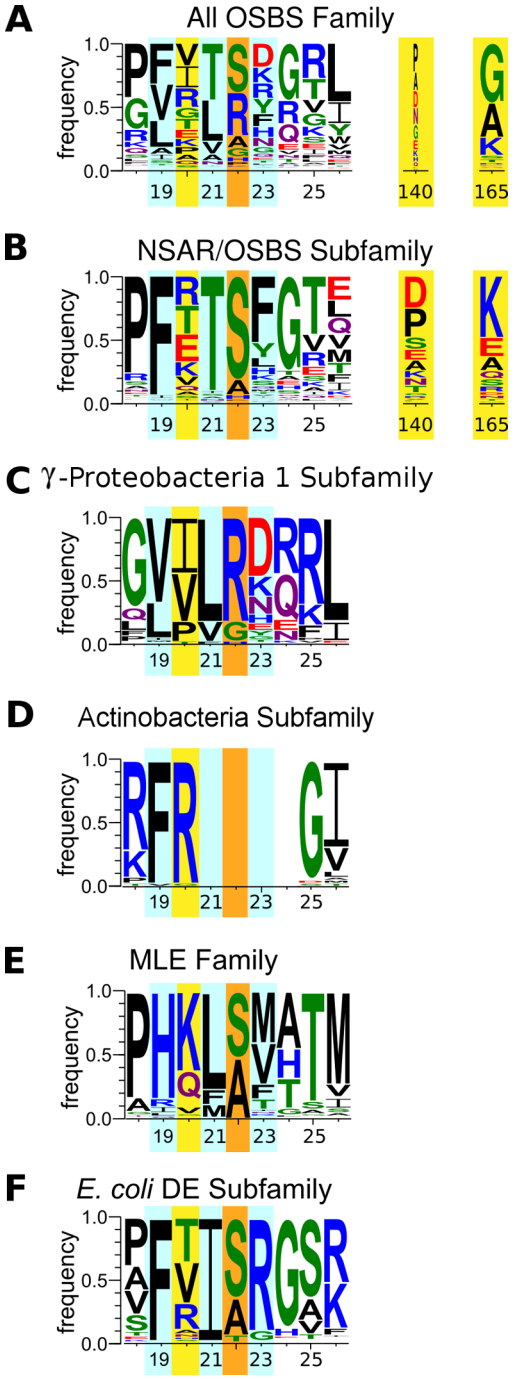


Figure 6. Sequence logos showing conservation of the 20s loop and potential salt bridge partners. Letter size is proportional to the frequency at which an amino acid is found in the sequence alignment. Sequence numbering is relative to the *Amycolatopsis* NSAR/OSBS protein. (A) All OSBS family members. (B) NSAR/OSBS subfamily. (C) γ-Proteobacteria 1 subfamily, which includes the *E. coli* OSBS. (D) Actinobacteria subfamily, which includes *T. fusca* OSBS. (E) Muconate lactonizing enzyme family, a related family in the enolase superfamily. (F) *E. coli* subfamily of dipeptide epimerase enzymes, which is a large family with diversity similar to that of the OSBS family. Includes the *E. coli* L-Ala-D/L-Glu epimerase that was mutated to confer OSBS activity.[14,40] Background colors indicate the structural position of the residues in the *Amycolatopsis* NSAR/OSBS crystal structure. Turquoise background indicates residues that point into the active site. Yellow indicates residues that form a salt bridge. Orange indicates residues that form a hydrogen bond with a backbone carbonyl.

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434–4444

ization for NSAR or OSBS activity, or even perhaps specificity for different succinylamino acids.

DISCUSSION

Role of the 20s Loop Is Not Conserved in the Enolase Superfamily. Because the 20s loop is a conserved structural element in the active sites of enolase superfamily enzymes, it has been thought that sequence divergence in the 20s loop contributes to divergence of specificity in the superfamily.[39] Our work here and elsewhere paints a more complex picture, in which the 20s loop plays different roles in the context of different proteins. In the *Amycolatopsis* NSAR/OSBS, the 20s loop is essential for achieving biologically relevant activity levels of both NSAR and OSBS activities. A single conserved residue (F19) is largely responsible for this effect. A conserved phenylalanine is found at the same position in the dipeptide epimerase family, and a histidine that forms a hydrogen bond to the ligand is generally conserved at this position in the muconate lactonizing enzyme family (Figure 6E and F).[38,49,52] Discovering that the critical phenylalanine is conserved in related families suggests that the 20s loop serves a structural role in *Amycolatopsis* NSAR/OSBS and is not the primary determinant of specificity.

However, the hydrophobic packing contributed by F19, the methyl group of T21, and F23 of the *Amycolatopsis* NSAR/OSBS 20s loop distinguishes it from the 20s loops of many enolase superfamily members. In muconate lactonizing enzyme, a histidine in the 20s loop makes a hydrogen bond with the ligand, implicating a role in determining substrate specificity.[52] In several members of the dipeptide epimerase family, specificity for different dipeptides is determined by hydrogen bonds between 20s loop residues and the ligand.[13,53] Likewise, most of the acid-sugar dehydratases have salt bridges or hydrogen bonds between the 20s loop and the carboxylate of a diacid sugar, but the positions of these charged and polar residues are not identical.[34,54−58]

In fact, increasing the hydrophobic packing of the 20s loop in *E. coli* L-Ala-D/L-Glu epimerase, in combination with a mutation in the barrel domain, changes its specificity from dipeptide epimerization to OSB synthesis.[14,40,59] The mutation in the barrel domain (D297G) removes a steric conflict with the succinyl group of SHCHC, raising k_{cat}/K_M^{OSBS} from undetectable to 12.5 M^{-1} s^{-1}. Two additional mutations in the 20s loop (I19F and R24W, which correspond to positions T21 and Q26 in *Amycolatopsis* NSAR/OSBS) increase the mutant's k_{cat}/K_M^{OSBS} 100-fold further to 2.1 × 10^3 M^{-1} s^{-1}, which is only 40-fold lower than that of the *Amycolatopsis* NSAR/OSBS. These mutations essentially abolish L-Ala-D/L-Glu epimerase activity. In *Amycolatopsis* NSAR/OSBS, the methyl group of T21, not its hydroxyl, is nearest the ligand, and Q26 is >6 Å from the ligand. Because the parts of SHCHC and *N*-succinyl-L-phenylglycine that are accessible to the 20s loop are hydrophobic, interaction with F19 and the lack of nearby polar or charged atoms in the *Amycolatopsis* NSAR/OSBS could help distinguish between its substrates and other enolase superfamily substrates like polar/charged dipeptides or diacid sugars.

The 20s loop mutations in *E. coli* L-Ala-D/L-Glu epimerase increased its OSBS activity mainly by decreasing K_M^{OSBS}. Conversely, mutating F19 and several other positions in *Amycolatopsis* NSAR/OSBS increased the K_M, especially for NSAR activity. Although *N*-succinylphenylglycine is probably not the natural substrate for NSAR enzymes, it was used because catalytic efficiency with succinylphenylalanine and derivatives of other natural amino acids are at least an order of magnitude lower, which made it difficult to measure activity using a low sensitivity method like polarimetry (Table S2, Supporting Information). If the mutations examined here have similar effects on the K_M for natural substrates, they could have a significant impact on flux through metabolic pathways that use NSAR activity *in vivo* because typical amino acid concentrations in cells are ∼0.5−3 mM, which are well below the ∼5−9 mM K_M's of some variants.[60]

The crystal structure suggests that salt bridges between R20 in the 20s loop and D140 and E165 in the barrel domain could stabilize the 20s loop in the closed position, helping to lower the K_M. Indeed, most salt bridge mutations increased the K_M, and the loss of catalytic efficiency due to R20E was restored by the triple mutant R20E/D140R/E165R. However, it is unclear whether restoration of wild-type OSBS efficiency by the triple mutant R20E/D140R/E165R was due to restoration of the salt bridges or whether the effects of R20E and D140R were masked by E165R, which conferred nearly wild-type efficiency by itself. In contrast, NSAR activity appears to be more sensitive to mutations in the salt bridge than OSBS activity. D140R, E165R, and D140R/E165R increased K_M^{NSAR} but not K_M^{OSBS}. In addition, k_{cat}/K_M^{NSAR} of the triple mutant R20E/D140R/E165R was similar to that of the wild-type enzyme, but K_M^{NSAR} was still elevated. These results suggest that these charged surface residues may have had a role in the evolution of NSAR activity. If so, we would expect to see a frequent occurrence of charged amino acids at positions equivalent to R20, D140, and E165 in other subfamily members whose biological function is NSAR. Indeed, 41% of the NSAR/OSBS subfamily proteins have the potential to form salt bridges at these positions. The salt bridge is present in *D. radiodurans* NSAR/OSBS, which is predicted to function as an NSAR *in vivo* because *D. radiodurans* has an alternative menaquinone synthesis pathway that does not require OSBS.[31,61] In contrast, the salt bridge is absent in *B. subtilis* OSBS and *Staphylococcus aureus* OSBS, which are encoded in menaquinone synthesis operons and do not have NSAR activity.[26,29,31] Testing this idea further will require predicting the biological functions of the ∼400 members of the NSAR/OSBS subfamily to supplement these anecdotal examples.

Thus, the 20s loop of *Amycolatopsis* NSAR/OSBS affects the enzyme's activity and possibly specificity two ways: via hydrophobic interactions between F19 and the ligand and through interactions between surface residues of the 20s loop and residues in the barrel domain that appear to help latch the 20s loop into the closed position. The observation that the 20s loop makes critical interactions in *Amycolatopsis* NSAR/OSBS is in line with the idea that the 20s loop helps determine specificity in the enolase superfamily.[39] The role of the 20s loop in determining specificity does not appear universal, however. The 20s loop of mandelate racemase, another member of the enolase superfamily, is relatively tolerant of mutations, suggesting that it primarily functions as a flexible flap that closes when the ligand is bound.[15]

In addition, mutagenesis of the 20s loops of two distantly related OSBS enzymes indicate that their 20s loops are more important for the structural integrity of the active site than for determining specificity.[28,38] *E. coli* OSBS and *T. fusca* OSBS share <30% amino acid sequence identity with the *Amycolatopsis* NSAR/OSBS. They are not part of the NSAR/OSBS subfamily, and they only catalyze the OSBS reaction.[26,29]

dx.doi.org/10.1021/bi500573v | *Biochemistry* 2014, 53, 4434−4444

The structures of these enzymes bound to OSB show that the active site is partially open, unlike that in *Amycolatopsis* NSAR/OSBS (Figure 2).[28,33,37] In *E. coli* OSBS, the 20s loop is also twisted so that only L19, which is homologous to T21, could contact the substrate.[37] In *T. fusca* OSBS, a deletion in the 20s loop leaves only F16 facing the active site.[28] Point mutations of L19 in *E. coli* OSBS and F16 in *T. fusca* OSBS reduce catalytic efficiency only ~6-fold, and deleting the 20s loop of both enzymes reduces catalytic efficiency by 300-fold, compared to 4500-fold for the *Amycolatopsis* NSAR/OSBS. Instead of interacting with the barrel domain, the 20s loops of *E. coli* and *T. fusca* OSBSs form cation-π interactions with the 50s loop, which is the other active site loop in the capping domain. Mutations that abolish this interaction in *E. coli* and *T. fusca* OSB synthases reduce catalytic efficiency 500-fold or 100-fold, respectively. In *Amycolatopsis* NSAR/OSBS, F23 contacts the 50s loop, but effects of mutations at this site were minimal for OSBS activity, although this mutation did increase K_M^{NSAR}, while maintaining catalytic efficiency similar to that of the wild type. Thus, the 20s loops of *E. coli* and *T. fusca* OSBSs form a ledge that partially occludes the entrance to the active site when the ligand is bound, but they do not strongly interact with the ligand. In contrast, the 20s loop in *Amycolatopsis* NSAR/OSBS completely closes the active site, and F19 forms an important part of the substrate binding pocket.

Role for the 20s Loop in the Evolution of NSAR Activity? Our primary reason for studying the NSAR/OSBS subfamily is to determine how catalytic promiscuity has been exploited in order to evolve a new enzyme function. Differences in the positions and sequences of the 20s loop between the promiscuous *Amycolatopsis* NSAR/OSBS and two distantly related, nonpromiscuous OSBS enzymes suggested that it could be an important determinant for NSAR specificity. Our data show that the 20s loop is necessary for NSAR activity. Although many mutations affected NSAR and OSBS activities equally, S22R had opposite effects on NSAR and OSBS activities, R20E and F19A/F23A had a larger effect on OSBS activity, and several mutations (F23A, D140R, E165R, D140R/E165R, and R20E/D140R/E165R) had a larger effect on K_M^{NSAR} (Figure 3B). While these mutations had a larger effect on one of the reactions, none resulted in severe reduction of NSAR activity without also affecting OSBS activity (or vice versa). For this reason, we do not expect that substitutions in the 20s loop were the only changes needed to evolve NSAR activity. The experiments described here lay the groundwork for future experiments on nonpromiscuous members of the NSAR/OSBS family, such as *B. subtilis* OSBS,[26] to determine what amino acid substitutions are sufficient for conferring NSAR activity.

■ ASSOCIATED CONTENT

Supporting Information

Detailed methods, primers, and templates used for site directed mutagenesis, kinetics parameters using *N*-succinyl-L-phenylalanine as an NSAR substrate, and NMR spectra of *N*-succinyl-L-phenylglycine. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: 979-458-0123. E-mail: margy.glasner@tamu.edu.

Funding

This work was supported by grant A-1758 from the Robert A. Welch Foundation (Principle Investigator: M. E. Glasner). Seed funding to the Natural Products LINCHPIN Laboratory to foster this collaboration was from the Office of the Vice President for Research, the College of Science, and the Department of Chemistry of Texas A&M University.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Mr. Denis Odokonyero for technical assistance and helpful discussions and Mr. Qutaiba Ababneh, Ms. Lauren Kustigian, and Mr. Ting-Yi Wang for conducting preliminary experiments. We also thank Dr. Michael Hicks, Ms. Stephanie Lucas, and Dr. Patricia Babbitt for discussions and support in developing this project.

■ ABBREVIATIONS

OSBS, *o*-succinylbenzoate synthase; NSAR, *N*-succinylamino acid racemase; SHCHC, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate; PMSF, phenylmethylsulfonyl fluoride

■ REFERENCES

(1) Fetrow, J. S. (1995) Omega loops: nonregular secondary structures significant in protein function and stability. *FASEB J. 9*, 708−717.

(2) Granum, D. M., Schutt, T. C., and Maupin, C. M. (2014) Computational evaluation of the dynamic fluctuations of peripheral loops enclosing the catalytic tunnel of a family 7 cellobiohydrolase. *J. Phys. Chem. B 118*, 5340−5349.

(3) Kempf, J. G., Jung, J. Y., Ragain, C., Sampson, N. S., and Loria, J. P. (2007) Dynamic requirements for a functional protein hinge. *J. Mol. Biol. 368*, 131−149.

(4) Kempner, E. S. (1993) Movable lobes and flexible loops in proteins. Structural deformations that control biochemical activity. *FEBS Lett. 326*, 4−10.

(5) Kursula, I., Salin, M., Sun, J., Norledge, B. V., Haapalainen, A. M., Sampson, N. S., and Wierenga, R. K. (2004) Understanding protein lids: structural analysis of active hinge mutants in triosephosphate isomerase. *Protein Eng., Des. Sel. 17*, 375−382.

(6) Leszczynski, J. F., and Rose, G. D. (1986) Loops in globular-proteins - a novel category of secondary structure. *Science 234*, 849−855.

(7) Massi, F., Wang, C. Y., and Palmer, A. G. (2006) Solution NMR and computer simulation studies of active site loop motion in triosephosphate isomerase. *Biochemistry 45*, 10787−10794.

(8) Ochoa-Leyva, A., Barona-Gómez, F., Saab-Rincón, G., Verdel-Aranda, K., Sánchez, F., and Soberón, X. (2011) Exploring the structure-function loop adaptability of a $(\beta/\alpha)_8$-barrel enzyme through loop swapping and hinge variability. *J. Mol. Biol. 411*, 143−157.

(9) Rozovsky, S., Jogl, G., Tong, L., and McDermott, A. E. (2001) Solution-state NMR investigations of triosephosphate isomerase active site loop motion: Ligand release in relation to active site loop dynamics. *J. Mol. Biol. 310*, 271−280.

(10) Rozovsky, S., and McDermott, A. E. (2001) The time scale of the catalytic loop motion in triosephosphate isomerase. *J. Mol. Biol. 310*, 259−270.

(11) Sullivan, S. M., and Holyoak, T. (2008) Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proc. Natl. Acad. Sci. U.S.A. 105*, 13829−13834.

(12) Banerjee, S., Pieper, U., Kapadia, G., Pannell, L. K., and Herzberg, O. (1998) Role of the omega-loop in the activity, substrate specificity, and structure of class A beta-lactamase. *Biochemistry 37*, 3286−3296.

(13) Lukk, T., Sakai, A., Kalyanaraman, C., Brown, S. D., Imker, H. J., Song, L., Fedorov, A. A., Fedorov, E. V., Toro, R., Hillerich, B., Seidel, R., Patskovsky, Y., Vetting, M. W., Nair, S. K., Babbitt, P. C., Almo, S. C., Gerlt, J. A., and Jacobson, M. P. (2012) Homology models guide discovery of diverse enzyme specificities among dipeptide epimerases in the enolase superfamily. *Proc. Natl. Acad. Sci. U.S.A. 109*, 4122−4127.

(14) Vick, J. E., Schmidt, D. M. Z., 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.

(15) Bourque, J. R., and Bearne, S. L. (2008) Mutational analysis of the active site flap (20s loop) of mandelate racemase. *Biochemistry 47*, 566−578.

(16) Malabanan, M. M., Amyes, T. L., and Richard, J. P. (2010) A role for flexible loops in enzyme catalysis. *Curr. Opin. Struct. Biol. 20*, 702−710.

(17) Schutz, C. N., and Warshel, A. (2001) What are the dielectric "constants" of proteins and how to validate electrostatic models? *Proteins 44*, 400−417.

(18) Teixeira, V. H., Cunha, C. A., Machuqueiro, M., Oliveira, A. S. F., Victor, B. L., Soares, C. M., and Baptista, A. M. (2005) On the use of different dielectric constants for computing individual and pairwise terms in Poisson-Boltzmann studies of protein ionization equilibrium. *J. Phys. Chem. B 109*, 14691−14706.

(19) Barnett, S. A., Amyes, T. L., Wood, B. M., Gerlt, J. A., and Richard, J. P. (2010) Activation of R235A mutant orotidine 5′-monophosphate decarboxylase by the guanidinium cation: effective molarity of the cationic side chain of Arg-235. *Biochemistry 49*, 824−826.

(20) Miller, B. G., Snider, M. J., Short, S. A., and Wolfenden, R. (2000) Contribution of enzyme-phosphoribosyl contacts to catalysis by orotidine 5′-phosphate decarboxylase. *Biochemistry 39*, 8113−8118.

(21) Li, C., Sato, K., Monari, S., Salard, I., Sola, M., Banfield, M. J., and Dennison, C. (2011) Metal-binding loop length is a determinant of the pK_a of a histidine ligand at a type 1 copper site. *Inorg. Chem. 50*, 482−488.

(22) Toma, S., Campagnoli, S., Margarit, I., Gianna, R., Grandi, G., Bolognesi, M., Defilippis, V., and Fontana, A. (1991) Grafting of a calcium-binding loop of thermolysin to *Bacillus subtilis* neutral protease. *Biochemistry 30*, 97−106.

(23) Tawfik, D. S. (2006) Biochemistry. Loop grafting and the origins of enzyme species. *Science 311*, 475−476.

(24) Afriat-Jurnou, L., Jackson, C. J., and Tawfik, D. S. (2012) Reconstructing a missing link in the evolution of a recently diverged phosphotriesterase by active-site loop remodeling. *Biochemistry 51*, 6047−6055.

(25) Meganathan, R. (2001) Biosynthesis of Menaquinone (Vitamin K_2) and Ubiquinone (Coenzyme Q): A Perspective on Enzymatic Mechanisms, in *Vitamins & Hormones* (Litwack, G., Ed.) pp 173−218, Academic Press, Waltham, MA.

(26) 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*, 4252−4258.

(27) 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.

(28) Odokonyero, D., Ragumani, S., Swaminathan, S., Lopez, M. S., Bonanno, J. B., Ozerova, N. D., Woodard, D. R., Machala, B. W., Burley, S. K., Almo, S. C., and Glasner, M. E. (2013) Divergent evolution of ligand binding in the *o*-succinylbenzoate synthase family. *Biochemistry 52*, 7512−7521.

(29) Odokonyero, D., Sakai, A., Patskovsky, Y., Malashkevich, V. N., Fedorov, A. A., Bonanno, J. B., Fedorov, E. V., Toro, R., Agarwal, R., Wang, C., Ozerova, N. D. S., Yew, W. S., Sauder, J. M., Swaminathan, S., Burley, S. K., Almo, S. C., and Glasner, M. E. (2014) Loss of quaternary structure is associated with rapid sequence divergence in the OSBS family. *Proc. Natl. Acad. Sci. U.S.A. 111*, 8345−8340.

(30) 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.

(31) 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.

(32) Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., de Crecy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A. C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G. D., Rodionov, D. A., Ruckert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., and Vonstein, V. (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res. 33*, 5691−5702.

(33) 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.

(34) Gulick, A. M., Hubbard, B. K., Gerlt, J. A., and Rayment, I. (2000) Evolution of enzymatic activities in the enolase superfamily: crystallographic and mutagenesis studies of the reaction catalyzed by D-glucarate dehydratase from *Escherichia coli*. *Biochemistry 39*, 4590−4602.

(35) 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) The role of lysine 166 in the mechanism of mandelate racemase from *Pseudomonas putida*: mechanistic and crystallographic evidence for stereospecific alkylation by (R)-alpha-phenylglycidate. *Biochemistry 33*, 635−643.

(36) Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R. S., Kenyon, G. L., and Gerlt, J. A. (1991) Mechanism of the reaction catalyzed by mandelate racemase. 2. Crystal structure of mandelate racemase at 2.5-Å resolution: identification of the active site and possible catalytic residues. *Biochemistry 30*, 9264−9273.

(37) 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.

(38) Zhu, W. W., Wang, C., Jipp, J., Ferguson, L., Lucas, S. N., Hicks, M. A., and Glasner, M. E. (2012) Residues required for activity in *Escherichia coli o*-succinylbenzoate synthase (OSBS) are not conserved in all OSBS enzymes. *Biochemistry 51*, 6171−6181.

(39) Gerlt, J. A., Babbitt, P. C., Jacobson, M. P., and Almo, S. C. (2012) Divergent evolution in enolase superfamily: strategies for assigning functions. *J. Biol. Chem. 287*, 29−34.

(40) 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.

(41) Wang, W., and Malcolm, B. A. (1999) Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange site-directed mutagenesis. *Biotechniques 26*, 680−682.

(42) Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics 26*, 680−682.

(43) Li, W., and Godzik, A. (2006) CD-HIT: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics 22*, 1658−1659.

(44) Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res. 32*, 1792–1797.
(45) Wang, W. C., Chiu, W. C., Hsu, S. K., Wu, C. L., Chen, C. Y., Liu, J. S., and Hsu, W. H. (2004) Structural basis for catalytic racemization and substrate specificity of an *N*-acylamino acid racemase homologue from *Deinococcus radiodurans*. *J. Mol. Biol. 342*, 155–169.
(46) Hayashida, M., Kim, S. H., Takeda, K., Hisano, T., and Miki, K. (2008) Crystal structure of *N*-acylamino acid racemase from *Thermus thermophilus* HB8. *Proteins 71*, 519–523.
(47) Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res. 18*, 6097–6100.
(48) Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res. 14*, 1188–1190.
(49) 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.
(50) 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*, 5824–5825.
(51) Bearne, S. L., and Wolfenden, R. (1997) Mandelate racemase in pieces: effective concentrations of enzyme functional groups in the transition state. *Biochemistry 36*, 1646–1656.
(52) Sakai, A., Fedorov, A. A., Fedorov, E. V., Schnoes, A. M., Glasner, M. E., Brown, S., Rutter, M. E., Bain, K., Chang, S., Gheyi, T., Sauder, J. M., Burley, S. K., Babbitt, P. C., Almo, S. C., and Gerlt, J. A. (2009) Evolution of enzymatic activities in the enolase superfamily: stereochemically distinct mechanisms in two families of *cis,cis*-muconate lactonizing enzymes. *Biochemistry 48*, 1445–1453.
(53) 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*, 10370–10378.
(54) Rakus, J. F., Fedorov, A. A., Fedorov, E. V., Glasner, M. E., Hubbard, B. K., Delli, J. D., Babbitt, P. C., Almo, S. C., and Gerlt, J. A. (2008) Evolution of enzymatic activities in the enolase superfamily: L-rhamnonate dehydratase. *Biochemistry 47*, 9944–9954.
(55) Rakus, J. F., Kalyanaraman, C., Fedorov, A. A., Fedorov, E. V., Mills-Groninger, F. P., Toro, R., Bonanno, J., Bain, K., Sauder, J. M., Burley, S. K., Almo, S. C., Jacobson, M. P., and Gerlt, J. A. (2009) Computation-facilitated assignment of the function in the enolase superfamily: a regiochemically distinct galactarate dehydratase from *Oceanobacillus iheyensis*. *Biochemistry 48*, 11546–11558.
(56) Yew, W. S., Fedorov, A. A., Fedorov, E. V., Rakus, J. F., Pierce, R. W., Almo, S. C., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: L-fuconate dehydratase from *Xanthomonas campestris*. *Biochemistry 45*, 14582–14597.
(57) Yew, W. S., Fedorov, A. A., Fedorov, E. V., Almo, S. C., and Gerlt, J. A. (2007) Evolution of enzymatic activities in the enolase superfamily: L-talarate/galactarate dehydratase from *Salmonella typhimurium* LT2. *Biochemistry 46*, 9564–9577.
(58) Yew, W. S., Fedorov, A. A., Fedorov, E. V., Wood, B. M., Almo, S. C., and Gerlt, J. A. (2006) Evolution of enzymatic activities in the enolase superfamily: D-tartrate dehydratase from *Bradyrhizobium japonicum*. *Biochemistry 45*, 14598–14608.
(59) 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) Evolutionary potential of $(\beta/\alpha)_8$-barrels: functional promiscuity produced by single substitutions in the enolase superfamily. *Biochemistry 42*, 8387–8393.
(60) Bennett, B. D., Kimball, E. H., Gao, M., Osterhout, R., Van Dien, S. J., and Rabinowitz, J. D. (2009) Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol. 5*, 593–599.
(61) Hiratsuka, T., Furihata, K., Ishikawa, J., Yamashita, H., Itoh, N., Seto, H., and Dairi, T. (2008) An alternative menaquinone biosynthetic pathway operating in microorganisms. *Science 321*, 1670–1673.
(62) 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.