Evolution of Enzymatic Activity in the Enolase Superfamily: Structural Studies of the Promiscuous *o*-Succinylbenzoate Synthase from *Amycolatopsis*[†,‡]

James B. Thoden,[§] Erika A. Taylor Ringia,[||] James B. Garrett,[||] John A. Gerlt,*[,||] Hazel M. Holden,[§] and Ivan Rayment*[,§]

Departments of Chemistry and Biochemistry, University of Illinois, Urbana, Illinois 61801, and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53705

Received January 29, 2004; Revised Manuscript Received March 8, 2004

ABSTRACT: Divergent evolution of enzyme function is commonly explained by a gene duplication event followed by mutational changes that allow the protein encoded by the copy to acquire a new function. An alternate hypothesis is that this process is facilitated when the progenitor enzyme acquires a second function while maintaining the original activity. This phenomenon has been suggested to occur in the *o*-succinylbenzoate synthase (OSBS) from a species of *Amycolatopsis* that catalyzes not only the physiological syn-dehydration reaction of 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate but also an accidental racemization of *N*-acylamino acids [Palmer, D. R., Garrett, J. B., Sharma, V., Meganathan, R., Babbitt, P. C., and Gerlt, J. A. (1999) *Biochemistry 38*, 4252−4258]. To understand the molecular basis of this promiscuity, three-dimensional structures of liganded complexes of this enzyme have been determined, including the product of the OSBS reaction and three *N*-acylamino acid substrates for the *N*-acylamino acid racemase (NAAAR) reaction, *N*-acetylmethionine, *N*-succinylmethionine, and *N*-succinylphenylglycine, to 2.2, 2.3, 2.1, and 1.9 Å resolution, respectively. These structures show how the active-site cavity can accommodate both the hydrophobic substrate for the OSBS reaction and the substrates for the accidental NAAAR reaction. As expected, the *N*-acylamino acid is sandwiched between lysines 163 and 263, which function as the catalytic bases for the abstraction of the α-proton in the (*R*)- and (*S*)-racemization reactions, respectively [Taylor Ringia, E. A., Garrett, J. B, Thoden, J. B., Holden, H. M., Rayment, I., and Gerlt, J. A. (2004) *Biochemistry 42*, 224−229]. Importantly, the protein forms specific favorable interactions with the hydrophobic amino acid side chain, α-carbon, carboxylate, and the polar components of the N-acyl linkage. Accommodation of the components of the N-acyl linkage appears to be the reason that this enzyme is capable of a racemization reaction on these substrates, whereas the orthologous OSBS from *Escherichia coli* lacks this functionality.

Divergent evolution of enzyme function is commonly explained by a gene duplication event followed by mutational changes that allow the protein encoded by the copy to acquire a new function. In one variation, the progenitor catalyzes a single reaction, but an alternate chemical reaction is achieved following a limited number of mutations. In a second variation, the gene duplication event is preceded by a period of "gene sharing" during which the enzyme is functionally promiscuous, i.e., it catalyzes multiple reactions. After gene duplication, the promiscuous reaction catalyzed by the copy provides a selective advantage if improved activity is accessible by a small number of mutations. Many examples of naturally promiscuous enzymes are known, thereby supporting the likely importance of the second variation (*1*).

In earlier work it was demonstrated that the *o*-succinylbenzoate synthase (OSBS)[1] from a species of *Amycolatopsis* catalyzes not only the physiological syn-dehydration reaction but also an accidental *N*-acylamino acid racemase (NAAAR) reaction (Scheme 1; *2*, *3*). This OSBS is a member of the muconate lactonizing enzyme (MLE) subgroup of the mechanistically diverse enolase superfamily (*4*). In this group, two lysine residues are located on opposite faces of

[†] This research was supported by Grants GM-52594 (to J.A.G. and I.R.), AR-35186 (to I.R.), and GM-55513 to H.M.H. from the National Institutes of Health. Use of the Argonne National Laboratory Structural Biology Center beamlines at the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Energy Research, under Contract W-31-109-ENG-38.

[‡] The X-ray coordinates of the *o*-succinylbenzoate synthase from *Amycolatopsis* complexed with *o*-succinylbenzoate (1SJB), *N*-acetylmethionine (1SJA), *N*-succinylmethionine (1SJC), and *N*-succinylphenylglycine (1SJD) have been deposited in the Brookhaven Protein Data Bank.

* Correspondence may be addressed to either author. (I.R.) Department of Biochemistry, 433 Babcock Dr., Madison, WI 53706; phone (608) 262-0437; fax (608) 262-1319; e-mail Ivan_Rayment@biochem.wisc.edu. (J.A.G.) Department of Biochemistry, University of Illinois, 600 South Mathews Ave., Urbana, IL 61801; phone (217) 244-7414; fax (217) 244-6538; e-mail j-gerlt@uiuc.edu.

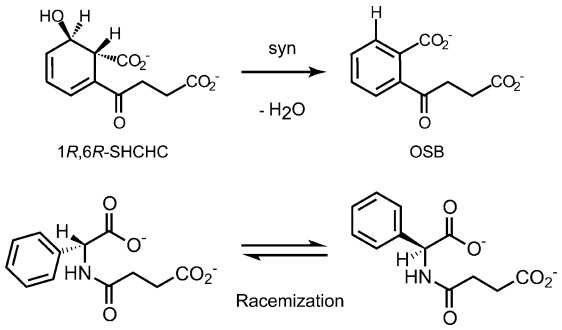
[§] University of Illinois.

[||] University of Wisconsin.

[1] Abbreviations: APS, Advanced Photon Source at Argonne National Laboratory; PEG, poly(ethylene glycol); NAAAR, *N*-acylamino acid racemase; MLE, muconate lactonizing enzyme; OSBS, *o*-succinylbenzoate synthase; OSB, 4-(2′-carboxyphenyl)-4-oxobutyrate or *o*-succinylbenzoate; SHCHC, 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate; SBC, The Structural Biology Center at the Advanced Photon Source, Argonne, IL; OSB, 4-(2′-carboxyphenyl)-4-oxobutyrate or *o*-succinylbenzoate; rms, root-mean-square; rmsd, root-mean-square deviation.

10.1021/bi0497897 CCC: $27.50
Published on Web 04/21/2004

Scheme 1: Reactions Catalyzed by the OSBS from *Amycolatopsis*



the active sites of these enzymes at the ends of the second and sixth β-strands in a $(\beta/\alpha)_7\beta$-barrel domain. In the OSBS reaction, Lys 163 at the end of the second β-strand initially functions as the base that abstracts the α-proton from the substrate and then as the acid that facilitates the departure of the β-hydroxyl group; Lys 263 at the end of the sixth β-strand likely stabilizes the enediolate intermediate. In the NAAAR reaction, Lys 163 and Lys 263 participate in a two-base-mediated 1,1-proton transfer reaction. This functional promiscuity is possible because the active sites can accommodate the substrates for both reactions, and the serendipitously positioned lysine residues can catalyze either reaction as dictated by the structure of the bound substrate.

In this paper, we report structural studies of the promiscuous OSBS/NAAAR in the presence of 4-(2′-carboxyphenyl)-4-oxobutyrate or *o*-succinylbenzoate (OSB), the product of the OSBS reaction, as well as three *N*-acylamino acid substrates for the NAAAR reaction, *N*-acetylmethionine, *N*-succinylmethionine, and *N*-succinylphenylglycine to 2.2, 2.3, 2.1, and 1.9 Å resolution, respectively. These structures reveal the molecular basis for the observed promiscuity and show how the active-site cavity can accommodate both the substrate for the OSBS reaction and the substrates for the accidental NAAAR reaction. They also show how this enzyme is able to utilize its diametrically opposed lysine residues (Lys 163 and Lys 263) for the different acid/base-catalyzed reactions. This suggests that one route for the evolution of the diverse reactions catalyzed by members of the enolase superfamily likely occurred via adventitious binding of the substrate for a different reaction in the active site of the progenitor followed by subsequent mutations that enhanced the promiscuous activity.

MATERIALS AND METHODS

Protein Expression and Purification of Wild-Type OSBS-Amycolatopsis. The 1.11 kb gene encoding the OSBS from *Amycolatopsis* sp. T-1-60 was cloned into the pET17b plasmid as described before (*2*). A frozen stock of BL-21 cells transformed with the gene encoding the OSBS in the pET17b vector was used to inoculate 5 mL of LB-Amp, which was shaken overnight at 37 °C. Two 2 L flasks of LB-Amp were inoculated from the overnight growth and shaken at 30 °C for 36 h. The suspension was centrifuged at 5000*g* for 10 min at 4 °C to harvest the cells. The cells were resuspended in ~200 mL of 20 mM Tris-HCl, pH 8.0, and then sonicated at 0 °C with a power of 6 for 5 s intervals followed by a 10 s delay, repeated 120 times. The cell debris was removed by centrifugation (15000*g* for 45 min at 4 °C). The supernatant was loaded onto a (2.6 × 70 cm) of DEAE-Sephacel resin (Cl^-) at 4 mL/min. The column was washed with 800 mL of 20 mM Tris-HCl, pH 8.0, and the protein was eluted with a linear gradient over 1800 mL of 0% to 100% 1 M NaCl in 20 mM Tris-HCl, pH 7.9. Fractions containing OSBS were identified by OD_{280} and SDS−PAGE, and those fractions were combined. The protein solution was then dialyzed into 20 mM Tris-HCl, pH 8.0, and the protein was then loaded in aliquots (no more than 100 mg per run) onto a 6 mL Resource-Q anion-exchange column (Cl^-) and eluted with a 240 mL linear 1 M NaCl gradient (15% to 50% 1 M NaCl in 20 mM Tris-HCl, pH 8.0). Fractions containing protein as assessed by SDS−PAGE and OD_{280} were combined and concentrated, and the mass (39 406.4 g/mol) was confirmed by ESI-MS to within experimental error (yield = 150 mg, assuming an extinction coefficient of 33 240 M^{-1} cm^{-1}, >98% pure).

Preparation of Se-Met OSBS. Selenomethionine-labeled protein was prepared in order to be able to solve the phase problem with multiple wavelength measurements at the selenium edge. A frozen stock of BL-21 cells containing the gene encoding the OSBS in the pET17b vector was used to inoculate minimal medium (2 L) containing $K_2HPO_4 \cdot 3H_2O$ (33 g), KH_2PO_4 (3 g), $(NH_4)_2SO_4$ (2 g), D-glucose (16 g), $MgSO_4 \cdot 7H_2O$ (0.4 g), $CaCl_2 \cdot 2H_2O$ (20 mg), $FeSO_4 \cdot 7H_2O$ (100 mg), thiamin (0.1 g), and ampicillin (0.2 g). The solution was shaken at 37 °C until $OD_{600} \approx 0.6$ AU (approximately 8−10 h). An aliquot was removed, and a methionine suppression solution (40 mL) containing L-Lys-HCl (200 mg), L-Thr-HCl (200 mg), L-Phe-HCl (200 mg), L-Leu-HCl (100 mg), L-Ile-HCl (100 mg), and L-Val-HCl (100 mg) was added. The flask was returned to 37 °C to shake for 30 min, after which time IPTG (238 mg) and SeMet (Acros; 100 mg) were added. The flask was allowed to shake at 30 °C for ~36 h. The suspension was centrifuged at 5000*g* for 10 min at 4 °C to harvest the cells, and the protein was then isolated and purified as described for the wild-type enzyme (yield = 80 mg assuming an extinction coefficient of 33 240 M^{-1} cm^{-1}, >98% pure). The extent of substitution was determined by ESI-MS, which showed that ~98% of the methionine residues in the protein contained seven selenium atoms.

Crystallization and Data Collection. A search for crystallization conditions at both room temperature was conducted by utilizing an in-house sparse matrix screen composed of 144 conditions and at 4 °C via the hanging drop method of vapor diffusion. The best crystals were observed growing at room temperature from poly(ethylene glycol) 8000 at pH 8.0. Large single crystals were subsequently obtained by macroseeding batch experiments (*5*) with final precipitant concentrations of 8−9% PEG-8000 buffered with 100 mM HEPPS (pH 8.0), 10 mM $MgCl_2$, 10 mM OSB or substrate, and enzyme concentrations of 7.0 mg/mL. Crystals achieved maximum dimensions of 1.2 mm × 0.7 mm × 0.6 mm in approximately 2−4 weeks. They belonged to the space group *R*32 with unit cell dimensions of $a = b = 216.0$ Å and $c = 261.0$ Å and contained four subunits in the asymmetric unit.

The crystals were harvested by first introducing a solution containing 12% PEG-8000, 12.5 mM $MgCl_2$, 200 mM NaCl, 125 mM KCl, and 25 mM OSB or *N*-acyl substrate directly into the batch experiments. Thereafter the crystals could be

Table 1: X-ray Data Collection Statistics

data set	beamline and date	no. of frames [width (deg)]	wavelength (Å)	resolution (Å)	indep reflections	completeness (%)	redundancy	avg I/avg $\sigma(I)$	R_{sym}[a]
MAD inflection	19-ID, 3-21-2001	225 (0.8)	0.97940	30.0−2.35	95666	99.4	8.4	34.9	7.2
				2.43−2.35[b]	9237	96.6	4.4	3.2	28.3
MAD remote	19-ID, 3-21-2001	225 (0.8)	1.02470	30.0−2.35	94977	98.5	8.0	34.2	6.4
				2.43−2.35	8691	90.7	3.4	2.6	29.1
MAD peak	19-ID, 3-21-2001	225 (0.8)	0.97930	30.0−2.35	95734	99.4	8.2	31.3	7.9
				2.43−2.35	9213	96.3	4.2	2.8	31.9
OSB	19-BM, 3-28-2001	350 (0.6)	0.97630	50.0−2.20	116834	98.7	5.8	19.0	6.1
				2.28−2.20	11288	95.7	4.1	1.7	38.7
N-acetyl-methionine	19-BM, 3-28-2001	300 (0.6)	0.97630	50.0−2.25	107728	99.2	7.9	28.1	5.2
				2.43−2.25	10378	96.4	4.2	1.8	36.4
N-succinyl-methionine	19-ID, 4-7-2003	250 (0.7) 90 (2.0)	0.65250	50.0−2.00	154522	99.6	9.9	37.6	8.0
				2.07−2.00	15313	99.4	5.3	3.0	38.7
N-succinyl-phenylglycine	19-ID, 4-7-2003	250 (0.7) 90 (2.0)	0.65250	50.0−1.87	186822	99.7	10.1	40.1	8.5
				1.94−1.87	18487	99.2	5.2	3.3	37.7

[a] $R_{sym} = (\Sigma|I - \bar{I}|/\Sigma I) \times 100$. [b] Statistics for the highest resolution bin.

Table 2: Relevant Refinement Statistics

complex	OSB	*N*-acetylmethionine	*N*-succinylmethionine	*N*-succinylphenylglycine
resolution limits (Å)	30.0−2.20	20.0−2.30	20.0−2.10	50.0−1.87
overall *R*-factor[a] (%)/no. of rflns	20.9/116371	21.5/100031	21.1/133316	20.2/186821
working *R*-factor (%)/no. ofrflns	20.4/104727	21.1/89902	20.8/120279	19.6/168140
free *R*-factor (%)/no. of rflns	27.2/11644	28.0/10006	27.4/13195	24.1/18681
no. of protein atoms	11 113[b]	11 131[c]	11 135	11 067
no. of heteroatoms	842[e]	573[d]	963[f]	1270[g]
bond lengths (Å)	0.014	0.012	0.013	0.012
bond angles (deg)	2.17	2.20	2.26	2.30
trigonal planes (Å)	0.008	0.006	0.006	0.008
general planes (Å)	0.012	0.016	0.012	0.013
torsional angles[h] (deg)	17.8	17.9	17.7	17.1
Ramachandran statistics				
core (%)	90.7	88.2	90.0	91.6
allowed (%)	9.0	11.1	9.6	8.1
generously allowed (%)	0.3	0.7	0.4	0.3

[a] *R*-factor $= (\Sigma|F_o - F_c|/\Sigma|F_o|) \times 100$ where F_o is the observed structure-factor amplitude and F_c is the calculated structure-factor amplitude. [b] These include multiple conformations for R204 in subunit A, D183 and E220 in subunit B, R10, E79, and E220 in subunit C, and R68 in subunit D. [c] These include multiple conformations for R68 and R204 in subunit A, R127 and R231 in subunit B, and R127 in subunit C. [d] These include 4 OSB, 4 magnesium ions, and 690 waters. [e] These include 4 *N*-acetylmethionines, 4 magnesium ions, and 522 waters. [f] These include 4 *N*-succinylmethionines, 4 magnesium ions, and 895 waters. [g] These include 4 *N*-succinylphenylglycines and 1162 waters. [h] The torsional angles were not restrained during the refinement.

transferred directly into the same solution. X-ray data were recorded from frozen crystals. Prior to freezing, the crystals were serially transferred into final cryosolution composed of 24% PEG-8000, 25 mM $MgCl_2$, 500 mM NaCl, 125 mM KCl, 25 mM substrate/inhibitor, and 15% ethylene glycol.

X-ray diffraction data were recorded by utilizing synchrotron radiation from the Structural Biology Center beamlines at the Advanced Photon Source, Argonne National Laboratory. For each complex, 180° of data were recorded with a frame width that varied from 0.6 to 0.8° depending on the mosaic spread of the crystals. The data were integrated and scaled with HKL2000 and Scalepack (*6*). Details of the data collection statistics are given in Table 1.

Structural Analysis. A three-wavelength MAD experiment was used to determine the protein structure of the complex with OSB. The positions of the selenium sites (27 of 28 possible) were determined with SOLVE (*7*). The positions, occupancies, and temperature factors of these sites were refined with SOLVE, yielding an overall figure-of-merit of 0.42. Manual inspection of the selenium sites revealed the positions of the four subunits. The relationship between the subunits was used for cyclical averaging with solvent flattening by use of DM in the CCP4 program package (*8*, *9*), giving a figure of merit of 0.92 for the averaged electron density map. One complete subunit was built into the averaged electron density map with TURBO (*10*). The model was then expanded back into the crystallographic cell and refined with TNT (*11*). Other structures were solved by Fourier difference analysis and refined in a similar manner. Least-squares refinement statistics for each of the models are presented in Table 2.

The quality of the models was assessed with the program Procheck (*12*). A Ramachandran showed that, for the OSB complex, 90.7% of nonglycine and nonproline residues fall in the most favorable region, 9.0% fall in the additionally allowed region, and 0.3% fall in the generously allowed region. The quality of the other models is listed in Table 2. A representative section of electron density is shown in Figure 1.

RESULTS AND DISCUSSION

Four structures of the promiscuous *o*-succinylbenzoate synthase from *Amycolatopsis* were determined. These structures include complexes with OSB, the product of the OSBS reaction, and three *N*-acylamino acids, *N*-acetylmethionine,

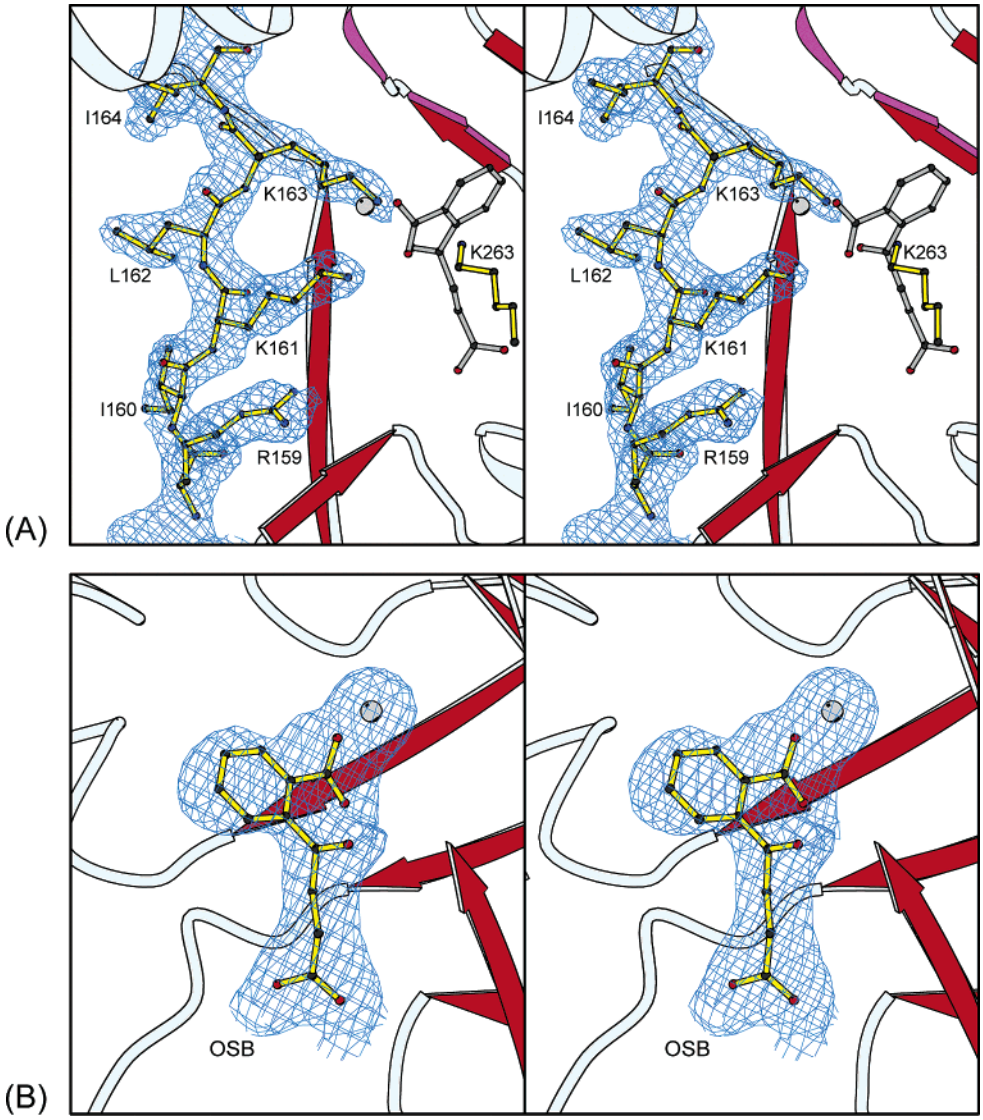


FIGURE 1: Stereoview of representative electron density from the complex with *o*-succinylbenzoate. In panel A a section of the protein electron density adjacent to the active site is depicted where the map was calculated with coefficients of the form $2F_o - F_c$ map and was contoured at 1σ. Panel B shows the omit electron density for OSB and its associated metal ion where these components were omitted from the refinement and phase calculation. This map was calculated with coefficients of the form $F_o - F_c$ map and was contoured at 3σ. The figure was prepared with the program Bobscript (*17*).

N-succinylmethionine, and *N*-succinylphenylglycine. As noted earlier, this OSBS is promiscuous and is able to catalyze both the OSBS reaction and the racemization of *N*-acylamino acid derivatives (*3*, *13*). Since the mechanism of the OSBS reaction is better understood than the racemization reaction, the structure of the product complex will be described first, followed by the three complexes with the *N*-acylamino acid derivatives.

Structure of the OSB Complex. The electron density for the OSB complex is essentially continuous for the entire length of the polypeptide chain, starting from Lys 2 and terminating at either Gly 367 or Ser 368 for all of the four subunits in the asymmetric unit. As expected, the structure consists of a large central domain composed of a $(\beta/\alpha)_7\beta$-barrel and a capping domain built from components from both the N- and C-termini (Figure 2).

The sequences of the $(\beta/\alpha)_7\beta$-barrel domains for the *Amycolatopsis* and *E. coli* OSBS share 27% identity and 42% similarity, but the sequences of the capping domain are more divergent (data not shown). Overall these proteins share 16% sequence identity. As expected, the topology of the OSBS from *Amycolatopsis* is essentially identical to that of the enzyme from *E. coli* (Figure 3) (*14*). Superposition of the structures of the OSB complexes of these enzymes confirms that the barrel domains are very similar. The rms difference between 95 structurally equivalent residues in the barrel is 1.3 Å, where the barrels extend from Arg 105 to Pro 290 in the enzyme from *E. coli* and from Pro 131 to Ser 318 in the enzyme from *Amycolatopsis*.

The OSBS from *Amycolatopsis* contains 368 amino acid residues and is somewhat larger than that from *E. coli*, which consists of 320 residues. Most of these additional residues are located in the capping domains with approximately 20 additional residues on either side of the central $(\beta/\alpha)_7\beta$-barrel domain. As a consequence, the structures of the capping domains are more divergent, although their overall topologies are very similar. The major differences arise in the lengths of the loops that connect the strands of sheet that dominate this domain. Many of these extensions contribute to the higher oligomeric state of the enzyme from *Amycolatopsis*.

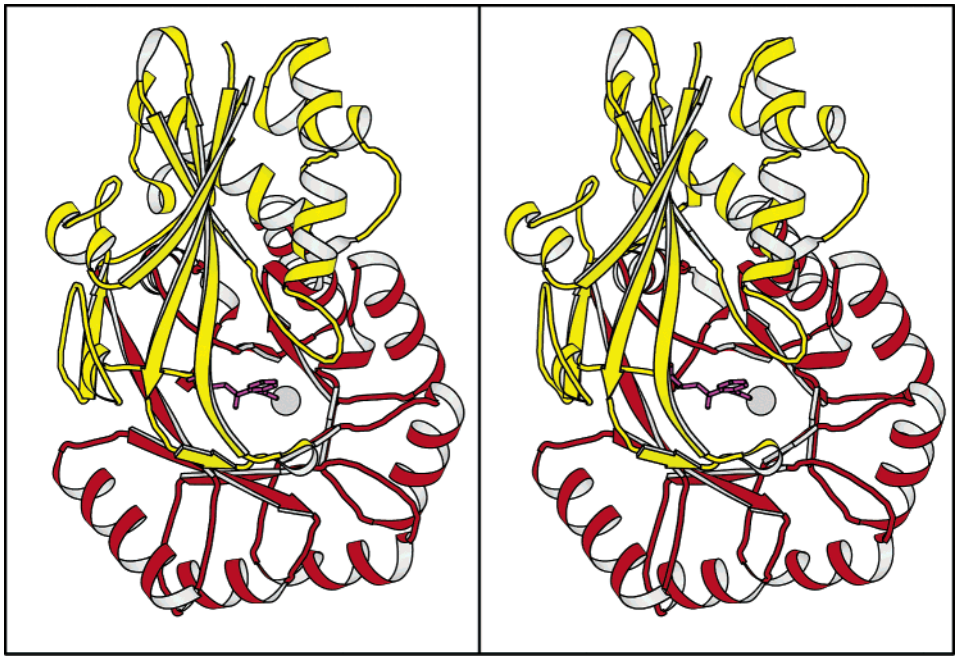


FIGURE 2: Stereo ribbon representation of a single subunit of OSBS from the complex with OSB. The capping domain is colored in yellow, whereas the $(\beta/\alpha)_7\beta$-barrel is depicted in red. The figure was prepared with the program Molscript (*18*).

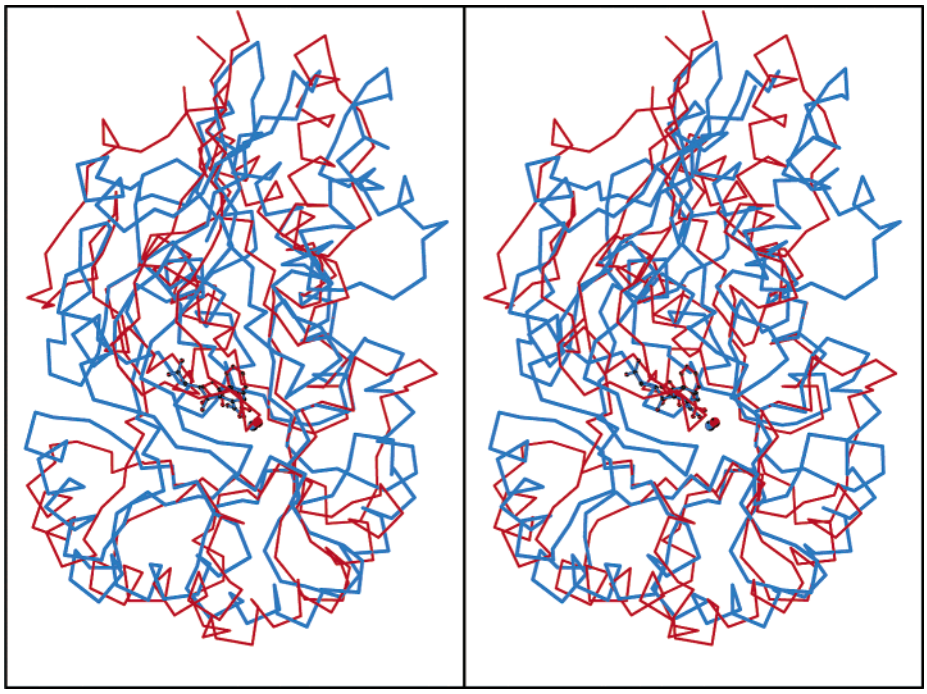


FIGURE 3: Stereo overlap of the α-carbons of the OSBS complexes from *Amycolatopsis* and *E. coli*. The enzyme from *Amycolatopsis* is depicted in blue whereas that from *E. coli* is colored in red. The coordinates for the product complex were taken from the RCSB with accession number 1FHV and superimposed with the program Align (*19*). The figure was prepared with the program Molscript (*18*).

In contrast to the monomeric enzyme from *E. coli*, the OSBS from *Amycolatopsis* assembles as an octamer with 422 symmetry (Figure 4). In the crystal lattice, the 4-fold axis of the octamer is coincident with a crystallographic 2-fold axis, which thus reduces the contents of the crystallographic asymmetric unit to four subunits. Extensive contacts exist between subunits in the octamer. Indeed, 2545 $Å^2$ of surface area is buried per subunit in the macromolecular assembly as measured with the program GETAREA (*15*). The insertions in the sequence between Ala 120 and Ser 129 and between Lys 85 and Lys 100, relative to the OSBS from *E. coli*, contribute significantly to the oligomer interface, although the extensions at the C-terminus starting at ~Val 333 are not involved in any protein−protein interfaces.

Comparison of the Binding of OSB in the OSBSs from Amycolatopsis and E. coli. The electron density for *o*-succinylbenzoate is well-defined in all subunits of the product complex of the OSBS from *Amycolatopsis* (Figure 1B). As expected from the previously determined structure of the OSBS from *E. coli* complexed with OSB, the active site contains a single magnesium ion that is coordinated to only one carboxyl oxygen atom of the product (*14*). In addition Nζ of the catalytic acid and base, Lys 163, lies 3.5 Å from

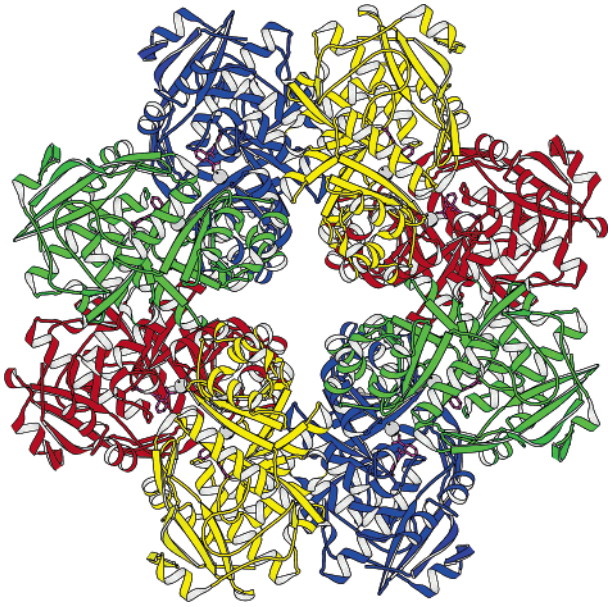


FIGURE 4: Ribbon representation of the complete octamer for OSBSfrom *Amycolatopsis* for the complex with OSB. View is looking down the noncrystallographic 4-fold axis, which is coincident with a crystallographic 2-fold axis. The active sites may be identified by the locations of the metal ions and product depicted in space-filling and ball-and-stick representations, respectively. The figure was prepared with the program Molscript (*18*).

C2 of OSB. Nζ of Lys 263 lies 3.3 Å from C2 of OSB on the opposite side of the product and is believed to provide electrostatic stabilization of the enediolate intermediate (*16*). The product as a whole is located in a large pocket created by the loops at the C-terminal ends of the β-strands of the TIM barrel. A mixture of hydrophobic and hydrophilic side chains form the surface of the pocket to complement the functionality of the substrate and product. The hydrophobic component is provided by Phe 19, Phe 23, Met 50, Tyr 55, Met 292, and Ile 293, whereas the succinyl carboxylate forms distant polar interactions with Arg 299 (mediated by water molecules) and the amide hydrogen of Ile 293. In addition, O3 of the succinyl moiety forms a hydrogen bond with Oγ of Ser 135.

The positions of the three carboxylate ligands for the metal ion, Asp 189, Glu 214, and Asp 239 (161, 190, and 213 in the enzyme from *E. coli*), are very similar in the OSB complexes of both enzymes. Likewise, the position of the catalytic acid/ base for the OSBS reaction, Lys 163 (133 in the enzyme from *E. coli*), is very similar in both enzymes, as is the location of the lysine on the opposite side of the active site, Lys 263 (163 in the enzyme from *E. coli*), which likely provides electrostatic stabilization of the enediolate intermediate (*3*, *16*). Overall, these five functional residues may be superimposed with an rms difference of only 0.32 Å.

Surprisingly, the orientation of OSB in the binding pocket is quite different in the two complexes (Figure 5). When the active sites are aligned on the basis of the secondary structural elements of the barrel, the metal binding ligands align reasonably well; however, the benzene rings of the OSBs are displaced from each other by an average distance of approximately 2.0 Å and are inclined to each other with an angle of ~45°. Furthermore, the orientation of the succinyl component of OSB is different in the structures (Figure 5A). In the case of the complex of the enzyme from *E. coli*, the succinyl carboxylate is directed toward the interior of the barrel, whereas in the complex with the enzyme from *Amycolatopsis* the carboxylate is directed away from the barrel axis. This latter orientation for the succinyl moiety is possible because the final β strand of the barrel (β8) is considerably shorter than that in the enzyme from *E. coli*. Furthermore, in the OSBS from *Amycolatopsis* the loop that connects β8 to the capping domain is shifted away from the barrel axis to create a cavity that can accept the succinyl side chain (Figure 5B). This difference in structure may contribute to the ability of the enzyme from *Amycolatopsis* to bind the *N*-acylamino acid substrates for the racemase reaction.

Structures of the OSBS from Amycolatopsis with Three N-Acylamino Acid Substrates. The structures of complexes of the OSBS from *Amycolatopsis* with *N*-acetylmethionine, *N*-succinylmethionine, and *N*-succinylphenylglycine have been determined to 2.3, 2.1, and 1.9 Å resolution, respectively. This series of substrates was chosen because they show increasing levels of activity in the racemase reaction (range of 3000-fold), with the rate for *N*-succinylphenylglycine approaching that of the OSBS reaction (*3*). The electron density for each of the substrates is unequivocal (Figure 6) and reveals how the active site can accommodate each of these molecules. The active site was expected to contain a mixture of enantiomers for each compound, although an equal population of enantiomers was not expected. Thus, each substrate was built as a single enantiomer that best matched the electron density. It was observed that the predominant enantiomers in the active site were *N*-acetyl-*(S)*-methionine, *N*-succinyl-(*R*)-methionine, and *N*-succinyl-(*R*)-phenylglycine.

As seen in Figure 7A, all three *N*-acylated amino acids bind in the active site in a manner that generally overlaps with OSB; however, the positions of the α-carbons vary. The location of that in *N*-succinylphenylglycine is anomalous relative to the *N*-acylmethionine substrates and will be discussed separately. As a general feature, the succinyl carboxylate moieties of *N*-succinylmethionine and *N*-succinylphenylglycine lie in the same location as that occupied by the equivalent carboxylate in OSB.

Structure of the N-Acetylmethionine Complex. As shown in Figure 7B, the *N*-acetylmethionine moiety lies sandwiched between lysines 163 and 263, which function as the catalytic bases for the abstraction of the α-proton in the (*R*)- and (*S*)-racemization reactions, respectively (*3*). In this complex the α-carbon is located 3.4 and 3.2 Å away from Lys 163 and Lys 263, respectively. The overall conformation of the polypeptide chain that surrounds the active site is essentially identical to that seen in the complex with OSB. Indeed, within experimental error the metal coordination site is identical in *N*-acetylmethionine and OSB complexes. As was observed in the complex with OSB, the amino acid carboxylate forms a monodentate interaction with the magnesium ion that will be discussed in more detail later.

The remainder of the *N*-acetylmethionine lies within the cavity that binds OSB. The methionine side chain is encompassed by the same groups required to bind the phenyl ring of OSB, including Phe 23, Met 50, Tyr 55, and Met 292. The orientation of the *N*-acetyl moiety is somewhat variable among the four molecules in the crystallographic

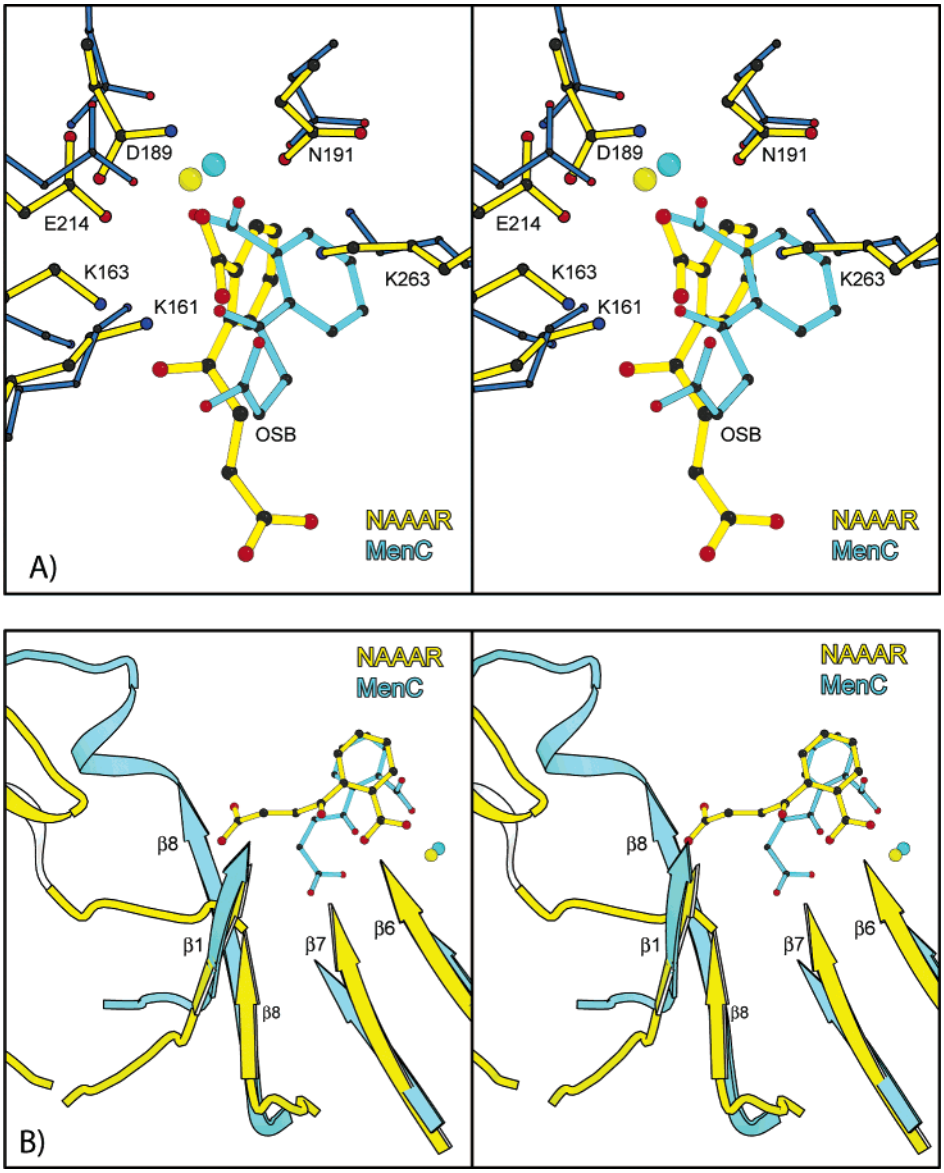


FIGURE 5: Stereo comparisons of the binding of OSB to the OSBSs from *Amycolatopsis* and *E. coli* (MenC), which reveal the differences in the coordination of the succinyl moiety in the active sites of the respective enzymes. Panel A shows a closeup view of the metal binding ligands and the active-site lysine residues. Panel B shows the difference in the position of the eighth β-strand that allows the succinyl moiety to adopt an extended conformation in the enzyme from *Amycolatopsis*. The enzyme from *Amycolatopsis* is depicted in yellow whereas that from *E. coli* is drawn in blue. The figure was prepared with the program Bobscript (*17*). The coordinates for the *E. coli* OSB complex were taken from the RCSB with accession number 1FHV (*14*).

asymmetric unit, but in no case does either the amide nitrogen or carbonyl oxygen form a hydrogen-bonding interaction with the polypeptide chain. Rather, this moiety extends into the cavity that contains the succinyl moieties of OSB and the *N*-succinyl substrates.

Structure of the N-Succinyl-(R)-Methionine Complex. Although both *N*-acetylmethionine and *N*-succinylmethionine bind to the active site in approximately the same location, the positions of the α-carbons differ even though the overall conformation of the polypeptide chain that surrounds the active site is essentially identical in both complexes. These differences allow better coordination of the *N*-succinyl moiety as shown in Figure 8A. In this complex the amide hydrogen is hydrogen-bonded to the carbonyl oxygen of Gly 291, whereas the succinyl carbonyl oxygen is hydrogen-bonded to Oγ of Ser 235. The α-carbon is located 2.9 and 3.5 Å away from Lys 163 and Lys 263, respectively, where the carbon is located closest to the *R*-specific base as expected for the *R*-enantiomer observed in the active site.

The succinyl carboxylate moiety of *N*-succinylmethionine lies in the same location occupied by the equivalent carboxylate in OSB (Figures 7A and 8A) and forms hydrogen bonds to two well-defined water molecules. One of these water molecule is hydrogen-bonded to Oϵ2 of Glu 294 and NH2 of Arg 299, whereas the second is hydrogen-bonded to NH1 of Arg 299. The polar interactions of the amide linkage and succinyl carboxylate provide an explanation for the lower values of K_M for the *N*-succinyl amino acids relative to the shorter *N*-acetylmethionine that lacks the water-mediated interaction with Arg 299 (*3*).

Interestingly the metal coordination in both the *N*-acetylmethionine and *N*-succinylmethionine complexes is monodentate as is observed for OSB complex. This is in contrast to the coordination of SHC observed in the K133R

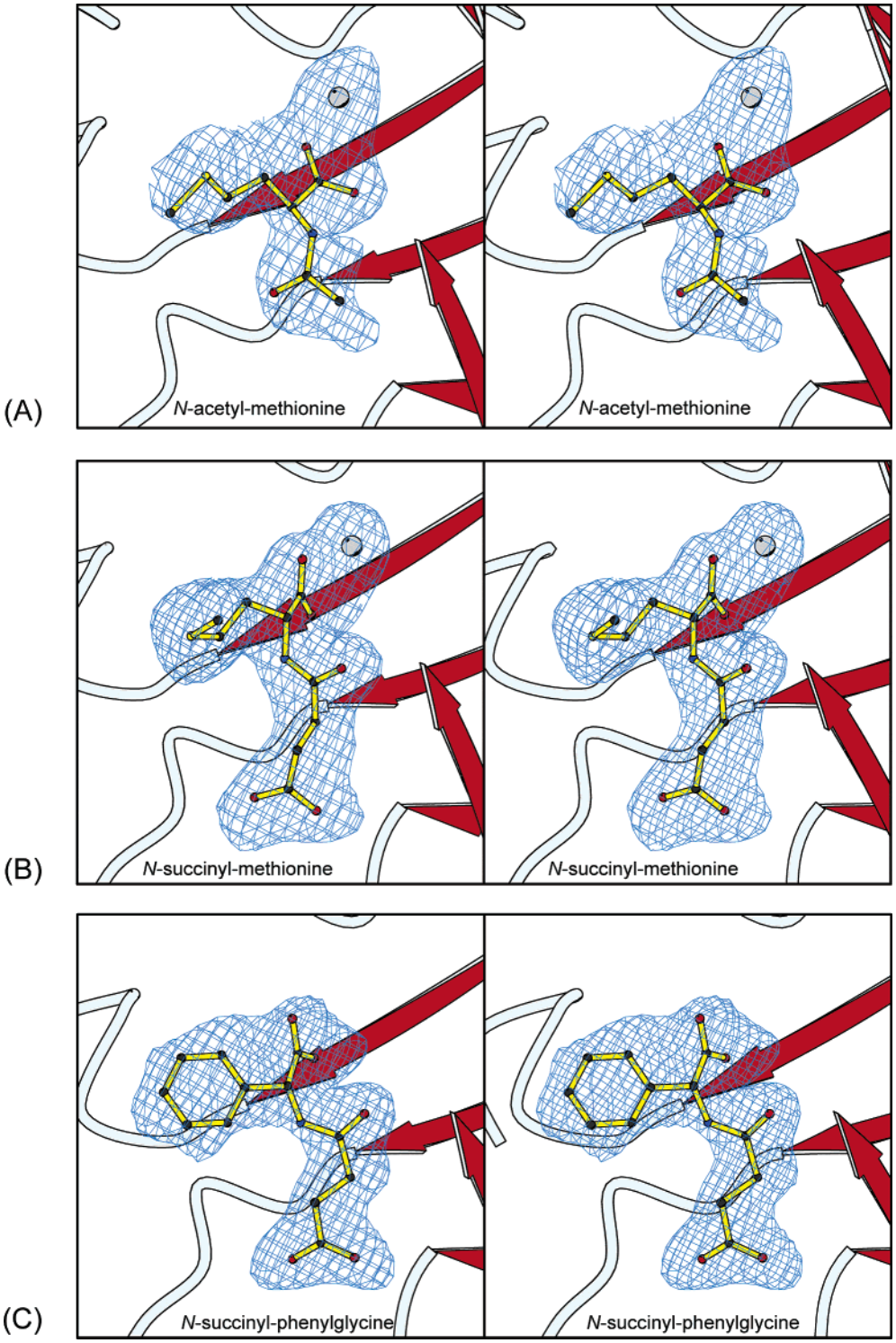


FIGURE 6: Stereoview of the omit electron density for the *N*-acyl substrates. Panels A−C show the density for *N*-acetylmethionine, *N*-succinylmethionine, and *N*-succinylphenylglycine, respectively. The maps were calculated with coefficients of the form $F_o - F_c$ map and were contoured at 3σ where the substrate and metal ion were omitted from the refinement and phase calculation The figures were prepared with the program Bobscript (*17*).

mutant of the OSBS from *E. coli* (*16*), where the coordination is clearly bidentate. The present structures are expected to be those of enzyme−substrate complexes; however, in all likelihood the coordination at the enediolate intermediate will be bidentate because this geometry would lead to greater electrostatic stabilization. If this assumption is correct, the substrate must move as the α-proton is abstracted, as was also suggested from the structure of SHCHC bound to the K133R mutant of the *E. coli* protein (*16*).

Structure of the N-Succinylphenylglycine Complex. The structure of the *N*-succinylphenylglycine complex is anomalous because the substrate is displaced from the positions observed in the complexes with OSB and both *N*-acylated methionines (Figures 7A and 8B). In addition, no metal ion is observed in this complex, even though it was included in the crystallization mixture at the same concentration as in the other complexes. This coordination is seen in all four subunits in the asymmetric unit. This is ironic because

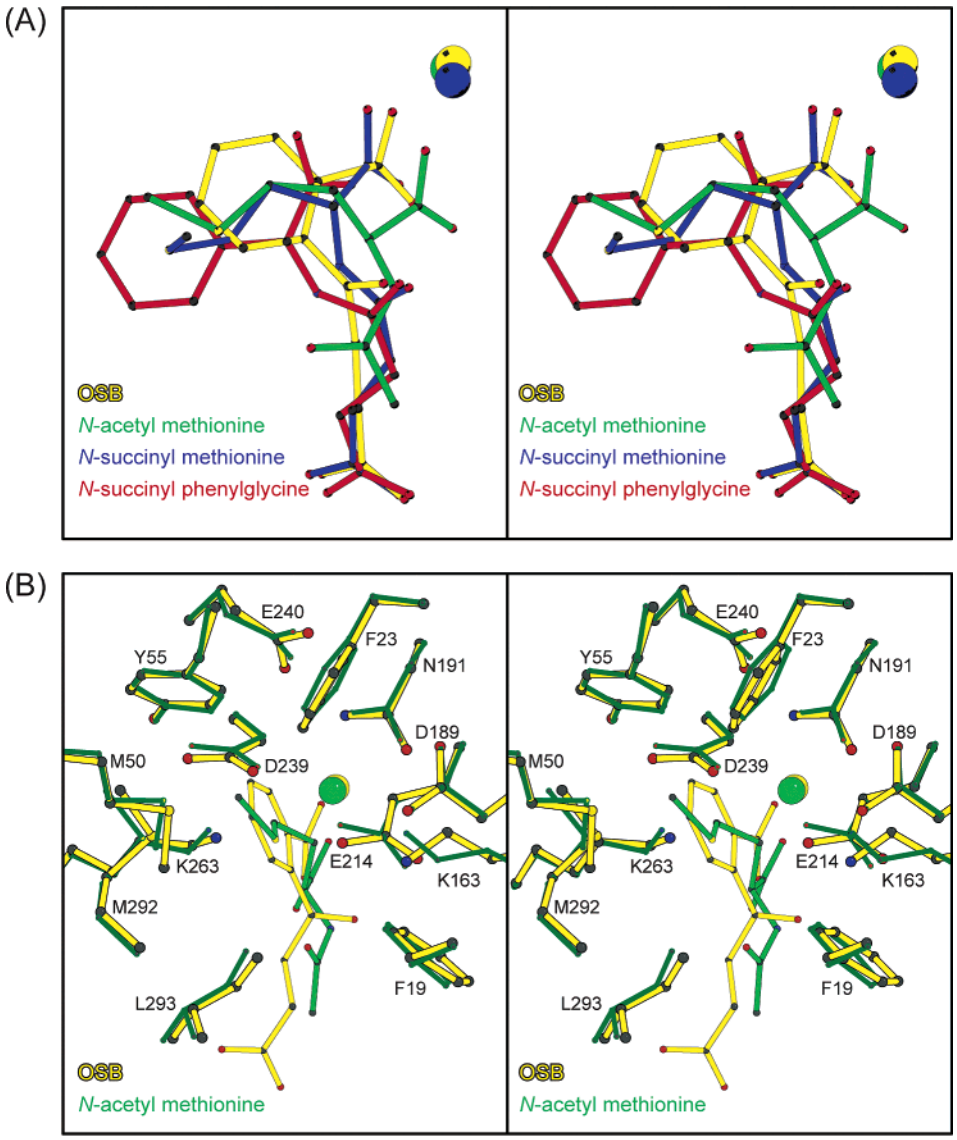


FIGURE 7: Stereo superposition of (A) OSB, *N*-acetylmethionine, *N*-succinylmethionine, and *N*-succinylphenylglycine and (B) the active-site residues of the OSB and *N*-acetylmethionine complexes. In these figures OSB, *N*-acetylmethionine, *N*-succinylmethionine, and *N*-succinylphenylglycine are depicted in yellow, green, blue, and red, respectively. The coordinates for the entire protein were aligned starting from the active-site residues utilizing the program Align (*19*).

N-succinylphenylglycine is a better substrate in the manganese-dependent racemase reaction than either *N*-acetylmethionine or *N*-succinylmethionine (*3*). There is no obvious explanation for the different conformation or absence of metal ion; however, it has previously been shown that, in the absence of substrate, OSBS from *E. coli* does not bind metal ions tightly (*14*). Thus since the substrate does not bind with its carboxylate in the correct place it is not surprising that no metal ion is observed. As such this structure should be viewed as a nonproductive complex that illustrates range of ligand binding conformations that are available to this active site.

As can be seen in Figure 8B, with the exception of Met 50, all of the side chains that constitute the active site are similar in the complexes with both *N*-succinylphenylglycine and *N*-succinylmethionine. In particular, the coordination of the succinyl moiety is identical in both complexes, thereby providing evidence for potential plasticity in substrate binding.

Comparison of N-Acylamino Acid Complexes with the OSBS from E. coli. Overlap of the structures of the complexes of the *N*-acyl substrates on the structure of the OSBS from *E. coli* complexed with either OSB (data not shown) or SHCHC (Figure 9) may explain why the enzyme from *Amycolatopsis* has promiscuous racemase activity whereas the enzyme from *E. coli* does not. Clearly both active sites can accommodate the metal−liganding carboxylate group and α-carbon. Likewise, they both have catalytic acids and bases in appropriate locations to perform a racemization reaction. The hydrophobic pocket in the enzyme from *Amycolatopsis* appears somewhat less restricted than that of the enzyme from *E. coli* due to the location and orientation of the loop between Pro 18 and Ser 27 (Gly 16−Lys 25 in the enzyme from *E. coli*). However, this difference should not prevent the acceptance of smaller amino acid side chains by the latter enzyme.

The most significant difference between these two enzymes resides in the manner in which the enzyme from *Amycolatopsis* binds the amide bond of the amino acid substrate and accommodates the *N*-acyl moiety (Figure 8A). As noted earlier, the conformation of the succinyl moiety in OSB is different in the complexes with the two enzymes

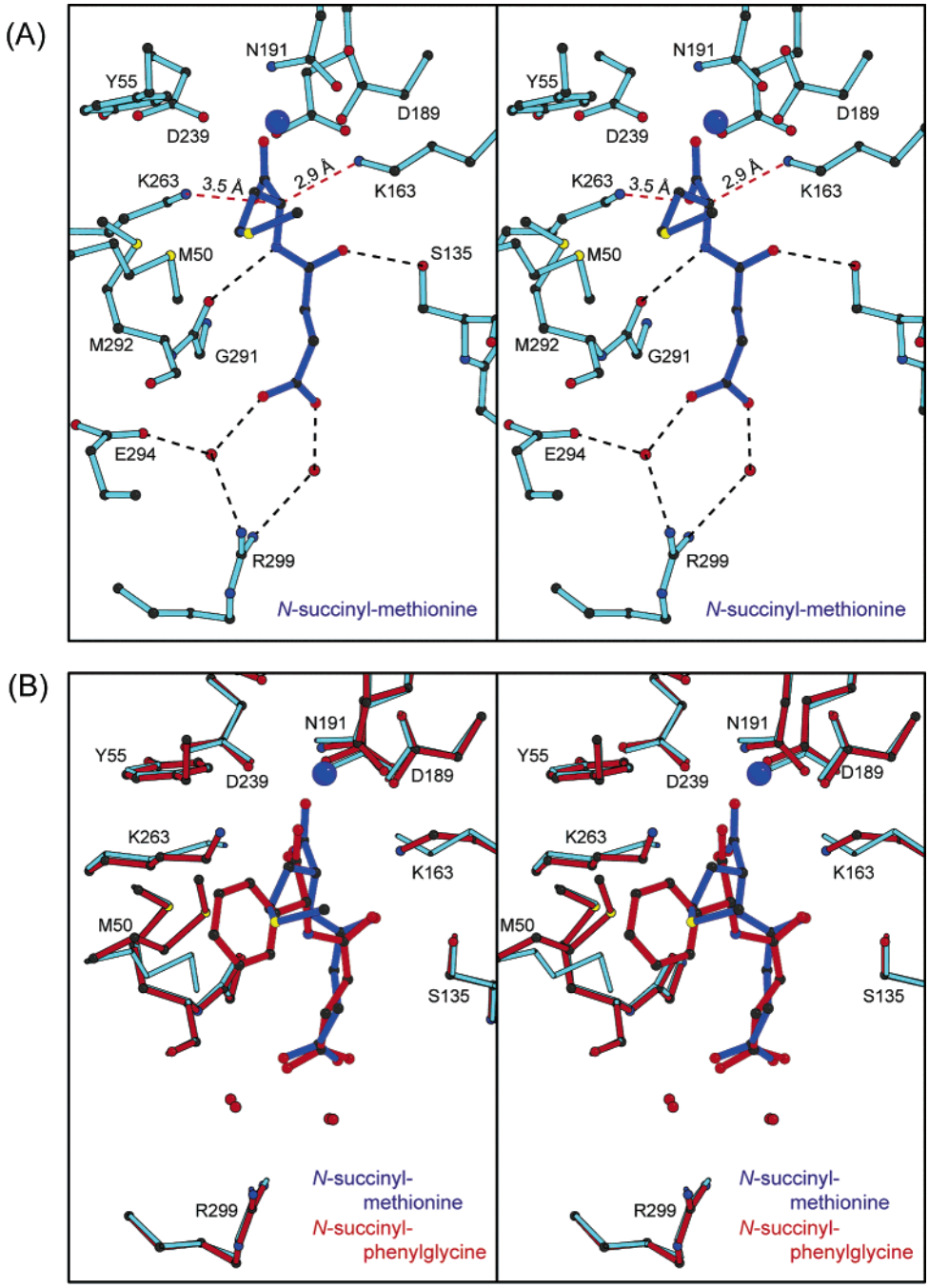


FIGURE 8: Stereoviews of (A) the *N*-succinylmethionine complex and (B) the superposition of the *N*-succinylmethionine and *N*-succinylphenylglycine complexes. Panel A shows the interaction distances between *N*-succinylmethionine and the ligand binding site. Panel B shows that *N*-succinylphenylglycine binds anomalously in the active site without inducing any major changes in conformation of the side chains. This demonstrates the ability of the active site to coordinate a variety of ligands. *N*-Succinylmethionine and *N*-succinylphenylglycine are depicted in red and blue, respectively. The figure was prepared with the program Bobscript (*17*).

where the succinyl moiety is in an extended conformation in the enzyme from *Amycolatopsis*. This extended conformation is more consistent with the planar conformation of the amide linkage observed in the *N*-acylmethionine substrates. In addition, at least for the *N*-acylsuccinyl substrates, it provides hydrogen-bonding partners for the amide hydrogen and carbonyl oxygen on opposite sides of the ligand-binding pocket. Of these, the interaction between Oγ of Ser 135 and the carbonyl oxygen is particularly notable. This type of interaction is not available in the enzyme from *E. coli* because the equivalent side chain is Leu 109. In principle, the interaction with the amide hydrogen provided by the carbonyl oxygen of Gly 291 in the enzyme from *Amycolatopsis* could be provided by the main-chain oxygen of Ser 263 in the enzyme from *E. coli*; however, the conformation and position of the backbone in the latter protein would force the substrate carbonyl oxygen into the hydrophobic environment provided by Leu 109. Furthermore, the main-chain atoms surrounding Ser 263 in the enzyme from *E. coli* lie ~1.8 Å away from the equivalent atoms around Gly 291 in the enzyme from *Amycolatopsis*. This implies that the active site in the enzyme from *E. coli* is somewhat wider than that of the enzyme from *Amycolatopsis*, which imposes more difficulty in coordinating an amide linkage of an *N*-acylamino acid by ligands on opposite side of the active-site pocket.

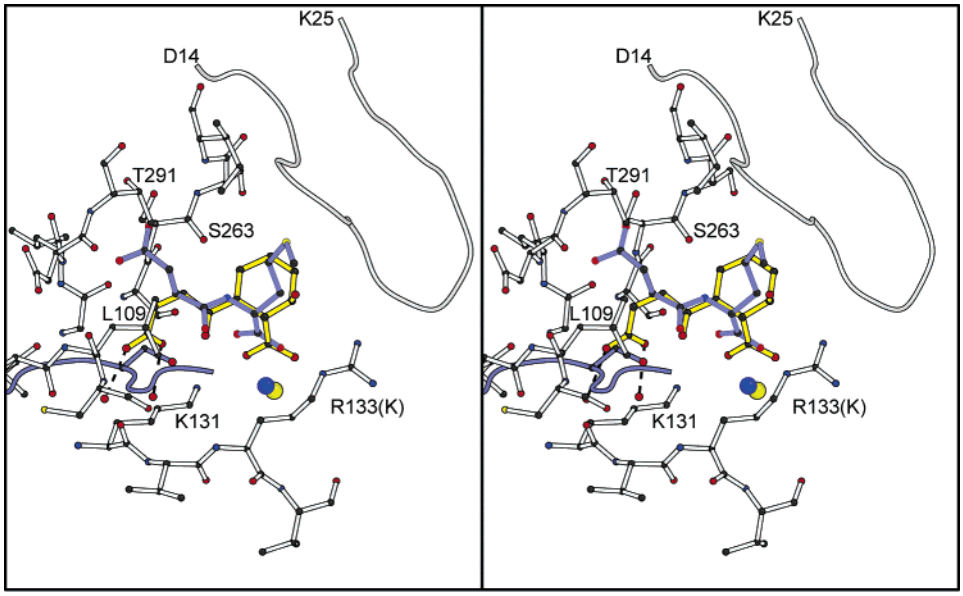


FIGURE 9: Stereoview of the superposition of the *N*-succinylmethionine from the *Amycolatopsis* complex on the structure of the 2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate (SHCHC) complex with the OSBS from *E. coli*. The *N*-succinylmethionine and its associated Ser 135 are depicted in blue, whereas the protein from *E. coli* and SHCHC are colored in yellow and white, respectively. The latter was obtained as a complex with the K133R mutant in which the catalytic base was eliminated that prevented the reaction from occurring. This reveals that the enzyme from *E. coli* should be able to readily accommodate most of the components of a *N*-acyl substrate, except for the carbonyl oxygen of the amide linkage. The coordinates for the *E. coli* complex were taken from the RCSB with accession number 1R6W (*16*).

Finally, it should be noted that although the protein from *E. coli* does not exhibit racemase activity against *N*-acylmethionine substrates, it may be able to racemize other, as yet unidentified, compounds that are more complementary with the binding pocket.

CONCLUSIONS

New enzymatic functions likely evolve most readily from a precursor that shares some functional characteristic with the new activity. Furthermore, the progenitor may have some basal level of the new activity prior to the gene duplication event, which facilitates improvement of function when selective pressure is applied. In the case of the OSBS from *Amycolatopsis*, this enzyme does possess two activities, although it is unknown whether the natural OSBS activity or the promiscuous NAAAR activity came first. Perhaps other orthologous OSBSs exhibit promiscuous behavior, although this has not yet been evaluated. The availability of active-site structures for the OSBSs from both *E. coli* and *Amyolatopsis* may allow identification of other substrates that will be racemized by one or both of these enzymes.

Comparison of the structures of the enzymes from *Amycolatopsis* and *E. coli* reveal different structural solutions for binding the substrate and product for the OSBS reaction. For example, the conformation of the succinyl moiety is not chemically coupled to the dehydration reaction so it is not forced to adopt a unique conformation. In addition, the identities of the amino acid residues that form the hydrophobic pocket for the hydrophobic regions of the substrate and product for the OSBS reaction are not conserved. We hypothesize that divergent evolution of OSBSs from different immediate ancestors likely provides opportunities for the coexistence or appearance of promiscuous reactions.

The challenge now is to discover, or even predict, whether other OSBSs or other members of the mechanistically diverse enolase superfamily that likely arose by divergent evolutions from a limited number of progenitors catalyze adventitous reactions that utilize conserved active-site functional groups to catalyze acid/base reactions with natural, but nonphysiological, or even unnatural carboxylate substrates.

ACKNOWLEDGMENT

We thank Drs. Randy Akire and Norma Duke of the Structural Biology Center, Argonne National Laboratory, for outstanding assistance in recording the data reported in this study.

REFERENCES

1. O'Brien, P. J., and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities, *Chem. Biol. 6*, R91−R105.
2. 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.
3. 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.
4. 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*, 16489−16501.
5. Rayment, I. (2002) Small-scale batch crystallization of proteins revisited. An underutilized way to grow large protein crystals, *Structure (Cambridge) 10*, 147−151.
6. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
7. Terwilliger, T. C. (2002) Automated structure solution, density modification and model building, *Acta Crystallogr. D: Biol. Crystallogr. 58*, 1937−1940.

8. Cowtan, K., and Main, P. (1998) Miscellaneous algorithms for density modification, *Acta Crystallogr. D54*, 487−493.
9. CCP4 (1994) The CCP4 Suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
10. Roussel, A., and Cambillau, C. (1991) Turbo Frodo, in *Silicon Graphics Geometry Partners Directory*, Silicon Graphics, Mountain View, CA.
11. Tronrud, D. E. (1997) TNT refinement package, *Methods Enzymol. 277*, 306−319.
12. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
13. Tokuyama, S., and Hatano, K. (1995) Purification and properties of thermostable N-acylamino acid racemase from *Amycolatopsis* sp. TS-1−60, *Appl. Microbiol. Biotechnol. 42*, 853−859.
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*, 10662−10676.
15. Fraczkiewicz, R., and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules, *J. Comput. Chem. 19*, 319−333.
16. 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*, 14427−14433.
17. Esnouf, R. M. (1999) Further additions to MolScript version 1.4, including reading and contouring of electron-density maps, *Acta Crystallogr. D55*, 938−940.
18. Kraulis, P. J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures, *J. Appl. Crystallogr. 24*, 946−950.
19. Cohen, G. H. (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions, *J. Appl. Crystallogr. 30*, 1160−1161.

BI0497897