

# Crystal Structure of the Pantothenate Synthetase from *Mycobacterium tuberculosis*, Snapshots of the Enzyme in Action<sup>†,‡</sup>

Shuishu Wang<sup>\*,§</sup> and David Eisenberg<sup>||</sup>

Public Health Research Institute, 225 Warren Street, Newark, New Jersey 07103, and Howard Hughes Medical Institute, UCLA-DOE Institute of Genomics and Proteomics, Molecular Biology Institute, University of California, Box 951570, Los Angeles, California 90095-1570

Received September 14, 2005; Revised Manuscript Received November 10, 2005

**ABSTRACT:** Pantothenate synthetase (PS) from *Mycobacterium tuberculosis* represents a potential target for antituberculosis drugs. PS catalyzes the ATP-dependent condensation of pantoate and  $\beta$ -alanine to form pantothenate. Previously, we determined the crystal structure of PS from *M. tuberculosis* and its complexes with AMPCPP, pantoate, and pantoyl adenylate. Here, we describe the crystal structure of this enzyme complexed with AMP and its last substrate,  $\beta$ -alanine, and show that the phosphate group of AMP serves as an anchor for the binding of  $\beta$ -alanine. This structure confirms that binding of  $\beta$ -alanine in the active site cavity can occur only after formation of the pantoyl adenylate intermediate. A new crystal form was also obtained; it displays the flexible wall of the active site cavity in a conformation incapable of binding pantoate. Soaking of this crystal form with ATP and pantoate gives a fully occupied complex of PS with ATP. Crystal structures of these complexes with substrates, the reaction intermediate, and the reaction product AMP provide a step-by-step view of the PS-catalyzed reaction. A detailed reaction mechanism and its implications for inhibitor design are discussed.

Pantothenate (vitamin B<sub>5</sub>) is an essential precursor for the biosynthesis of coenzyme A and acyl carrier proteins, both of which play critical roles in many cellular processes, including energy metabolism and fatty acid metabolism (1). Microorganisms and plants can synthesize pantothenate, while animals obtain this essential nutrient from their diet (2). Therefore, the pantothenate biosynthetic pathway offers targets for developing drugs against microbial pathogens. Recently, Jacobs and co-authors (3) reported that a *Mycobacterium tuberculosis* (MTB)<sup>1</sup> mutant defective in the de novo biosynthesis of pantothenate is highly attenuated in both immunocompromised and immunocompetent mice. This observation indicates that a functional pantothenate biosynthetic pathway is essential for virulence of MTB.

The pantothenate biosynthetic pathway comprises four steps catalyzed by enzymes encoded by the *panB*, *panC*, *panD*, and *panE* genes (4). The *panC* gene encodes a pantothenate synthetase (PS), which catalyzes the last step of pantothenate biosynthesis, the ATP-dependent condensation of pantoate and  $\beta$ -alanine to form pantothenate. The

*panC* gene product in several organisms has been identified and characterized, including those in *Escherichia coli* (5), *Saccharomyces cerevisiae*, higher plants *Oryza sativa* and *Lotus japonicus* (6), fungus *Fusarium oxysporum* (7), and *M. tuberculosis* (8). The PS enzymes from *E. coli*, higher plants, and MTB form dimers in solution.

The *E. coli* PS enzyme structure belongs to the cytidyltransferase superfamily (5). It has two distinct domains, a large N-terminal domain having a Rossmann fold and a smaller C-terminal domain containing a helical layer above a three-stranded antiparallel  $\beta$ -sheet. On the basis of a structural comparison of the *E. coli* PS with other members of cytidyltransferase superfamily having known structures, von Delft et al. (5) deduced the ATP and pantoate binding sites of the *E. coli* PS and proposed a hinged domain mechanism for opening and closing of the enzyme active site cavity.

We have determined the crystal structure of the MTB PS enzyme and its complexes with a nonhydrolyzable ATP analogue (AMPCPP), pantoate, and a pantoyl adenylate reaction intermediate (9). The MTB PS has the same fold as the *E. coli* enzyme, and the structures of substrate complexes confirm the binding sites for ATP and pantoate predicted from the *E. coli* PS structure. However, in the MTB PS structure, the domains of each subunit have a closed conformation, in contrast to those of the *E. coli* PS structure (5), and there is no significant movement between domains among all structures of the apoenzyme and various complexes. A flexible region, which forms a wall of the active site cavity, becomes ordered in the reaction intermediate complex and closes the active site cavity, thus acting as a gate to the active site cavity of the MTB PS enzyme.

<sup>†</sup> This work was supported by NIH Grant GM 62410 to D.E. D.E. is an Investigator of the Howard Hughes Medical Institute.

<sup>‡</sup> Coordinates and observed structure factor amplitudes for the structures described in this paper have been deposited in the Protein Data Bank (entries 2A88, 2A84, 2A86, and 2A7X).

\* To whom correspondence should be addressed. Phone: (973) 854-3470. Fax: (973) 854-3101. E-mail: shuishu@phri.org.

<sup>§</sup> Public Health Research Institute.

<sup>||</sup> University of California.

<sup>1</sup> Abbreviations: MTB, *Mycobacterium tuberculosis*; PS, pantothenate synthetase; AMPCPP,  $\alpha,\beta$ -methylene adenosine 5'-triphosphate; AMP, adenosine 5'-monophosphate; ATP, adenosine 5'-triphosphate; IPTG, isopropyl  $\beta$ -D-1-thiogalactopyranoside; NCS, noncrystallographic symmetry; PMSF, phenylmethanesulfonyl fluoride.

Table 1: X-ray Diffraction Data for *M. tuberculosis* Pantothenate Synthetase and Complexes<sup>a</sup>

|                                     | apo (1)       | AMP (2)       | ATP (3)       | $\beta$ -alanine (4) |
|-------------------------------------|---------------|---------------|---------------|----------------------|
| space group                         | C2            | C2            | C2            | P2 <sub>1</sub>      |
| resolution (Å)                      | 50–1.7        | 50–1.7        | 50–1.55       | 50–1.85              |
| $R_{\text{merge}}^{b,c}$ (last bin) | 0.059 (0.372) | 0.039 (0.220) | 0.060 (0.343) | 0.083 (0.547)        |
| completeness (%) <sup>b</sup>       | 97.3 (88.8)   | 88.0 (47.0)   | 87.6 (30.8)   | 99.8 (98.1)          |
| $I/\sigma^b$                        | 21.3 (2.8)    | 33.8 (5.2)    | 30.7 (3.9)    | 23.6 (2.5)           |
| redundancy <sup>b</sup>             | 3.3 (2.8)     | 4.2 (3.1)     | 6.6 (3.5)     | 2.7 (2.2)            |
| cell dimensions                     |               |               |               |                      |
| <i>a</i> (Å)                        | 119.82        | 120.25        | 118.95        | 47.99                |
| <i>b</i> (Å)                        | 44.98         | 44.94         | 44.88         | 70.64                |
| <i>c</i> (Å)                        | 82.14         | 82.23         | 82.13         | 81.68                |
| $\beta$ (deg)                       | 125.50        | 125.56        | 125.26        | 99.21                |

<sup>a</sup> All data collected at the cryogenic temperature of 100 K. <sup>b</sup> The numbers in parentheses are for the last bin of data, which is from 1.76 to 1.7 Å, from 1.76 to 1.7 Å, from 1.61 to 1.55 Å, and from 1.90 to 1.85 Å for data sets 1–4, respectively. <sup>c</sup>  $R_{\text{merge}} = \sum |I_{hkl} - \langle I_{hkl} \rangle| / \sum I_{hkl}$ , where  $\langle I_{hkl} \rangle$  is the average of  $I_{hkl}$  over all symmetry equivalents.

Kinetic analysis of the MTB PS suggests that the enzyme-catalyzed reaction proceeds through two steps: the formation of an enzyme-bound intermediate, pantoyl adenylate, from ATP and pantoate, followed by nucleophilic attack on the intermediate by  $\beta$ -alanine to form pantothenate and AMP (8). This two-step mechanism is supported by the structure of the reaction intermediate complex, which shows that pantoyl adenylate can exist stably in the enzyme active site. Structures of complexes with AMPCPP, pantoate, and pantoyl adenylate give a direct view of the first step of the PS-catalyzed reaction. Some important active site residues that were suggested by the crystal structures have been studied recently by site-directed mutagenesis and kinetic analysis, and their roles in substrate binding and catalysis have been confirmed (10). However, the mechanism of the second half of the reaction has remained elusive, because of the lack of knowledge of binding interactions of the last substrate,  $\beta$ -alanine.

In this work, we determined a structure for MTB PS complexed with  $\beta$ -alanine. This structure gave direct evidence that the binding site for this last substrate in the pathway exists only after formation of the pantoyl adenylate intermediate. We also obtained a new crystal form, in which the flexible region (gate to the active site) is locked in an inactive form. An ATP complex structure in this crystal form confirmed ATP binding interactions deduced from the AMPCPP complex and the function of the His47 side chain. Together, these crystal structures lead to a detailed view of the overall reaction mechanism. A model of the tetrahedral intermediate of the second half of the reaction was constructed from the  $\beta$ -alanine complex. This tetrahedral intermediate provides a basis for design of PS inhibitors.

## MATERIALS AND METHODS

**Cloning, Expression, and Purification.** The detailed procedures for cloning of the *panC* gene, expression, and purification of the protein have been described (9). Briefly, the MTB *panC* gene (Rv3602c) encoding the pantothenate synthetase was amplified from the genomic DNA of MTB strain H37Rv and inserted into a pET30a plasmid (Novagen) between restriction sites *Nco*I and *Hind*III. BL21(DE3) cells containing the pET30-PanC plasmid were grown in LB medium containing 50  $\mu$ g/mL kanamycin at 37 °C and induced with 0.4 mM IPTG for 3 h. Cells were collected and lysed by lysozyme treatment followed by sonication in 20 mM HEPES (pH 7.8), 500 mM NaCl, and 0.5 mM PMSF.

The protein was purified from a Ni<sup>2+</sup>-charged HiTrap chelating column (Amersham) with 20 mM HEPES (pH 7.8), 500 mM NaCl, and a linear gradient of imidazole. The PS protein was subjected to enterokinase (New England Biolabs) digestion for cleavage of the N-terminal fusion tag and further purified with a second round of chromatography on a Ni<sup>2+</sup>-HiTrap column followed by Superdex 75 column (Amersham) gel filtration.

**Crystallization and Data Collection.** Crystallization was carried out as described previously (9). Crystals in space group P2<sub>1</sub> were obtained from drops set up with well solutions containing 10–15% PEG 3000, 5% glycerol, 2% ethanol, 20 mM MgCl<sub>2</sub>, 150 mM Li<sub>2</sub>SO<sub>4</sub>, and 100 mM imidazole (pH 8.0) at 20 °C. We obtained a new crystal form, which is in space group C2, from similar conditions, except with 2% 2-propanol in place of ethanol and with 200 mM Li<sub>2</sub>SO<sub>4</sub>.

To obtain crystals in complex with  $\beta$ -alanine, crystals of both P2<sub>1</sub> and C2 space groups were soaked in the mother liquor with both AMP and  $\beta$ -alanine (~10 mM each) added for ~14 h. In another experiment, the C2 crystals were also soaked in the mother liquor with both ATP and pantoate (~10 mM each), in the hope of placing the pantoyl adenylate in the active site.

Before data collection, crystals were soaked for 2–5 min in a cryogenic solution similar to well solutions or substrate soaking solutions with glycerol added to 30%, and they were then flash-frozen in a cryo stream of N<sub>2</sub> gas at 100 K. Diffraction data were collected at 100 K on a Rigaku FRD generator with an R-Axis IV<sup>++</sup> detector. Data reduction and scaling were carried out with DENZO and SCALEPACK (11). Data processing statistics are listed in Table 1. The P2<sub>1</sub> crystal form has two PS molecules per asymmetric unit, while the C2 crystal contains one molecule per asymmetric unit. The data were converted to CCP4 format and structure factor amplitudes calculated by TRUNCATE and other programs in the CCP4 suite (12).

**Molecular Replacement, Model Building, and Refinement.** The structure of the apoenzyme (9) was used to calculate initial phases for the data sets by molecular replacement with AMORE (13). Structural models were refined against the diffraction data using CNS (14) and/or REFMAC (15). After each cycle of refinement, models were manually adjusted with electron density maps using O (16). All models were refined for a few cycles of REFMAC with TLC refinement (17) until convergence. The same subsets of data for  $R_{\text{free}}$

Table 2: Atomic Refinement of Models for *M. tuberculosis* Pantothenate Synthetase and Complexes

|                                                                    | apo (1)                               | AMP (2)                                                     | ATP (3)                                                                         | $\beta$ -alanine (4)                                                                                                               |
|--------------------------------------------------------------------|---------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------|
| resolution ( $\text{\AA}$ )                                        | 20–1.7                                | 20–1.7                                                      | 20–1.55                                                                         | 20–1.85                                                                                                                            |
| $R_{\text{work}}$ ( $R_{\text{free}}$ ) <sup>a</sup>               | 15.7 (18.8)                           | 15.3 (17.7)                                                 | 15.0 (17.0)                                                                     | 15.6 (19.5)                                                                                                                        |
| rmsd for bonds ( $\text{\AA}$ )                                    | 0.019                                 | 0.019                                                       | 0.018                                                                           | 0.022                                                                                                                              |
| rmsd for angles (deg)                                              | 1.64                                  | 1.70                                                        | 1.71                                                                            | 1.78                                                                                                                               |
| no. of protein atoms                                               | 2067                                  | 2062                                                        | 2065                                                                            | 4216                                                                                                                               |
| average $B$ factors, protein atoms ( $\text{\AA}^2$ ) <sup>b</sup> | 27.1                                  | 24.6                                                        | 22.8                                                                            | 30.5                                                                                                                               |
| no. of other molecules                                             | 2 glycerols (36.4)                    | 1 AMP (19.6)                                                | 1 ATP (17.3)                                                                    | 2 AMP molecules (24.5, 27.7) <sup>d</sup>                                                                                          |
| [average $B$ factors ( $\text{\AA}^2$ )]                           | 1 sulfate (45.8)<br>287 waters (41.3) | 1 sulfate (69.9)<br>2 glycerols (40.8)<br>313 waters (37.7) | 2 Mg atoms (23.5, 53.4) <sup>c</sup><br>3 glycerols (39.2)<br>294 waters (35.7) | 2 $\beta$ -alanines (42.3, 54.3) <sup>e</sup><br>2 sulfates (93.9)<br>5 glycerols (39.8)<br>5 ethanols (46.0)<br>313 waters (37.5) |

<sup>a</sup>  $R$  factors were calculated using data in the resolution range for refinement without a  $\sigma$  cutoff.  $R_{\text{free}}$  was calculated with a subset of data (8%) never used in the refinement.  $R_{\text{work}}$  was calculated against the data used in the refinement. <sup>b</sup>  $B$  factors reported for the protein atoms are total  $B$  factors after TLC refinements. <sup>c</sup> The first  $\text{Mg}^{2+}$  with a  $B$  factor of  $23.5 \text{ \AA}^2$  is coordinated by the phosphates of ATP. Positive electron density was found between side chains of Asp88 and Asp89 with distances to the oxygen atoms of 2.14 and 2.50  $\text{\AA}$ , respectively, and a  $\text{Mg}^{2+}$  was modeled into the electron density. A water molecule with a distance of 2.46  $\text{\AA}$  could be another ligand to this ion. The high  $B$  factor indicates a low occupancy. <sup>d</sup> The two average  $B$  factors that are listed are for the AMP molecules in subunits A and B, respectively. <sup>e</sup> Two  $\beta$ -alanine molecules were found in the structure. The first one with a lower  $B$  factor was found in the active site of subunit A; the other was found at the dimer interface with a low occupancy. The latter might be an artifact from a high concentration of  $\beta$ -alanine in the soaking solution.

calculation were kept between the CNS and REFMAC programs. No NCS restraints were applied for the  $P2_1$  data set in the final REFMAC refinement cycles. Final refinement statistics for the refined coordinate sets are reported in Table 2.

## RESULTS

**Binding Interactions of  $\beta$ -Alanine.** Our earlier work (9) indicated that the binding site for  $\beta$ -alanine exists only after the reaction intermediate, pantoyl adenylate, is formed in the active site. On the basis of the structure of the reaction intermediate complex, it is likely that the phosphate group of pantoyl adenylate serves as an anchor for the binding of  $\beta$ -alanine by offering hydrogen-bonding partners and/or favorable charge–charge interactions. Furthermore, an AMP molecule bound at the active site can also assist the binding of  $\beta$ -alanine from its phosphate group. In this work, when we soaked crystals of PS in a solution containing  $\beta$ -alanine and AMP, we obtained a crystal structure of the PS enzyme in complex with  $\beta$ -alanine and AMP from the  $P2_1$  crystals, which has a dimer in the asymmetric unit.

Figure 1 shows the initial  $F_o - F_c$  difference electron density for  $\beta$ -alanine. Clear electron density was seen for  $\beta$ -alanine only in subunit A of the dimer. The  $\beta$ -alanine molecule binds in the active site of subunit A in the upper part of the active site cavity, with its amino group near the phosphate group of AMP (Figure 2A). The amino group forms hydrogen bonds to the phosphate group of AMP and to two water molecules. Both water molecules are fixed in position through one hydrogen bond to the phosphate group and a second hydrogen bond to the side chain of Asp161 or Tyr82. The carboxyl group of  $\beta$ -alanine has one hydrogen bond to the N $\epsilon$ 2 atom of the Gln72 side chain. This carboxyl is also close to the side chains of Arg198 and His135, with distances of  $\sim 4.0$  and  $3.4 \text{ \AA}$ , respectively, which is favorable for charge–charge interactions and  $\pi$ -electron interactions. One side of  $\beta$ -alanine faces side chains of Met40 and Tyr82; the other side faces several ordered water molecules in the active site cavity. The binding of  $\beta$ -alanine is not as tight as that of other substrates, and its binding site in the crystal is

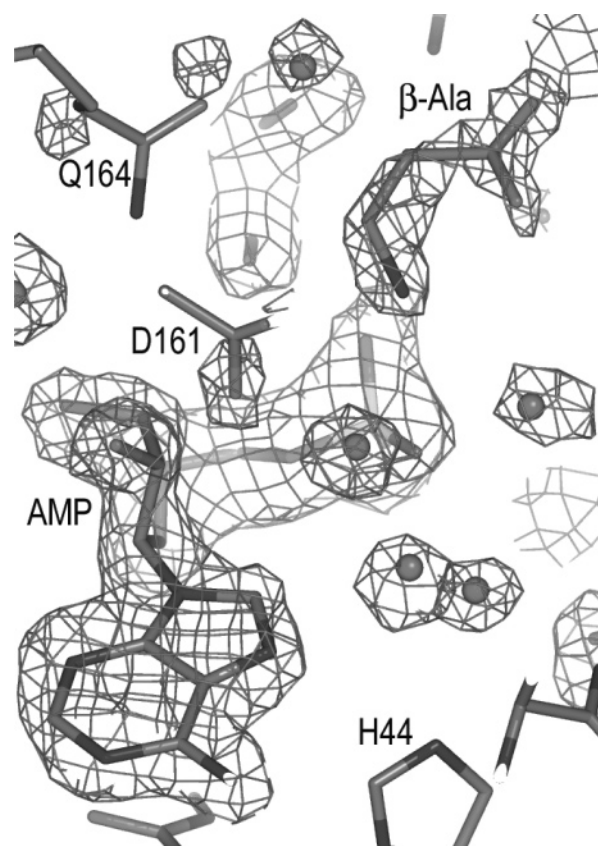


FIGURE 1: Section of the initial difference electron density map ( $F_o - F_c$ ) in the active site of subunit A of the  $P2_1$  crystal showing electron density for  $\beta$ -alanine and AMP. The difference density was calculated at  $1.9 \text{ \AA}$  after rigid body refinement with the apoenzyme structure (PDB entry 1MOP) as a model and contoured at the  $2.2\sigma$  level. Superimposed on the electron density is the final refined model. The side chains that are shown have little movement relative to those in the apoenzyme and, thus, have little initial difference electron density. This figure was prepared with PYMOL.

not fully occupied. However, because the initial model is essentially the correct structure, the initial  $F_o - F_c$  electron density gives unmistakable density for the binding of  $\beta$ -alanine in the active site of subunit A.



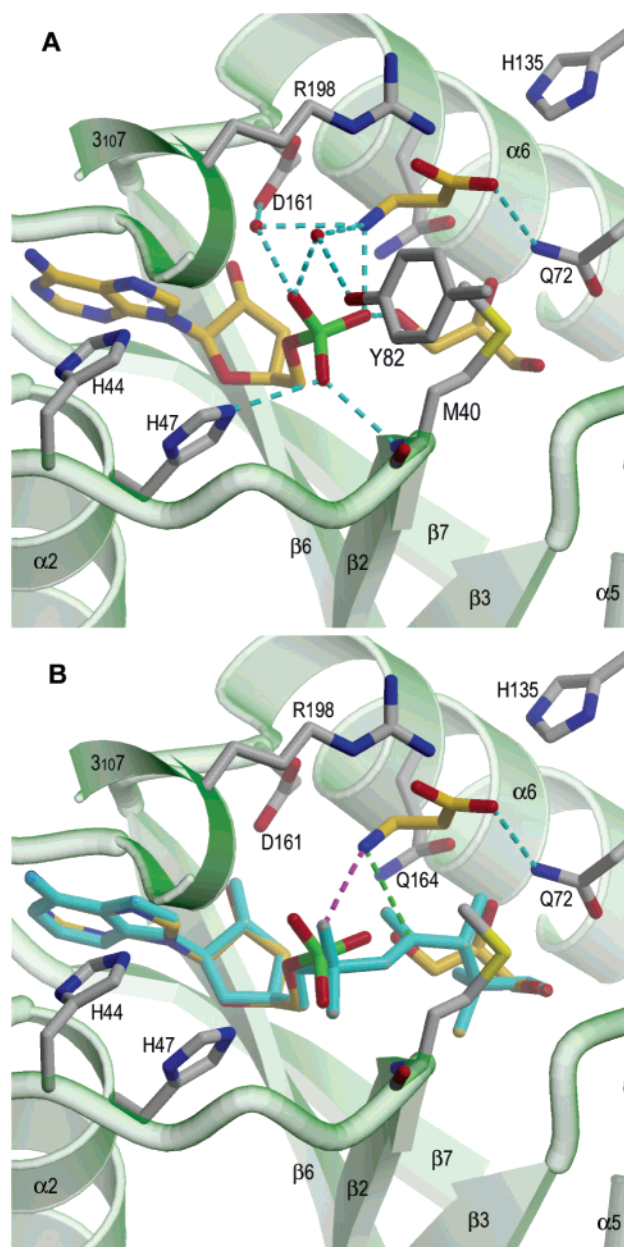


FIGURE 2: (A) Binding interactions of  $\beta$ -alanine in the active site. The AMP,  $\beta$ -alanine, and a glycerol at the pantoate binding site are colored with carbon atoms in gold. Two water molecules that have hydrogen bonds to  $\beta$ -alanine are shown as red spheres. Some hydrogen bonds are shown as dashed lines. Other hydrogen bonds, such as those between the glycerol and Gln72 and Gln164, are not shown for clarity. (B) Superposition of the active site residues of subunit A between the  $\beta$ -alanine complex and the pantooyl adenylate complex (PDB entry 1N2I). The active site residues align well with an rmsd for C $\alpha$  atoms of less than 0.15 Å. The protein structure is from the  $\beta$ -alanine complex, as in panel A, except that Tyr82 and the water molecules are not shown. The pantooyl adenylate molecule colored cyan is from the pantooyl adenylate complex. A hydrogen bond from the carboxylate group of  $\beta$ -alanine to the Gln72 side chain is shown as a cyan dashed line. The distance between the amino group of  $\beta$ -alanine and the phosphate oxygen of pantooyl adenylate is 2.8 Å, suggesting a potential hydrogen bond (shown as a magenta dashed line) for initial binding of  $\beta$ -alanine. In this binding position, the nitrogen atom of  $\beta$ -alanine is 3.38 Å from the carbonyl carbon of the pantooyl group. This figure was prepared with Molscript (23) and Raster3D (24).

No clear electron density for  $\beta$ -alanine was found in the active site of subunit B, presumably because of the disordered loop in the crystal that contains Tyr82 (9). Similarly, in the

C2 form crystal, which has one molecule per asymmetry unit and the flexible loop disordered (see below for more details), no  $\beta$ -alanine was found in the active site when crystals were soaked in a solution containing both AMP and  $\beta$ -alanine.

When we superimposed the structure of the  $\beta$ -alanine complex with that of the pantooyl adenylate complex (Figure 2B), the amino group of  $\beta$ -alanine was positioned above the carbonyl group of pantooyl adenylate such that the nitrogen atom was 3.38 Å from the carbonyl carbon. With changes in only one torsion angle (see the movie in the Supporting Information), the amino group of  $\beta$ -alanine can move toward the carbonyl carbon for a nucleophilic attack and attain a nearly perfect geometry for a tetrahedral intermediate. In comparison to that of AMP, the phosphate group of pantooyl adenylate is slightly shifted, presumably because of its bonding restraint. This shift allows the O1 $\alpha$  atom of pantooyl adenylate to form a hydrogen bond with the amino group of  $\beta$ -alanine, serving as an anchor for the initial binding of  $\beta$ -alanine.

**Structure of the C2 Crystal Form.** A new crystal form was obtained which differs from the crystals in space group  $P2_1$ . This crystal form is in space group C2 with one polypeptide chain per asymmetric unit. The protein molecule in the C2 crystal is present as a dimer that is essentially identical to that in the  $P2_1$  crystal, but in the C2 crystal, the 2-fold axis of the dimer coincides with one crystal symmetry axis. The interface between subunits of the dimer is essentially identical for the two crystal forms. However, in the C2 crystal, the flexible region between secondary structure elements  $\beta 3$  and  $\alpha 3$ , which forms one wall of the active site cavity, differs significantly from that in the  $P2_1$  crystal (Figure 3). A large portion, from residue 76 to 86, is completely disordered. The first turn of helix  $\alpha 3$  unwinds, and helix  $3_{10}3$  in the  $P2_1$  crystal also unwinds and forms an extended structure in which most residues are partially disordered, as indicated by high B factors. This structural difference is induced by crystal packing and the intrinsic flexibility of the loop, which functions as a gate to the active site cavity (9).

**ATP Binding Interactions.** In the  $P2_1$  crystal, the disordered loop becomes ordered when the reaction intermediate forms in the active site (9). To determine if this would also happen in the C2 crystal, we soaked crystals in a solution containing both ATP and pantoate. We obtained a complex having an ATP rather than an intermediate in the active site. The ATP molecule has an average temperature factor of 17.3 Å<sup>2</sup>, virtually identical to the average temperature factor of the surrounding protein atoms, indicating that it is at full occupancy. There was no clear electron density for pantoate or pantooyl adenylate in the active site cavity. In the pantoate binding pocket, there was positive difference electron density, into which we modeled a glycerol molecule. This glycerol has a high temperature factor, suggesting that it has a low occupancy. The inability of pantoate to react with ATP in this C2 crystal seems to be due to the conformation of the flexible loop, which is different from that in the  $P2_1$  crystal (Figure 3). The side chain of Gln72 on this flexible loop plays an important role in binding pantoate by hydrogen bonding interactions, but in the C2 crystal, it swings out of the active site and is partially disordered. Except for this Gln72 side chain, the binding site for pantoate is intact.

The binding interactions of ATP with protein are similar to those of AMPCPP. However, unlike the methylene group

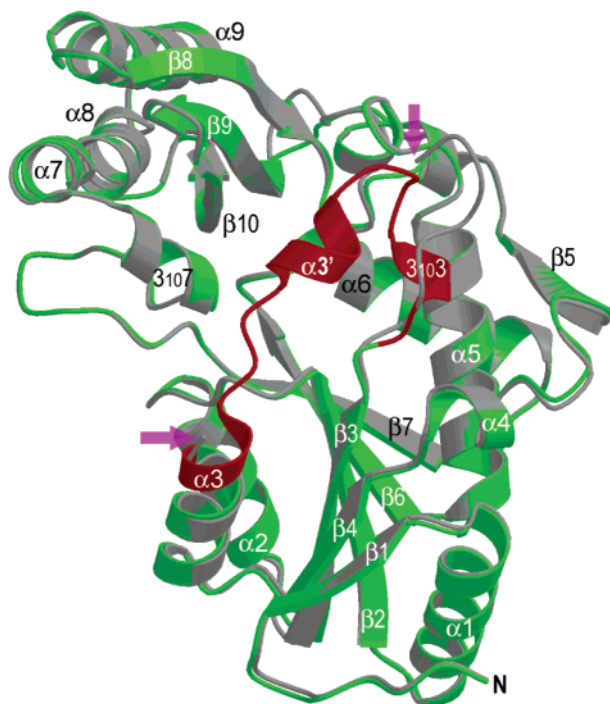


FIGURE 3: Superposition of the protein molecule in the asymmetric unit of the C2 crystal and that of subunit A of the P2<sub>1</sub> crystal. The structures are shown as ribbon diagrams, which are colored gray for the C2 crystal and green and red for the P2<sub>1</sub> crystal. Colored in red is the structurally different region. Most of the residues align very well with an rmsd of 0.33 Å for 255 C $\alpha$  atoms. However, a large segment that forms a wall of the active site cavity is disordered in the structure of the C2 crystal (two magenta arrows mark the ends of the disordered segment). The first turn of helix  $\alpha$ 3 (horizontal arrow) unwinds, and residues that form the 3<sub>10</sub>3 helix in the P2<sub>1</sub> crystal have an extended structure in the C2 crystal. The 3<sub>10</sub>3 helix contains Gln72, which plays an important role in binding pantoate by having hydrogen bonds to the hydroxyl groups of pantoate. Strand  $\beta$ 5 forms a two-stranded intersubunit  $\beta$ -sheet of the dimer.

in AMPCPP, the bridge oxygen between the  $\alpha$ - and  $\beta$ -phosphate groups of ATP is capable of forming a hydrogen bond with the N $\epsilon$ 2 atom of His47 (Figure 4). This hydrogen bond, in turn, allows the  $\alpha$ -phosphate to form a hydrogen bond with the amide nitrogen of Met40. The magnesium ion that is bound to ATP has a nearly perfect octahedral coordination. Its distances to the O2 $\alpha$ , O2 $\beta$ , and O1 $\gamma$  atoms of the phosphate groups are 2.26, 1.99, and 2.12 Å, respectively; three other ligands are water molecules with distances of 1.96, 2.11, and 2.28 Å. Two of the water ligands have hydrogen bonds to the side chain of Asp161.

**AMP Binding Interactions.** We obtained an AMP complex from both P2<sub>1</sub> and C2 crystals by soaking the crystals in a solution containing both AMP and  $\beta$ -alanine. As described above, the P2<sub>1</sub> crystal has both AMP and  $\beta$ -alanine in the active site of one subunit but only AMP in the other, since one wall of the active site cavity is disordered in the latter. The C2 crystal has a larger disordered region, which includes Gln72; thus, it cannot bind  $\beta$ -alanine. Binding interactions of AMP with the active site residues in these complexes are essentially identical. The adenosine group of AMP binds in the same position as that of AMPCPP and ATP; it fits snugly in its binding pocket at the bottom of the active site cavity through hydrophobic and hydrogen bonding interactions. The adenine group is flanked by Gly46 on helix  $\alpha$ 2 and Lys160

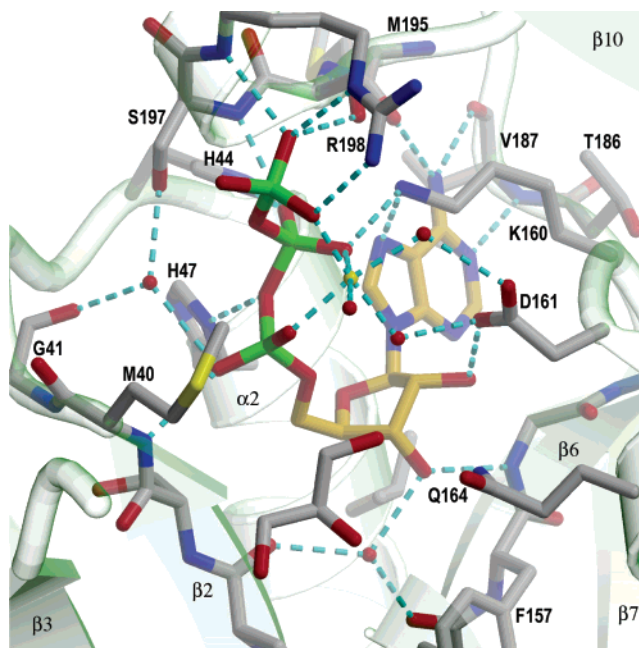


FIGURE 4: Binding interactions of ATP with active site residues. The Mg<sup>2+</sup> ion is shown as a yellow sphere, and water molecules are shown as red spheres. Hydrogen bonds between ATP and protein atoms as well as some water-mediated hydrogen bonds are shown as dashed lines. The Mg<sup>2+</sup> ion has nearly perfect octahedral coordination with three ligands from oxygen atoms of the phosphate groups and another three from water molecules. The N $\epsilon$ 2 atom of His47 has a hydrogen bond to the bridge oxygen between the  $\alpha$ - and  $\beta$ -phosphate. A glycerol molecule found at the pantoate binding site is also shown.

on the loop after  $\beta$ 6. Its N1 and N6 atoms have hydrogen bonds to main chain atoms. The hydroxyl groups of ribose form hydrogen bonds with the Asp161 side chain and a few main chain atoms at the bottom of the active site cavity. The phosphate group, however, has torsional flexibility, and it rotates slightly relative to the  $\alpha$ -phosphate of ATP, allowing one of its oxygen atoms to form hydrogen bonds to the side chain of His47 and to the amide nitrogen of Met40 simultaneously (Figure 2A).

## DISCUSSION

**A Dimer as the Biological Functional Unit.** In this work, we obtained a new crystal form in space group C2, in which there is one molecule per asymmetric unit. However, the same dimer as in the P2<sub>1</sub> crystal was found, with the 2-fold symmetry axis of the dimer coinciding with the crystal  $y$  axis. This result confirms that the biological unit of this enzyme is a dimer. The MTB PS enzyme exists in solution as a dimer (8, 9), as do PS enzymes from higher plants (6, 7). The crystal structures suggest that active sites of the dimer are independent of each other in catalytic function; reactions can occur in both active sites of the dimer simultaneously. When we grew crystals in the presence of both ATP and pantoate or soaked crystals in solutions containing both ATP and pantoate, the reaction intermediate was found in both active sites. This suggests that dimerization serves a function other than allostery. Although we cannot exclude the possibility of cooperative binding of  $\beta$ -alanine in the two active sites, the enzyme surface feature suggests that regulation of enzyme function could occur through interactions of the dimer with other macromolecules. There is a large groove

that is more than 6 Å wide and 8 Å deep at the dimer interface near the bottom of the molecule. In addition, a group of arginine residues from each subunit forms a large continuous patch of positively charged surface across the groove at the dimer interface. The *E. coli* pantothenate synthetase [PDB entry 1IHO (5)] also has a large groove at its dimer interface, and positively charged side chains decorate the surface near and in the groove.

**Functions of the Active Site Residues.** The HIGH (18, 19) and KMSKS (20) sequence motifs that are important in binding of ATP in tRNA synthetases are present in the structure of pantothenate synthetase (5). In the MTB PS structure, the HIGH motif is at the N-terminus of helix  $\alpha_2$ . Both His residues in the motif interact with the phosphates of ATP. The first histidine, His44, has a hydrogen bond extending from its N $\epsilon$ 2 to O1 $\beta$  of  $\beta$ -phosphate. The His47 side chain has a hydrogen bond from its N $\epsilon$ 2 to the bridge oxygen between the  $\alpha$ - and  $\beta$ -phosphates, suggesting that it functions as a general acid by donating a proton to the leaving pyrophosphate group. Both histidine residues have the N $\delta$ 1 atom forming a hydrogen bond to a main chain NH group, and thus, the N $\epsilon$ 2 atoms are hydrogen bond donors. Their imidazole rings are in the proximity of each other with the closest separation being  $\sim 3.2$  Å. This indicates favorable  $\pi$ - $\pi$  interactions. These two side chains also play a role in binding the adenine group and stabilizing the reaction intermediate, because the imidazole rings are close to the adenine (shortest distances of 3.4 and 4.3 Å for His44 and His47, respectively). Mutation of either histidine residue to alanine causes a  $>1000$ -fold reduction in enzyme activity (10). The Gly46 residue is highly conserved in this motif. It flanks one side of adenine (with Lys160 on the other side), where any side chain larger than a hydrogen atom would have steric hindrance. The KSMKS motif is not conserved at the sequence level in PS enzymes, but it is evident from structural alignments (5). In the MTB PS structure of the ATP complex, residues Ser196, Ser197, and Arg198, which are at the N-terminus of helix  $3_{10}7$  and are part of the KSMKS motif, bind the  $\beta$ - and  $\gamma$ -phosphate groups. The Lys160 side chain packs against one side of adenine. Mutation of this residue to alanine reduces the enzyme activity more than 1000-fold, and it decreases the affinity of the enzyme for ATP (10). Another residue that binds ATP is Asp161, which has hydrogen bonds to two water ligands of the magnesium ion, and to O2\* of the ribose group. Overall, the ATP molecule is rigidly held in the active site by many hydrophobic,  $\pi$ -electron, hydrogen bonding, and charge-charge interactions, which keep it in an optimal conformation for a nucleophilic attack by pantoate. Residues involved in binding the adenosine group are also important in stabilizing the pantoyl adenylate intermediate.

As shown in our earlier work (9), binding of pantoate involves hydrogen bonding interactions with side chains of Gln72 and Gln164 and hydrophobic interactions with side chains of Pro38, Phe157, and Met40. Mutation of either glutamine residue to alanine greatly reduces the rate of pantoyl adenylate formation (10), because these two side chains bind pantoate and keep the nucleophile at an optimal position for a nucleophilic attack. We observed only ATP in the active site with full occupancy when we soaked the C2 crystals in a solution containing both ATP and pantoate. In this structure, the Gln72 side chain is outside the active

site cavity because of crystal packing. This result indicates that loss of this glutamine side chain completely abolishes the enzyme catalytic activity for the first half of the reaction inside the crystal. The glutamine side chains are also involved in binding of the reaction intermediate. Thus, mutation of either residue could affect the stability of the pantoyl adenylate intermediate. It is noteworthy that Asn69 was also found to be important for the enzyme activity (10). This residue is not directly involved in binding substrates or the intermediate. However, it forms a hydrogen bond to the side chain of Gln72 and thus is important for keeping the Gln72 side chain in position. In addition, both Asn69 and Gln72 are on the one-turn  $3_{10}3$  helix, and mutation of Asn69 might also affect the stability of this  $3_{10}$  helix and hence the position of the Gln72 side chain.

The structure of the  $\beta$ -alanine complex indicates that Gln72 and Met40 side chains also bind  $\beta$ -alanine. The side chain of Arg198, which is important for binding the  $\gamma$ -phosphate of ATP, moves close to the carboxylate group of  $\beta$ -alanine for favorable charge-charge interactions. The side chain of Tyr82 also plays a role in binding of  $\beta$ -alanine. This Tyr82, together with the side chain of Met40, packs on one side of the bound  $\beta$ -alanine. In the other subunit of the dimer, where residues 74 to 83 are disordered, no  $\beta$ -alanine was found in the active site. Also in this case, disorder of this active site cavity wall weakens binding interactions for  $\beta$ -alanine because its binding site is exposed to bulk solvent and thus the charge-charge interactions are dampened. The binding of  $\beta$ -alanine is not as tight as that of other substrates. One side of  $\beta$ -alanine faces the large cavity of the active site, which is filled with a few ordered water molecules in the crystal structure. However, the binding interactions limit molecules that can bind well and have a nucleophilic attack on the reaction intermediate. Therefore, structural analogues of  $\beta$ -alanine are all poor substrates (8).

**Overall Enzyme-Catalyzed Reaction Mechanism.** In the first step of the PS-catalyzed reaction, ATP reacts with pantoate to form pantoyl adenylate. The structure of the ATP complex reported here confirms our earlier model of ATP binding interactions based on the AMPCPP complex (9). The ATP molecule binds in a position essentially identical to that of the AMPCPP. Our earlier prediction that the N $\epsilon$ 2 atom of His47 can form a hydrogen bond to the bridge oxygen between the  $\alpha$ - and  $\beta$ -phosphates is confirmed in the ATP complex. This hydrogen bond allows the  $\alpha$ -phosphate to move toward the pantoate binding site and have a hydrogen bond to the NH group of Met40 from its O1 $\alpha$  atom. When the structures of the ATP complex and the pantoate complex were superimposed, the carboxyl oxygen of pantoate was within 2.7 Å of the  $\alpha$ -phosphorus atom, and it was at a perfect position for an in-line nucleophilic attack. Both ATP and pantoate can bind alone in the active site. However, when both are present in the same active site, the nucleophilic reaction must occur. As discussed in our previous work (9), it is likely that ATP binds first. Then pantoate binds, allowing a nucleophilic attack by the  $\alpha$ -phosphate while it binds in its binding pocket (8, 21). This produces a trigonal bipyramidal transient reaction intermediate (see the movie in the Supporting Information). Although transient in nature, this intermediate is stabilized by the enzyme because all the favorable interactions are preserved, and the steric and charge repulsions are eliminated. Formation of the bipyramidal



intermediate involves only the movement of the phosphorus atom toward the carboxyl oxygen to be coplanar with three equatorial oxygen atoms. The intermediate then dissociates to form the stable reaction intermediate pantooyl adenylate and pyrophosphate. The side chain of His47 could donate a proton to the leaving pyrophosphate to facilitate the reaction. Pantooyl adenylate is stable only when tightly bound in the enzyme active site. In solution, it decomposes rapidly to give pantooyl lactone and AMP (22). The dissociation of the trigonal bipyramidal intermediate necessitates the simultaneous leaving of the pyrophosphate, because there would be unfavorable charge and steric repulsions once pyrophosphate is formed. This reaction, however, is reversible, as was demonstrated by positional isotope exchange experiments in which  $^{18}\text{O}$  labels were scrambled when incubating  $[\beta\gamma\text{-}^{18}\text{O}_6]\text{-ATP}$  and pantoate with the MTB PS enzyme (21).

In the second step of the reaction,  $\beta$ -alanine reacts with the enzyme-bound pantooyl adenylate to yield the final reaction product, pantothenate. Upon formation of the pantooyl adenylate intermediate in the active site cavity, the flexible wall becomes ordered and thus closes the active site cavity. This protects the highly reactive pantooyl adenylate and also creates the binding site for  $\beta$ -alanine, the last substrate. The  $\beta$ -alanine molecule is relatively small, capable of passing through the small opening at the top of the active site cavity. When we soaked crystals with both pantoate and  $\beta$ -alanine, we observed very weak binding of  $\beta$ -alanine in the active site cavity (9). Its carboxyl group has charge interactions with side chains of Arg198 and Arg132, while its amino group has a hydrogen bond to the side chain of Asp161, being  $\sim 7.5$  Å from the carboxyl group of pantoate. In the presence of pantooyl adenylate in the active site, its phosphate group forms a hydrogen bond to the amino group of  $\beta$ -alanine. This allows the carboxylate group of  $\beta$ -alanine to have better interactions with the side chains of Arg198 and His135 and a hydrogen bond to Gln72. There is a shift of the Arg198 side chain of  $\sim 1.8$  Å toward the carboxylate group of  $\beta$ -alanine relative to that in the pantooyl adenylate complex. This binding position for  $\beta$ -alanine puts its amino group at a good position for a nucleophilic attack on the carbonyl carbon of the pantooyl group. A model of the tetrahedral intermediate can be readily prepared from the crystal structure with only a simple rotation of a torsion angle of  $\beta$ -alanine (see the movie in the Supporting Information). This transient tetrahedral intermediate is stabilized by favorable interactions with the enzyme active site residues (Figure 5). Once the intermediate dissociates to form AMP and pantothenate, the planarity of the peptide bond in pantothenate forces the molecule to have a conformation that cannot have favorable interactions with the side chains that bind pantoate and  $\beta$ -alanine. In addition, there are steric clashes with the phosphate group of AMP and the Asp161 side chain. Therefore, pantothenate will leave the active site once it is formed. Pantothenate has a very low affinity for the enzyme. We have tried soaking and growing crystals in the presence of pantothenate, but we failed to obtain a complex with this compound. This is consistent with the finding that pantothenate is a poor inhibitor of the enzyme (8). The AMP molecule, on the other hand, has good binding interactions in the active site. However, it typically is present at a low concentration in cells and thus can easily diffuse out of the active site. Moreover, ATP has many favorable

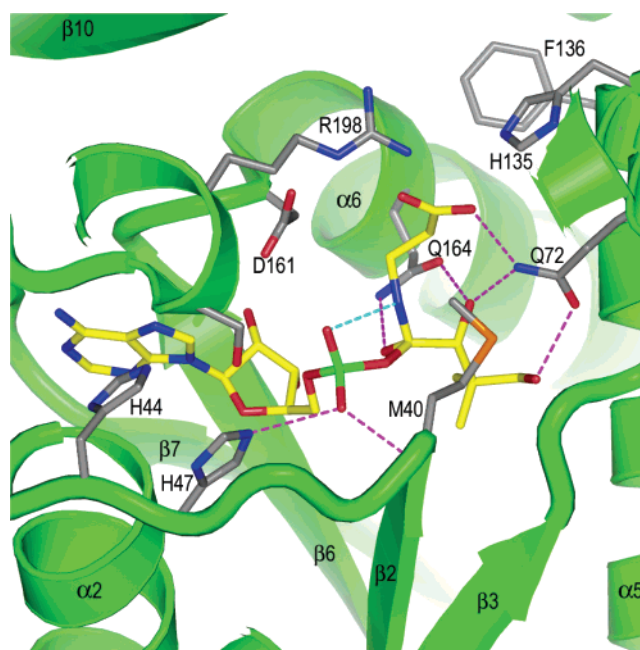


FIGURE 5: Model of the tetrahedral intermediate of the second half-reaction catalyzed by PS. The model was constructed on the basis of Figure 2B of the superposition between the pantooyl adenylate complex and the  $\beta$ -alanine complex. A tetrahedral geometry can be obtained by a torsional rotation of the amino nitrogen of  $\beta$ -alanine to move it to  $\sim 1.5$  Å above the carbonyl carbon, and another torsional rotation of the carbonyl oxygen of pantoate. Hydrogen bonds to  $\beta$ -alanine, pantooyl, and phosphate groups are shown as magenta dashed lines. A cyan dashed line between O1 $\alpha$  and the amino nitrogen outlines a potential ring system for the design of inhibitors (see the text for details). The distance between these two atoms in this model is  $\sim 3$  Å. However, a hydrogen bond is unlikely on the basis of the geometry.

interactions in the active site of the enzyme, and it can readily displace AMP.

**Inhibitor Design.** Because pantooyl adenylate binds tightly in the active site cavity, it is reasonable to expect that nonreactive analogues of the intermediate would be good inhibitors (9). The tetrahedral intermediate from the reaction of  $\beta$ -alanine with pantooyl adenylate (Figure 5) can be another template for inhibitor design. It is well accepted that enzymes lower the transition state energy by stabilizing reaction intermediates along the reaction coordinates. In this case, the tetrahedral intermediate is tightly bound by hydrophobic interactions, hydrogen bonds, and charge–charge interactions. The easiest way to design an inhibitor based on this tetrahedral intermediate is to remove the adenosine group and retain only the carboxyl and pantooyl groups. The tetrahedral conformation can be stabilized by creating a five- or six-membered ring through the tetrahedral carbon, O12, P $\alpha$ , O1 $\alpha$ , and the amino nitrogen (atoms connected into a ring by the cyan dashed line in Figure 5) to make the structure rigid. Members of the ring can be all carbon, or mixtures of carbon, nitrogen, and oxygen. A hydroxyl or carbonyl group at the O2 $\alpha$  position can have hydrogen bonds to the side chain of His47 and the NH group of Met40, thus increasing the binding affinity.

#### ACKNOWLEDGMENT

We thank Drs. Loren Day, Karl Drlica, David Dubnau, Leonard Mindich, Richard Pine, Issar Smith, and David Wah for helpful comments on and revisions of the manuscript.

## SUPPORTING INFORMATION AVAILABLE

A movie showing detailed reactions in the active site of the PS enzyme. The movie was prepared with PYMOL (<http://pymol.sourceforge.net/>) based on the crystal structures of the apoenzyme, ATP complex, pantoate complex, pantoyl adenylate complex,  $\beta$ -alanine-AMP complex, and AMP complex. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## REFERENCES

- Jackowski, S. (1996) *Biosynthesis of pantothenic acid and coenzyme A in Escherichia coli and Salmonella typhimurium: Cellular and molecular biology*, 2nd ed., Vol. 1, pp 687–694, American Society for Microbiology, Washington, DC.
- Maas, W. K. (1960) The biosynthesis of pantothenic acid, *Proc. Int. Congr. Biochem.* 11, 161–168.
- Sambandamurthy, V. K., Wang, X., Chen, B., Russell, R. G., Derrick, S., Collins, F. M., Morris, S. L., and Jacobs, W. R., Jr. (2002) A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis, *Nat. Med.* 8, 1171–1174.
- Merkel, W. K., and Nichols, B. P. (1996) Characterization and sequence of the *Escherichia coli* panBCD gene cluster, *FEMS Microbiol. Lett.* 143, 247–252.
- von Delft, F., Lewendon, A., Dhanaraj, V., Blundell, T. L., Abell, C., and Smith, A. G. (2001) The crystal structure of *E. coli* pantothenate synthetase confirms it as a member of the cytidyltransferase superfamily, *Structure* 9, 439–450.
- Genschel, U., Powell, C. A., Abell, C., and Smith, A. G. (1999) The final step of pantothenate biosynthesis in higher plants: Cloning and characterization of pantothenate synthetase from *Lotus japonicus* and *Oryza sativum* (rice), *Biochem. J.* 341 (Part 3), 669–678.
- Perez-Espinosa, A., Roldan-Arjona, T., and Ruiz-Rubio, M. (2001) Pantothenate synthetase from *Fusarium oxysporum* f. sp. *lycopersici* is induced by  $\alpha$ -tomatine, *Mol. Genet. Genomics* 265, 922–929.
- Zheng, R., and Blanchard, J. S. (2001) Steady-state and pre-steady-state kinetic analysis of *Mycobacterium tuberculosis* pantothenate synthetase, *Biochemistry* 40, 12904–12912.
- Wang, S., and Eisenberg, D. (2003) Crystal structures of a pantothenate synthetase from *M. tuberculosis* and its complexes with substrates and a reaction intermediate, *Protein Sci.* 12, 1097–1108.
- Zheng, R., Dam, T. K., Brewer, C. F., and Blanchard, J. S. (2004) Active site residues in *Mycobacterium tuberculosis* pantothenate synthetase required in the formation and stabilization of the adenylate intermediate, *Biochemistry* 43, 7171–7178.
- Otwinowski, Z., and Minor, W. (1996) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol.* 276, 307–326.
- Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760–763.
- Navaza, J. (1994) AmoRe: An automated package for molecular replacement, *Acta Crystallogr. A50*, 157–163.
- Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54* (Part 5), 905–921.
- Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of macromolecular structures using FFT, *Acta Crystallogr. D55* (Part 1), 247–255.
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47* (Part 2), 110–119.
- Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement, *Acta Crystallogr. D57*, 122–133.
- Barker, D. G., and Winter, G. (1982) Conserved cysteine and histidine residues in the structures of the tyrosyl and methionyl-tRNA synthetases, *FEBS Lett.* 145, 191–193.
- Bork, P., Holm, L., Koonin, E. V., and Sander, C. (1995) The cytidyltransferase superfamily: Identification of the nucleotide-binding site and fold prediction, *Proteins* 22, 259–266.
- Hountondji, C., Dessen, P., and Blanquet, S. (1986) Sequence similarities among the family of aminoacyl-tRNA synthetases, *Biochimie* 68, 1071–1078.
- Williams, L., Zheng, R., Blanchard, J. S., and Raushel, F. M. (2003) Positional isotope exchange analysis of the pantothenate synthetase reaction, *Biochemistry* 42, 5108–5113.
- Wieland, T., Loewe, W., Kreiling, A., and Pfeleiderer, G. (1963) [On Pantothenic Acid Synthetase from *E. coli*. V. Pantoyladenylate as the Acylating Component in the Enzymatic Synthesis of Pantothenic Acid], *Biochem. Z.* 339, 1–7.
- Kraulis, P. J. (1991) MOLSCRIPT: A Program to Produce Both Detailed and Schematic Plots of Protein Structures, *J. Appl. Crystallogr.* 24, 946–950.
- Merritt, E. A., and Murphy, M. E. (1994) Raster3D Version 2.0. A program for photorealistic molecular graphics, *Acta Crystallogr. D50*, 869–873.

BI051873E