SNAPSHOTS OF THE PANTOTHENATE SYNTHETASE FROM MYCOBACTERIUM TUBERCULOSIS
ALONG THE REACTION COORDINATE

Shuishu Wang
Public Health Research Institute, 225 Warren Street, Newark, New Jersey 07103

David Eisenberg
Howard Hughes Medical Institute, UCLA-DOE Institute of Genomics and Proteomics, Molecular Biology Institute, UCLA, Box 951570, Los Angeles, CA 90095-1570

1. ABSTRACT

Pantothenate synthetase (PS) from *Mycobacterium tuberculosis* represents a potential target for developing new anti-tuberculosis drugs. PS catalyzes the ATP-dependent condensation of pantoate and β-alanine to form pantothenate. We determined the crystal structures of PS from *M. tuberculosis* and its complexes with substrates ATP, pantoate, and β-alanine, as well as an ATP analog AMPCPP, a reaction intermediate pantoyl adenylate and a reaction product AMP, with resolutions from 1.6 to 2 Å. The PS structure reveals a dimer, and each subunit has two domains with tight association between domains. The active site cavity is on the N-terminal domain, covered by the C-terminal domain. There is a flexible loop that forms one wall of the active site cavity, and it opens and closes the active site cavity. The enzyme binds ATP and pantoate tightly in the active site, and brings the carboxyl oxygen of pantoate near the α-phosphorus atom of ATP for an in-line nucleophilic attack. The pantoyl adenylate thus formed is stabilized by tight binding interactions with enzyme active site residues, and protected from hydrolysis by closing of the active site cavity with the flexible loop becoming ordered. Binding of β-alanine can occur only after formation of the pantoyl adenylate intermediate. Crystal structures of these complexes provide a step-by-step view of the PS catalyzed reaction, and allow us to construct transient reaction intermediates and deduce detailed reaction mechanism. PS catalyzes the reaction by stabilizing the reaction intermediates along the reaction coordinate, and structural models of these intermediates can provide basis for inhibitor design.

2. INTRODUCTION

Pantothenate (vitamin B5) 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 coauthors [3] reported that a *Mycobacterium tuberculosis* (MTB) 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, and thus represents a new target for anti-tuberculosis agents.

The pantothenate biosynthetic pathway in bacteria 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 β-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 cytidylyltransferase 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 β-sheet. Based on structural comparison of the *E. coli* PS with other members of cytidylyltransferase 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.
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 β-alanine to form pantothenate and AMP [8]. The existence of pantoyl adenylate as an enzyme bound intermediate is suggested by the transfer of the 18O label from the carboxyl group of pantoate to the reaction product AMP. However, pantoyl adenylate cannot be directly isolated or identified by TLC methods, presumably due to rapid lactonization [9].

Here, we present the crystal structure of the MTB PS enzyme and structures of its complexes with substrates, a substrate analog (AMPCPP), a reaction product (AMP), and a pantoyl adenylate reaction intermediate [10, 11]. The MTB PS has the same fold as the E. coli enzyme. 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 apo-enzyme 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. The structure of the β-alanine complex gives direct evidence that the binding site for this last substrate exists only after formation of the pantoyl adenylate intermediate. Together these crystal structures lead to a detailed view of the overall reaction mechanism. Models of the transient reaction intermediate can be easily constructed from the experimentally observed complexes.

3. MATERIALS AND METHODS

3.1 Protein Production
The detailed procedures for cloning of the panC gene, expression and purification of the protein have been described [10]. 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), which produces a recombinant protein with an N-terminal 6xHis tag that can be cleaved off with enterokinase. BL21(DE3) cells containing the pET30-panC plasmid were grown in LB medium containing 50 µg/ml kanamycin at 37 °C and induced with 0.4 mM IPTG for 3 hours. 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 Ni2+-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 to cleave off the N-terminal fusion tag. Electrospray mass spectroscopy indicated that 9 residues from the C-terminus of the protein were also cleaved off by enterokinase digestion. However, enzyme activity assays suggested that this C-terminal truncation does not have any detectable effect on the enzyme activity [10]. The enterokinase digested protein was further purified with a second round of chromatography on a Ni2+-HiTrap column followed by Superdex 75 column (Amersham) gel filtration.

3.2 Crystallization and Data Collection
Crystallization was carried out as described previously [10]. Crystals in space group P21 were obtained from drops set up with well solutions containing 10-15% PEG 3000, 5% glycerol, 2% ethanol, 20 mM MgCl2, 150 mM Li2SO4 and 100 mM imidazole pH 8.0 at 20 °C. Those in space group C2 were from similar conditions, except with 2% isopropanol in place of ethanol and with 200 mM Li2SO4. Cocrystralization with β-alanine was done with a well solution of 15% PEG 3000, 2% ethanol, 20 mM MgCl2, 100 mM imidazole pH 8.0 and 20 mM β-alanine, from which crystals in space group P212121 were obtained. However, these crystals do not have β-alanine in the active site. Cocrystralization with both ATP and pantoate was carried out in the same crystallization condition of the P21 crystals but with both ATP and pantoate added, from which isomorphous P21 crystals were obtained that have a pantoyl adenylate intermediate in the active site.

Crystals in complex with substrates were obtained by soaking apo-enzyme crystals in solutions containing substrates or the ATP analog AMPCPP. Soaking experiments were carried out by adding solutions containing substrates or AMPCPP directly to the drops, or to drops of pseudo mother liquor with transferred crystals, and incubating overnight or longer. Crystals were soaked in solutions containing ~10 mM each of AMPCPP alone, both AMPCPP and pantoate, both pantoate and β-alanine, both ATP and pantoate, and both AMP and β-alanine. Crystals were also soaked in solutions containing up to 20 mM pantothenate, but no pantothenate was found in the crystal structure.
Snapshots of the Pantothenate Synthetase from *Mycobacterium Tuberculosis* along the Reaction Coordinate

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 N₂ gas at 100 K. Diffraction data were collected at 100 K on a Rigaku FRD generator with an R-AXIS IV⁺⁺ detector. Data reduction and scaling were carried out with the programs DENZO and SCALEPACK [12]. Data processing statistics are reported in references [10] and [11]. Table 1 summarizes a few of the data sets, of which the structures are mentioned in this paper. The P₂₁ crystal has two PS molecules per asymmetric unit; while the C₂ crystal contains one molecule per asymmetric unit. In the C₂ crystal, the active site is locked in an inactive conformation by crystal packing (see below), and the pantoate and β-alanine binding sites are compromised. Therefore, only AMP was found when we soaked the C₂ crystals with both AMP and β-alanine (data set 8, in Table 1), and those soaked with both ATP and pantoate gave an ATP complex (data set 9) instead of a reaction intermediate complex.

### Table 1

<table>
<thead>
<tr>
<th>Complex</th>
<th>Experiment</th>
<th>SG</th>
<th>#/A.U.</th>
<th>resolution</th>
<th>PDB ID</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apo enzyme</td>
<td></td>
<td>P₂₁</td>
<td>2</td>
<td>1.6 Å</td>
<td>1MOP</td>
<td>[10]</td>
</tr>
<tr>
<td>AMP-CPP and pantoate in A; AMP-CPP in B</td>
<td>AMP-CPP, pantoate, soak</td>
<td>P₂₁</td>
<td>2</td>
<td>1.6 Å</td>
<td>1N2E</td>
<td>[10]</td>
</tr>
<tr>
<td>Pantoyl adenylate</td>
<td>ATP, pantoate, cocystal</td>
<td>P₂₁</td>
<td>2</td>
<td>2.0 Å</td>
<td>1N2H</td>
<td>[10]</td>
</tr>
<tr>
<td>Pantoyl adenylate</td>
<td>ATP, pantoate, soak</td>
<td>P₂₁</td>
<td>2</td>
<td>1.7 Å</td>
<td>1N2I</td>
<td>[10]</td>
</tr>
<tr>
<td>Pantoate</td>
<td>Pantoate, β-alanine, soak</td>
<td>P₂₁</td>
<td>2</td>
<td>1.8 Å</td>
<td>1N2J</td>
<td>[10]</td>
</tr>
<tr>
<td>AMP, β-alanine in A; AMP only in B</td>
<td>AMP, β-alanine soak</td>
<td>P₂₁</td>
<td>2</td>
<td>1.85 Å</td>
<td>1A86</td>
<td>[11]</td>
</tr>
<tr>
<td>Apo enzyme</td>
<td></td>
<td>C₂</td>
<td>1</td>
<td>1.7 Å</td>
<td>2A88</td>
<td>[11]</td>
</tr>
<tr>
<td>AMP</td>
<td>AMP, β-alanine soak</td>
<td>C₂</td>
<td>1</td>
<td>1.7 Å</td>
<td>2A7X</td>
<td>[11]</td>
</tr>
<tr>
<td>ATP</td>
<td>ATP, pantoate, soak</td>
<td>C₂</td>
<td>1</td>
<td>1.6 Å</td>
<td>2A84</td>
<td>[11]</td>
</tr>
</tbody>
</table>

3.3 Structural Determination and Refinement

The crystal structure of the apo-enzyme in the P₂₁ crystal was determined [10] by the molecular replacement method using AMORE [13], with subunit A of the *E. coli* PS structure (PDB ID code 1IHO [5]) as a model. The two domains were separated as two independent search models. The positions of the two N-terminal domains and one C-terminal domain were found from the rotation and translation searches, and the other C-terminal domain was generated by a non-crystallographic-symmetry (NCS) operation on the position of the known C-terminal domain with the NCS symmetry operators derived from the two N-terminal domains. The structure of the apo-enzyme was used to calculate phases for data sets of various complexes by rigid body refinement into new data sets for isomorphous crystals, or by molecular replacement with AMORE. The final refinement statistics are reported in references [10, 11]. Structural models were refined against 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]. The same subsets of data for R_free calculation were kept between the CNS and REFMAC programs.

Besides the 9 residues that were digested off by enterokinase, there are 10 more residues from the C-terminus that do not have clear electron density. Based on its susceptibility to protease digestion and lack of clear electron density, it is reasonable to assume that these residues at the C-terminus of the protein are flexible, and do not have a defined structure in solution. It is clear from the structure that both N- and C-termini are away from the active site cavity, and therefore are unlikely to affect the enzyme catalytic activity. This is consistent with the results of enzyme activity assays [10].

4. STRUCTURE OF THE MTB PANTOTHENATE SYNTHETASE

4.1 Dimer Structure

The crystal structure of the MTB PS enzyme indicates that it is a dimer (Figure 1), resembling a butterfly when viewed from one side. The dimer interface is extensive, having a buried area of about 2150 Å². A cluster of 8 hydrophobic side chains at the center of the dimer interface, Phe174, Leu177, Leu144, and Val118, form the core of the dimer interaction. This cluster is surrounded by hydrogen bonds and salt bridges, as well as the specific β strand
interactions of the two-stranded intersubunit β sheet that is composed of the strand β5 from each subunit. The MTB PS enzyme also exists in solution as a dimer [8, 10], as do PS enzymes from higher plants [6, 7].

The two subunits of the dimer in the P2₁ crystal are similar to each other, with an rmsd of 0.9 for all Cα atoms except the disordered residues. Most of the deviation lies at residues 74 to 88, and residues 260 to 265, where the crystal packing environments are different. Residues 74 to 88 (β3-3'-α3'-α3) form one wall of the active site cavity (see Figure 2 and the active site cavity section b ). In subunit A these residues are ordered with residues 74 to 76 having high B factors, while in subunit B residues 75 to 80 (helix α3' and nearby residues) do not have electron density, and the rest of the residues up to residue 88 are partially disordered with high B factors. This flexible wall of the active site cavity becomes ordered in the complex with the pantoyl adenylate reaction intermediate, suggesting their function as a gate to the active site cavity (see below pantoyl adenylate binding interactions).

Figure 1. Ribbon diagram of the M. tuberculosis pantothenate synthetase dimer. (a) A side view of the dimer structure shows that it resembles the shape of a butterfly. (b) An orthogonal view of (A) from top, with the two-fold NCS symmetry axis (labeled with a dot) approximately perpendicular to the paper plane. Secondary structure elements for subunit A (left) are labeled. Those for subunit B are identical except that the short helix α3' is disordered. The figure was prepared from the apo-enzyme structure of the P2₁ crystal, with Molscript [17] and Raster3D [18].
Snapshots of the Pantothenate Synthetase from *Mycobacterium Tuberculosis* along the Reaction Coordinate

The C2 crystal differs from the P2₁ crystal in that there is one polypeptide chain per asymmetric unit [11]. However, the protein molecule in the C2 crystal is also present as a dimer that is essentially identical to that in the P2₁ crystal, but the two-fold axis of the dimer coincides with one crystal symmetry axis. In the C2 crystal, a larger portion of the flexible wall, from residue 76 to 86, is completely disordered. The first turn of helix α₃ unwinds, and helix 3₁₀³, which contains residue Glu⁷₂ that is important for binding pantoate (see below pantoate binding site), also unwinds and forms an extended structure. This natural difference is induced by crystal packing and the intrinsic flexibility of the loop.

4.2 Domain Interactions
Each subunit of the dimer has two well-defined domains (Figure 1). The N-terminal domain has a Rossmann fold, with a central parallel β sheet composed of strands β₁ to β₄, β₆, and β₇, and helices on both sides of the β-sheet. Strand β₇ leads to the C-terminal domain, which starts with a hairpin loop followed by helix 3₁₀⁶. The rest of the C-terminal domain has a simple two-layer structure: a layer of helices α₇, α₈, and α₉ followed by a layer of antiparallel β sheet composed of strands β₈ to β₁₀.

The two domains have extensive interactions. There is a hydrophobic core composed of residues Leu₁₂₃, Leu₁₂₇, Tyr₁₆₂, Val₁₆₆ in the N-terminal domain and residues Tyr₂₄₉, Leu₂₅₇, and Leu₂₆₉ in the C-terminal domain. Surrounding this hydrophobic core are three salt bridges, Glu₁₅₉ to Arg₂₆₇, Glu₁₂₈ to Arg₂₇₈, Glu₁₂₆ to Arg₂₅₃, and 6 additional hydrogen bonds linking the two domains. The relative positions of the two domains stay the same for the two subunits in the dimer and for all crystals including the P₂₁, P₂₁₂₁, and C2 crystals. These structural features suggest that domain hinge motion is unlikely to be the mechanism of opening and closing the active site cavity for MTB PS.

4.3 Active Site Cavity
Typical for nucleotide binding (Rossmann fold) proteins, the active site is located at the C-terminus of the central parallel β sheet. The cavity is large and deep, and its bottom is at a cleft between strands β₂ and β₆ (Figure 2). The segment between β₃ and α₃ (loop-3₁₀₃-loop-α₃'-loop region) forms one side of the walls surrounding the cavity in subunit A of the P₂₁ crystal. Part of this loop is disordered in subunit B, thus opens up the active site cavity. Other walls of the cavity are from β₂-loop-α₂, 3₁₀₅'-loop-α₅, β₆-loop-α₆, and the loop after β₇. Helix 3₁₀⁷ and the β strands from the C-terminal domain partially cover the top of the active site cavity. The bottom of the active site cavity is mainly hydrophobic, while the top half of the cavity has several charged residues, including His₄₄ and His₄₇ at the N-terminus of α₂ (the HIGH motif [5, 19]), Lys₁₆₀ and Asp₁₆₁ at the N-terminus of α₆, and Arg₁₉₈ of 3₁₀⁷ (KMSKS motif [5]). At the entrance to the active site cavity, several charged side chains stick to the center,
leaving only a small opening to the cavity (Figure 2, right panel). Four arginine side chains, including Arg198, form a patch of positive charges at the entrance. These positive charges may steer the negatively charged substrates into the active site cavity. As we will see later, Arg198 plays an important role in binding the phosphate groups of ATP.

4.4 ATP Binding Interactions

The active site is well designed for binding ATP. Figure 3 shows the initial difference electron density for the ATP molecule in the active site of the C2 crystal. The ATP molecule has an average temperature factor of 17.3 Å², virtually identical to the average temperature factor of the surrounding protein atoms, indicating that it is at full occupancy. A magnesium ion was found binding to the phosphate groups of ATP. It has a nearly perfect octahedral coordination. Its distances to the O2α, O2β, and O1γ atoms of the phosphate groups are 2.26 Å, 1.99 Å, and 2.12 Å, respectively; three other ligands are water molecules with distances of 1.96 Å, 2.11 Å, and 2.28 Å. The adenosine group of ATP binds tightly at the bottom of the active site cavity with hydrogen bonds and hydrophobic interactions [10, 11]. The phosphate groups stick toward the top of the active site cavity, and are located near the N-terminal ends of helices α2 and 3₁₀. Binding of the phosphate groups involves mainly main chain NH groups and charged side chains.

Two His residues of the HIGH motif [5, 19, 21] interact with the phosphates of ATP. The first histidine, His44, has a hydrogen bond extending from its Nε2 to O1β of β-phosphate. The His47 side chain has a hydrogen bond from its Nε2 to the bridge oxygen between the α- and β-phosphate, suggesting that it functions as a general acid by donating a proton to the leaving pyrophosphate group. Both histidine residues have the Nδ1 atom forming a hydrogen bond to a main chain NH group, and thus the Nε2 atoms are hydrogen bond donors. Their imidazole rings are in close proximity to each other with the closest separation being ~3.2 Å, indicating favorable π-π 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 (closest distances of ~3.4 and ~4.3 Å for His44 and His47, respectively). Mutation of either histidine residue to alanine causes a greater than 1000-fold reduction in enzyme activity [22]. The Gly46 residue is highly conserved in this motif. It flanks one side of adenine (with Lys160 on the loop after β6 on the other side), where any side chain larger than a hydrogen atom would have steric hindrance. The KSMKS motif [23] is not conserved at the sequence level in PS enzymes, but it is evident from structural alignments [5]. Residues Ser196, Ser197, and Arg198, which are at the N-terminus of helix 3₁₀ and are part of the KSMKS motif, bind the β- and γ-phosphate groups. The side chains of Lys160 and Arg198 are partially disordered in the apo-enzyme structure, but become ordered and move to enable salt-bridges to the β- and γ-phosphate groups, respectively. The side chain of Lys160 also packs against one side of adenine. Mutation of this
residue to alanine reduces the enzyme activity more than 1000-fold, and it decreases affinity of the enzyme for ATP [22]. 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.

The binding interactions of ATP with protein are similar to those of AMPCPP [10]. However, the methylene group in AMPCPP replaces the bridge oxygen between the $\alpha$- and $\beta$-phosphate groups of ATP, and thus it is not capable of forming a hydrogen bond with the Nε2 atom of His47. Consequently, the $\alpha$-phosphate moves slightly away from amide nitrogen of Met40, and thus no hydrogen bond is formed between them in the AMPCPP complex. Therefore, ATP has slightly higher binding affinity to the enzyme active site than its non-reactive analog AMPCPP.

### 4.5 Pantoate Binding Site

In the P2$_1$ crystal that was soaked with pantoate and $\beta$-alanine, there is a full occupancy of pantoate in both active sites [10]. The binding site for pantoate is located in a pocket at the bottom of the active site cavity, next to the binding site of the adenosine group (Figure 4a). This is the same binding site proposed for the E. coli PS protein [5], although the pantoate molecule has a different conformation and therefore different binding interactions. The pantoate molecule is tightly bound at the bottom of the active site cavity, with side chains of two glutamine residues, Gln72 from helix 3$_{10}$3 and Gln164 from helix $\alpha$3, forming hydrogen bonds to its hydroxyl groups and one carboxyl oxygen. The two methyl groups face the hydrophobic groups at the bottom of the active site cavity, Pro38 and Phe157 side chains. The side chain of Met40 packs against the hydrophobic side of the pantoate molecule. Mutation of either glutamine residue to alanine greatly reduces the rate of pantoyl adenylate formation [22].

Soaking P2$_1$ crystals in solutions containing both AMPCPP and pantoate gave a full occupancy of AMPCPP, but with a glycerol in the pantoate binding site in subunit B, which has the flexible wall (residues 75-88) disordered. However, subunit A has a partial occupancy of both AMPCPP and pantoate, as indicated by their electron density and B factors. Increasing the concentration of pantoate in the soaking solution gave only a slightly higher occupancy of pantoate but a lower occupancy of AMPCPP [10]. All these observations indicate that AMPCPP and pantoate cannot coexist in the active site. This must be due to the steric and charge repulsions between the $\alpha$-phosphate group and the carboxyl group of pantoate. Modeling both ATP and pantoate in one active site based on the fully occupied complexes will put one carboxyl oxygen atom of pantoate ~2.7 Å from the $\alpha$-phosphorus atom of ATP [11]. This is essentially identical to what is observed in subunit A that has partial occupancy of both AMPCPP and pantoate. The carboxyl oxygen of pantoate is in a good position for an in-line nucleophilic attack on the $\alpha$-phosphate. Because residues of the flexible wall (gate residues) are disordered in subunit B, the active site cavity is open and allows AMPCPP to diffuse easily into the active site. The extensive binding interactions of AMPCPP to the protein atoms make its binding thermodynamically more favorable than the binding of pantoate to the protein. Therefore, AMPCPP has a full occupancy in the active site cavity of subunit B of the complex with both AMPCPP and pantoate. On the other hand, the gate residues are less flexible in the subunit A due to crystal packing, thus making it kinetically less favorable for AMPCPP to bind, and allowing pantoate to have a partial occupancy in the active site of subunit A.

In the C2 crystal, the side chain of Gln72 on helix 3$_{10}$3 swings out of the active site due to crystal packing, and the binding of pantoate is abolished. We observed only ATP in the active site with a full occupancy (see above ATP binding interactions) when we soaked the C2 crystal in a solution containing both ATP and pantoate [11]. There was no electron density for pantoate or pantoyl adenylate in the active site cavity. In the pantoate binding pocket, there was positive difference electron density, into which we modeled a glycerol molecule. This glycerol has a high temperature factor, suggesting that it has a low occupancy. Except for this Gln72 side chain, the rest of the binding site for pantoate is intact. In the P2$_1$ crystal, soaking with both ATP and pantoate results in a pantoyl adenylate intermediate in the active site cavity, and the disordered loop in subunit B becomes ordered [10] (see more details below). This 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 the reaction intermediate. Thus mutation of either residue could affect the stability of pantoyl adenylate intermediate. It is noteworthy that Asn69 was also found to be important for the enzyme activity [22]. 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, since both Asn69 and Gln72 are on
Shuishu Wang and David Eisenberg

the one-turn $3_{10}$ helix, mutation of Asn69 might also affect the stability of this $3_{10}$ helix and hence the position of the Gln72 side chain.

**Figure 4.** (a) Binding interactions of AMPCPP and pantoate in the active site of subunit A of the P2 crystal soaked in a solution containing both AMPCPP and pantoate. Both molecules have a partial occupancy (see text for details), and the magnesium ion is not well defined in this active site. However, binding position and interactions are essentially identical to those of AMPCPP and pantoate with a full occupancy when each is bound in the active site alone. The distance between the $\alpha$-phosphorus atom and the carboxyl oxygen is ~2.96 Å, shown in blue dashed line. (b)
Snapshots of the Pantothenate Synthetase from *Mycobacterium Tuberculosis* along the Reaction Coordinate

A model of the trigonal bipyramidal intermediate based on crystal structures of the ATP complex, AMPCPP complex and pantoate complex. Formation of the intermediate requires only the movement of the α-phosphorus atom towards the carboxyl oxygen. All favorable interactions with active residues are preserved. His47 is poised to function as a general acid.

4.6 Binding Interactions of the Reaction Intermediate, Pantoyl Adenylate

Soaking the P2₁ crystals with solutions containing both ATP and pantoate, or crystallization in the presence of both substrates, resulted in a reaction intermediate, pantoyl adenylate, in the active site [10]. Cocrystallization resulted in both active sites in the dimer being fully occupied with the pantoyl adenylate molecule. However, crystals soaked with ATP and pantoate have one active site (subunit B) fully occupied but the other partially occupied, indicated by their difference in B factors. Subunit A of the dimer also has a lower occupancy of AMPCPP when crystals were soaked in solutions containing AMPCPP. This is due to crystal packing, which makes the flexible wall of the active site cavity (3₁₀₇-loop-α₃’-loop region) ordered in subunit A and thus the active site less accessible for bulky molecules ATP and AMPCPP.

Pantoyl adenylate has extensive binding interactions with the active site residues (Figure 5). The molecule is almost linear and fits snugly in the bottom of the active site cavity. Both pantoate and adenylate moieties sit in the same positions in the active site as the pantoate molecule of the pantoate complex and the adenosine group of ATP in the ATP complex, and preserve all strong interactions with protein atoms. The tight binding of pantoyl adenylate stabilizes this highly reactive intermediate. One important feature of the pantoyl adenylate complex is that the disordered loop in subunit B becomes ordered, in both cocrystallization and soaking the apo-enzyme crystals with both ATP and pantoate. This suggests that the flexible wall can function as a gate, opening for substrates, especially the bulky ATP molecule, to bind and closing the active site cavity to protect the highly reactive pantoyl adenylate.

![Figure 5](image-url)

**Figure 5.** Binding interactions of the pantoyl adenylate intermediate. The intermediate binds tightly at the bottom of the active site cavity with many hydrogen bonds and hydrophobic interactions. Hydrogen bonds to the O₃* of ribose and some water mediate hydrogen bonds are not shown for clarity.

4.7 Binding Interactions of Beta-Alanine

Cocrystallization and soaking crystals with β-alanine [10] failed to yield a complex with this substrate, suggesting that the binding site for β-alanine exists only after pantoyl adenylate is formed in the active site. Based on the structure of the pantoyl adenylate complex, it is likely that the phosphate group of pantoyl adenylate serves as an anchor for the initial binding of β-alanine by offering hydrogen-bonding partners and/or favorable charge-charge interactions. Furthermore, an AMP molecule bound at the active site may also assist the binding of β-alanine from
its phosphate group. Therefore, we soaked crystals of PS in a solution containing both β-alanine and AMP, and we obtained a crystal structure of the PS enzyme in complex with β-alanine and AMP from the P2₁ crystals [11].

Figure 6. (a) Structural superposition of the active site of the pantoyl adenylate complex with that of the β-alanine/AMP complex. The active site residues align well with an rmsd of Ca less than 0.15 Å. The pantoyl adenylate molecule is colored in cyan. The protein structure shown in the figure is from the β-alanine/AMP complex. Hydrogen bonds to the phosphate group of AMP and to β-alanine are shown in yellow dashed lines. Some water mediated hydrogen bonds are not shown for clarity. The Tyr82 side chain is from a loop in the front that is clipped away to reveal the bound AMP and β-alanine. Based on the structural alignment, the amino group of β-alanine is ~2.8 Å from the phosphate oxygen of pantoyl adenylate, and ~3.4 Å from the carbonyl carbon, shown in blue dashed lines. (b) Model of the tetrahedral intermediate structure based on the structural superposition in (A). The Tyr82 side chain is not shown for clarity. A tetrahedral geometry can be obtained by a torsional rotation of the amino nitrogen of β-alanine to move it to ~1.5 Å above the carbonyl carbon. All favorable interactions with protein atoms are preserved, and thus the intermediate is stabilized by the enzyme.
The β-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 6a). 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 β-alanine has one hydrogen bond to the Na2 atom of the Asp161 or Tyr82 side chain. This carboxyl is also close to the side chains of Arg198 and His135, with distances of ~4.0 Å, respectively, indicating favorable charge-charge interactions and π-electron interactions. One side of β-alanine faces side chains of Met40 and Tyr82; the other side faces the large cavity of the active site, which is filled with a few ordered water molecules in the crystal structure. Binding of β-alanine is not as tight as that of other substrates and its binding site in the crystal is not fully occupied. However, the binding interactions limit molecules that can bind well and have a nucleophilic attack on the reaction intermediate. Therefore, structural analogs of β-alanine are all poor substrates [8].

In subunit B of the dimer, where residues 74 to 83 (including Tyr82) are disordered, only AMP but no β-alanine was found in the active site [11]. Also in this case, disorder of this active site cavity wall weakens binding interactions for β-alanine because its binding site is exposed to bulk solvent and thus the charge-charge interactions are dampened. Similarly, in the C2 crystal, which has one molecule per asymmetry unit and the flexible loop disordered, only AMP was found in the active site when crystals were soaked in a solution containing both AMP and β-alanine.

Superposition of the structure of the β-alanine/AMP complex with that of the pantoyl adenylate complex (Figure 6a) shows that the amino group of β-alanine was positioned above the carbonyl group of pantoyl adenylate such that the nitrogen atom was ~3.4 Å from the carbonyl carbon. The O1a atom of pantoyl adenylate can form a hydrogen bond with the amino group of β-alanine, serving as an anchor for the initial binding of β-alanine. With changes in only one torsion angle, the amino group of β-alanine can move towards the carbonyl carbon for a nucleophilic attack and attain a nearly perfect geometry for a tetrahedral intermediate (Figure 6b).

### 4.8 AMP Binding Interactions

As described above, we obtained AMP complexes from both P21 and C2 crystals by soaking crystals in a solution containing both AMP and β-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 the pantoyl adenylate, AMPCP, 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 α2 and Lys160 on the loop after β6. 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 α-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 6a).

### 5. OVERALL ENZYME-CATALYZED REACTION MECHANISM

Putting all structural data together allows us to have a detailed view of the overall enzyme-catalyzed reaction mechanism. When the active site is empty, the disordered loop opens the active site and thus allows the ATP molecule to come in and bind. Because ATP is relatively bulky and is held rigidly in the active site by many binding interactions, it is likely that ATP binds first. The smaller sized pantoate then comes in, and initiates a nucleophilic attack on the α-phosphate while it binds in its binding pocket [8, 24]. When both ATP and pantoate are present in the same active site, the nucleophilic reaction must occur (Figure 4a). Formation of the transient trigonal bipyramidal intermediate involves only the movement of the phosphate atom towards the carboxyl oxygen of pantoate to be coplanar with the three equatorial oxygen atoms (Figure 4b). Although transient in nature, this intermediate is stabilized by the enzyme because all the favorable interactions are preserved, and the unfavorable steric hindrance and charge repulsions between ATP and pantoate are eliminated. The trigonal bipyramidal intermediate then dissociates, and simultaneously the pyrophosphate group leaves the active site. The magnesium ion and the positively charged side chains around the β- and γ-phosphate groups draw the negative charges towards the leaving pyrophosphate. His47 is also likely to facilitate the reaction by donating a proton to the leaving pyrophosphate. This leaves a reaction intermediate, pantoyl adenylate in the active site and completes the first half of the enzyme catalyzed reaction. Pantoyl adenylate is stabilized by many strong binding interactions, as we have seen earlier. It is also protected by the closing of the active site when the flexible loop becomes ordered. Pantoyl adenylate is stable only when tightly bound in the enzyme active site. In solution, it decomposes rapidly to give...
pantoyl lactone and AMP [9]. 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 that $^{18}$O labels scrambled when incubating $[^{18}$O$_{6}$]-ATP and pantoate with the MTB PS enzyme [24].

The presence of pantoyl adenylate intermediate and the ordering of the flexible wall create 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. The phosphate group of pantoyl adenylate can have a hydrogen bond to the amino group of $\beta$-alanine, thus serving as an initial anchor for binding $\beta$-alanine. The carboxylate group of $\beta$-alanine has favorable interactions with the side chains of Arg198 and His135 and a hydrogen bond to Glu72. There is a shift of the Arg198 side chain of ~1.8 Å towards the carboxylate group of $\beta$-alanine relative to that in the pantoyl 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 pantoyl group. The distance of the amino nitrogen is only ~3.4 Å from the carboxyl carbon of pantoyl adenylate. A model of the tetrahedral intermediate can be readily prepared from the crystal structures of the pantoyl adenylate complex and the $\beta$-alanine/AMP complex (Figure 6b).

Formation of the tetrahedral intermediate requires only a torsional rotation of the amino nitrogen atom of $\beta$-alanine toward the carbonyl carbon of pantoyl adenylate. Similar to that of the trigonal bipyramidal intermediate, formation of the tetrahedral intermediate preserves all favorable binding interactions of active site residues with pantoyl adenylate and $\beta$-alanine. Therefore, the intermediate is stabilized by the enzyme. Dissociation of the tetrahedral intermediate then forms pantothenate and AMP. The planarity of the peptide bond of pantothenate causes significant rearrangement of the molecule, and several hydrogen bonds to active site residues are broken. In addition, the newly formed pantothenate has steric clashes and charge repulsions with AMP and some active site residues. Therefore, pantothenate must leave the active site once it is formed. The AMP molecule then diffuses away to regenerate the enzyme active site for next cycle of catalysis. Pantothenate has very low binding affinity to 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 has a low concentration inside cells, and thus can easily diffuse out of the active site. Moreover, ATP has many favorable interactions in the active site of the enzyme, and it can readily displace AMP.

6. REFERENCES

Snapshots of the Pantothenate Synthetase from *Mycobacterium Tuberculosis* along the Reaction Coordinate