Degradation Pathway of the Phosphonate Ciliatine: Crystal Structure of 2-Aminoethylphosphonate Transaminase[†,‡]

Celia C. H. Chen,[§] H. Zhang,[§] A. D. Kim,[∥] A. Howard,[⊥] G. M. Sheldrick,[@] D. Mariano-Dunaway,[∥] and O. Herzberg*[,§]

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, Maryland 20850, Department of Chemistry, University of New Mexico, Albuquerque, New Mexico 87131, Illinois Institute of Technology, Chicago, Illinois 60616, Advanced Photon Source, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439, and Department of Structural Chemistry, University of Göttingen, Tammannstrase 4, 37077 Göttingen, Germany

Received June 3, 2002; Revised Manuscript Received August 4, 2002

ABSTRACT: Phosphonates allow certain organisms to thrive in otherwise hostile environments, and 2-aminoethylphosphonate (AEP) is a precursor of many cellular phosphonates. AEP transaminase (AEPT) is an enzyme essential to phosphonate synthesis and degradation pathways. The crystal structure of AEP transaminase was determined by multiwavelength anomalous diffraction of 66 selenium atoms. The refined structure at 2.2 Å resolution revealed an overall fold and active site location similar to those of the dimeric, two-domain structure of type I aminotransferases. The active site contains a cofactor, pyridoxal 5′-phosphate (PLP), and the product phosphonoacetaldehyde. Comparison with other type I aminotransferase structures shows that the PLP−protein interactions are conserved. Modeling of bound substrates and products reveals the structural basis for AEP recognition and the stereospecificity of proton elimination at the α-carbon and indicates conformational changes along the reaction pathway.

Phosphonates are utilized by certain organisms to serve in specialized functions which include chemical warfare, phosphorus storage, cell−cell communication, and host recognition. 2-Aminoethylphosphonate (AEP),[1] the most common naturally occurring phosphonate (*1*−*5*) also known as ciliatine, is an important precursor used in the biosynthesis of phosphonolipids, phosphonoproteins, and phosphonoglycans (*6*−*9*). The subject of this report, AEP transaminase (AEPT, EC 2.6.1.37), couples the interconversion of AEP and phosphonoacetaldehyde (P-Ald) with the interconversion of pyruvate and L-alanine (L-Ala).

AEP is resistant to acid- and base-catalyzed hydrolysis as well as to the action of phosphotransferases (*10*). For utilization of AEP as a source of nitrogen, carbon, and phosphorus, the soil-dwelling bacteria have acquired an AEP degradation pathway. In this two-step pathway, AEP is first converted to P-Ald through the action of AEPT (Figure 1) and then to orthophosphate and acetaldehyde through the action of phosphonatase. This pathway has been studied mainly in *Pseudomonas aeruginosa* (*11*), *Bacillus cereus* (*12*), and *Salmonella typhimurium* (*13*), although it is present in other bacteria as well (*14*, *15*).

The closest known sequence homologue of AEPT is the putative acetone phosphate transaminase of the rhizobitoxine biosynthetic pathway (*16*), the sequence of which is 30% identical. Rhizobitoxine is a nodulation enhancer produced by the legume symbiont *Bradyrhizobium alkanii* and plant pathogen *Pseudomonas andropogonis* (*17*, *18*). The high level of sequence identity between AEPT and this transaminase suggests a common ancestor that was drawn upon more than once to provide a new transaminase for an adaptive pathway. The closest homologue with known structure is the phosphoserine aminotransferase (PSAT) from *Escherichia coli* with a level of sequence identity of only 17.5% (*19*).

Like many other aminotransferases, AEPT is a pyridoxal 5′-phosphate (PLP)-dependent enzyme (*11*, *20*). In these enzymes, PLP serves as a midway station for the transferred amino group by forming alternately an internal aldimine with the enzyme and an external aldimine with the substrate (Figure 1). Reported structures of aminotransferases (*21*−*25*) provide a rich context in which to view the crystal structure of AEPT from *S. typhimurium* that is described here.

EXPERIMENTAL PROCEDURES

Protein Production and Purification. The Se-Met-containing protein of *S. typhimurium* AEPT was prepared and purified according to established procedures (*20*, *26*). Electron spray mass spectrometry was used to confirm that all 11 Met residues (the N-terminal Met is removed post-translationally) were replaced with Se-Met. The steady-state

[†] Supported by the NSF Grant MCB9813271 (O.H.) and NIH Grant GM28688 (D.M.-D.).

[‡] Protein Data Bank entry 1M32.

* To whom correspondence should be addressed. Telephone: (301) 738-6245. Fax: (301) 738-6255. E-mail: osnat@carb.nist.gov.

[§] University of Maryland Biotechnology Institute.

[∥] University of New Mexico.

[⊥] Illinois Institute of Technology and Argonne National Laboratory.

[@] University of Göttingen.

[1] Abbreviations: AAT, aspartate aminotransferase; AEP, 2-aminoethylphosphonate; AEPT, 2-aminoethylphosphonate transaminase; GAM, methyl glutamate; MmPEG5K, monomethyl polyethylene glycol 5000; PAH, phosphonoacetohydroxamate; P-Ald, phosphonoacetaldehyde; PLP, pyridoxal 5′-phosphate; PMP, pyridoxamine 5′-phosphate; PSAT, phosphoserine aminotransferase; Se-Met, selenomethionine.

10.1021/bi026231v CCC: $22.00
Published on Web 10/11/2002

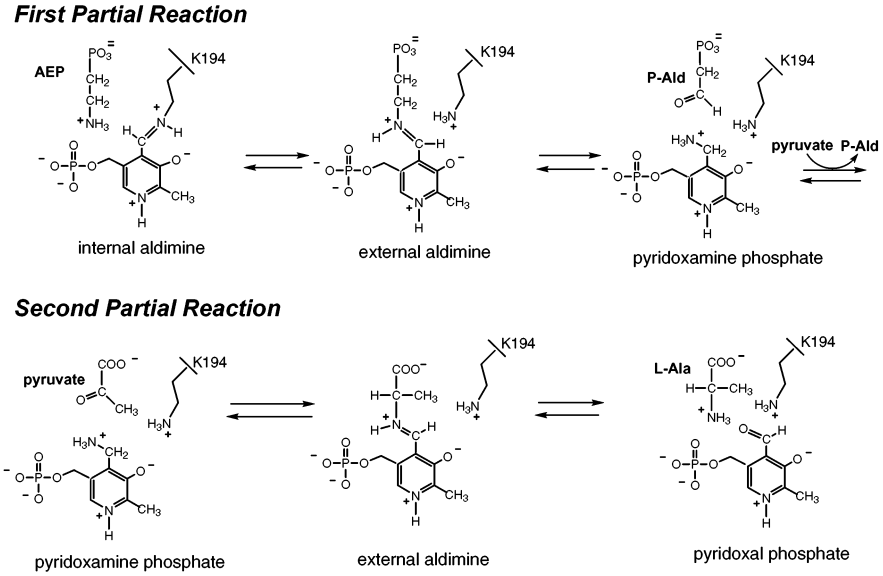

FIGURE 1: Scheme of the AEPT reaction sequence. PLP and AEP atom numbers used in the text are shown.

kinetic constants measured for Se-Met AEPT are similar to those measured for the native recombinant AEPT (*26*), thus indicating that the substitution of Se-Met residues for the Met residues in AEPT had not altered catalytic functioning.

Crystallization. Crystals of Se-Met-containing AEPT were obtained at 4 °C by the vapor diffusion method in hanging drops. Protein drops were equilibrated against a reservoir solution containing 0.2 M ammonium acetate, 0.1 M sodium citrate (pH 5.0), and 10−13% monomethyl polyethylene glycol 5000 (mmPEG5K). The protein solution contained 0.45 mM (18 mg/mL) of protein/PLP, 8 mM phosphate buffer (pH 7.5), 0.8 mM DTT, and 20 mM phosphonoacetaldehyde. The hanging drops consisted of approximately equal volumes of protein and reservoir solutions. Yellow crystals in the form of thin plates with a longest dimension of up to 0.5 mm were obtained within 2 weeks. For data collection, the crystals were mounted in cryoloops and briefly dipped in solutions containing the reservoir solution except that the concentration of mmPEG5K was increased to 15%, and 20% glycerol was added as a cryoprotectant. The mounted crystals were flash-cooled in liquid propane at near-liquid nitrogen temperature. The crystals belong to space group $P2_1$, with the following unit cell dimensions: $a = 49.6$ Å, $b = 156.4$ Å, $c = 170.2$ Å, and $\beta = 90.7°$.

X-ray Data Collection. Two MAD data sets (AEPT5 and AEPT27) exploiting the absorption edge of selenium [inflection point (*i*), peak (*p*), and high-energy remote (*r*) wavelengths] and a single wavelength data set (AEPT4) at a high-energy remote wavelength were collected at 100 K. All data were processed with the program HKL (*27*) (Table 1).

Self-Rotation Search for Noncrystallographic Symmetry Elements. The AMORE self-rotation function (*28*) showed 2-fold axes both parallel and perpendicular to the unit cell *a* axis, but none parallel to the crystallographic 2-fold screw axis. Three-, four-, or sixfold axes were not found.

Determination of Se Positions. Development versions of the program Xprep and ShelxD were used to assess the data quality and determine the position of Se atoms (*29*, *30*) with AEPT5 data. The correlations between the anomalous signal in the three-wavelength analysis were very high which indicated that the data were of excellent quality, and useful to the limit of the measured resolution. The calculated f_a values were used in ShelxD to successfully determine all 66 Se sites in the asymmetric unit. Because of the high quality of the MAD data and the large number of Se sites in the asymmetric unit, these data have been used as a test case for improving phase determination computer programs (*31*).

A set of 66 sites was identified. The correlation coefficient was ≥0.605 for correct solutions; the correlation coefficient of the highest false solution was 0.309. The peak values of the 66 correct solutions were clearly higher than that of the 67th solution corresponding to the first noise peak.

Examination of the 66 positions in real space revealed that they could be separated into six groups of 11 positions related by noncrystallographic 2-fold symmetry rotation axes. The average rmsd between sets of 11 positions was in the range of 0.2−0.3 Å.

The program Shake and Bake (SnB2) (*32*) was also used in an attempt to determine the locations of Se sites. In our hands, with a search for 44 or 88 sites, with all default settings, there was no discrimination between the correct 66 sites and incorrect sites. When 88 sites were requested, the 66 correct sites ranked 1−64, 67, and 68 with no significant drop in peak heights for the next group of peaks. Later studies showed that if the default values were fine-tuned, it was also possible to find the 66 sites using SnB2 (*31*). Moreover, determination of all 66 sites is not necessary to produce phases that lead to an interpretable electron density map.

The position of the 66 Se sites enabled accurate determination of the noncrystallographic symmetry axes, revealing that the asymmetric unit contains three dimers. The packing of the molecules in the asymmetric unit was unusual, showing three dimers with one dimer roughly perpendicular to the other two, confirming the findings of the self-rotation search.

Structure Determination and Model Building. The MAD data set AEPT27 had a resolution (2.3 Å) that was higher than that of AEPT5 (2.55 Å) (Table 1). To facilitate automatic model building, phase refinement continued with the AEPT27 data. The electron density map obtained after MLPHARE refinement (*33*) on the occupancies of the Se

Table 1: Data Collection and Phasing Statistics

	AEPT5[a]			AEPT27[b]			AEPT4[a]
space group	$P2_1$			$P2_1$			$P2_1$
cell dimensions	$a = 49.52$ Å			$a = 49.60$ Å			$a = 49.59$ Å
	$b = 155.60$ Å			$b = 156.37$ Å			$b = 155.26$ Å
	$c = 169.92$ Å			$c = 170.20$ Å			$c = 168.63$ Å
	$\beta = 90.62°$			$\beta = 90.67°$			$\beta = 90.57°$
no. of molecules in the asymmetric unit	6			6			6
	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3	λ
MAD data statistics							
wavelength (Å)	0.9790	0.9791	0.9500	0.9787	0.9791	0.9500	1.0000
resolution (Å)	2.55	2.55	2.55	2.28	2.28	2.28	2.15
no. of observed reflections	478023	510180	511438	406115	406338	389833	496002
no. of unique reflections	154141	154828	155653	211943	212073	214397	264452
completeness (%)[c]	93.9 (72.2)	94.1 (69.1)	94.7 (79.2)	90.7 (76.2)	90.7 (76.3)	91.8 (86.4)	96.8 (83.0)
R_{merge} (%)[d]	3.3 (8.3)	3.3 (8.8)	3.4 (6.2)	4.5 (12.9)	4.5 (13.1)	3.9 (12.1)	3.7 (15.2)
$\langle I/\sigma \rangle$	19.6 (15.9)	19.3 (13.6)	19.3 (17.2)	6.9 (2.2)	8.1 (2.5)	9.6 (3.0)	15.5 (5.7)
phasing statistics							
phasing power[e]	–			1.61 (0.91)	1.19 (0.66)	–	
dispersive R[e]				0.71 (0.98)	0.80 (0.97)	- - -	
anomalous R[e]				0.77 (0.98)	0.70 (0.99)	0.77 (0.98)	
figure of merit				0.483			

[a] Collected at the Advanced Photon Source (National Argonne Laboratory) on the IMCA-CAT 17-ID beamline with with Mar 165 mm charged couple device. [b] Collected at the National Synchrotron Light Source (Brookhaven National Laboratory, Upton, NY) on beamline X12C with a Brandeis 2 × 2 CCD detector. [c] Values in parentheses are for the highest-resolution shell: 2.64−2.55 Å for AEPT5, 2.36−2.28 Å for AEPT27, and 2.23−2.15 Å for AEPT4. [d] $R_{merge} = \sum_{hkl}[(\sum_j |I_j - \langle I \rangle|)/\sum_j |I_j|]$. [e] Phasing power $= \sum_j |F_H|/\sum_j |E|$. Dispersive $R = \sum_{hkl}||F_P + F_H(\text{calcd})| - F_{PH}|/\sum_{hkl}|F_{PH} - F_P|$. Anomalous $R = \sum_{hkl}|F_{obs}^{\pm} - F_{cal}^{\pm}{}_c|/\sum_{hkl}|F_{obs}^{\pm}|$. E is the lack of closure error. F_P is the reference data set λ_1. F_{PH} represents the data collected at λ_2 or λ_3. $F^{\pm}$ is the structure factor difference between Friedel pairs. The values in parentheses are for the highest-resolution shell (2.58−2.28 Å).

atoms and DM solvent flattening (*34*) was of excellent quality.

Three consecutive runs of the automated polypeptide chain building program ARP/wARP (*35*, *36*) yielded 114 chains with a total of 1436 residues in the entire unit cell, and a connectivity factor of 0.81. Application of the crystallographic and noncrystallographic symmetry operators coalesced the 114 chains into eight chains corresponding to one molecule. This initial model included 319 of the 367 residues in the molecule (87%). The automated side chain building of ARP/wARP generated 65 side chain residues. The remaining molecule was built manually using the computer graphics program Turbo-Frodo (*37*).

Refinement. The structure was refined by CNS (*38*), using the highest-quality data from crystal AEPT4 (Table 1). One cycle of simulated annealing molecular dynamics at 2500 K was followed by alternating positional and individual temperature factor refinement cycles. Maximum likelihood targets were used (*39*). The resulting models were inspected and modified using the computer graphics program Turbo-Frodo. The six molecules in the asymmetric unit were refined independently. Water molecules, PLP, P-Ald, and phosphate ions were added as the refinement progressed and the electron density maps improved. The topology and parameter files for PLP and phosphonoacetohydroxamate (PAH) were obtained from the hetero compound database (*40*) at the HIC-UP site (http://xray.bmc.uu.se/hicup), and the PAH files were modified for P-Ald. The occupancies of 140 disordered side chain atoms were set to zero, and alternate conformations were included for 14 side chain residues. Water molecules were selected automatically in the difference electron density map using 3σ cutoff criteria. Solvent molecules in the active site were incorporated only at the final stages of the refinement. The final statistics are presented in Table 2.

Table 2: Refinement Statistics

space group	$P2_1$
resolution (Å)	20−2.2
wavelength (Å)	1.00
no. of unique reflections where $F \geq 2\sigma(F)$	127266
completeness (%)	98.6
no. of protein atoms	16698
no. of nonprotein atoms	3056 H_2O, 6 PLP, 2 P-Ald, 4 PO_4
R_{cryst} (%)[a,b]	16.6 (18.8)
R_{free} (%)[a,c]	20.0 (23.6)
R_{total} (%)[a,d]	17.1 (19.8)
rms deviation from ideal geometry	
bond lengths (Å)	0.005
bond angles (deg)	1.2
average B-factor (Å^2)	
protein main chain	19
protein side chain	20
H_2O	37
PLP	23
P-Ald and phosphates	69
Ramachandran plot (%)	
most favored	91.6
allowed	8.0
generously allowed	0.3
disallowed	0.0

[a] The values in parentheses are for the highest-resolution shell (2.3−2.2 Å). [b] $R_{cryst} = \sum_{hkl}||F_o| - |F_c||/\sum_{hkl}|F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively. [c] R_{free} is computed from 12 668 (9.8%) reflections. [d] R_{total} is computed from 128 378 reflections, with all data included.

RESULTS

Quality of the Model. The final model with a final R of 0.166 (R_f of 0.200) (Table 2), refined at 2.2 Å using the data designated as AEPT4 (Table 1), includes a total of 2169 amino acid residues forming six independently refined molecules with the first four or five residues at the N-

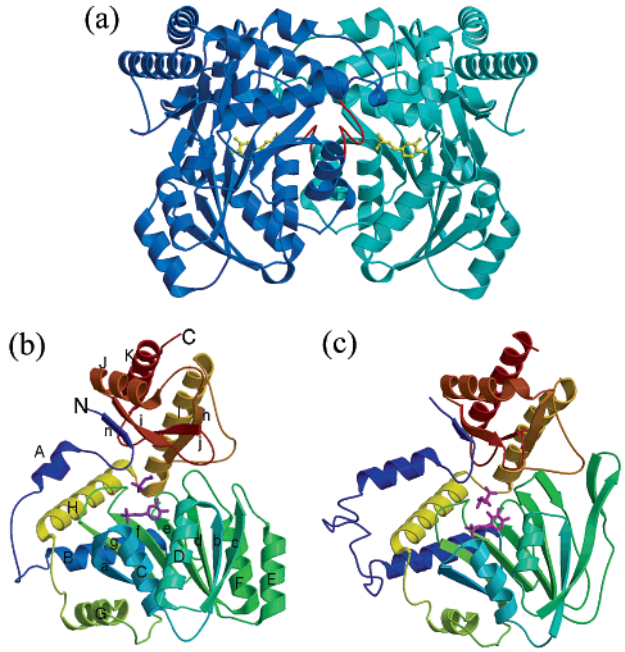

FIGURE 2: Structure of the AEPT molecule and a comparison with PSAT. (a) Dimer of AEPT. The PLP molecules are colored yellow, and the loops of the neighboring molecules which contained Thr243 residues that contribute to the stabilization of the PLPs are colored red. The noncrystallographic 2-fold axis is vertical and in the plane of the paper. (b) Ribbon diagram of the AEPT molecule, rainbow colored from the N- to the C-terminus. PLP and P-Ald are colored purple. (c) Ribbon diagram of the PSAT molecule.

terminus and the last residues at the C-terminus disordered. In addition, there are 140 disordered atoms from side chain residues, and 14 side chains with alternate conformations (see refinement) for the six molecules in the asymmetric unit. The model also includes six PLP molecules (the crystallization solution contained a stoichiometric amount of PLP), one per protein molecule, two P-Ald molecules (the crystallization solution contained a 50-fold excess of P-Ald), four phosphate molecules that probably correspond to disordered P-Ald molecules, and a total of 3040 water molecules.

Although the six molecules in the asymmetric unit were refined independently, they are essentially identical given the data resolution. The rmsd between the α-carbon atom positions (*41*) varies between 0.1 and 0.4 Å. For the positions of 84 atoms belonging to 11 residues at the active site, the rmsd varies between 0.1 and 0.2 Å.

Overall Structure. The overall fold of the AEPT molecule (Figure 2b) confirms that the protein is a member of the α-family of PLP-dependent enzymes, as classified by Alexander et al. (*42*). The molecule contains two domains: a large domain comprising residues 17−262 (helices A−H and strands a−g) and a small domain comprising residues 5−16 and 263−366 (helices I−K and strands h−j and n).

The prominent feature of the large domain is a core made up of a seven-stranded *β*-sheet characteristic of the aspartate transaminase enzyme family (fold type I of PLP-dependent enzymes) (*22*, *23*, *43*). The *β*-strands are arranged in the order agfedbc with all but strand g running parallel. The *β*-sheet is flanked on both faces by six α-helices (B−G) of which helix C is known as the "anchoring helix", present in all PLP-dependent enzymes (*44*). Helix H connects the large domain to the small domain.

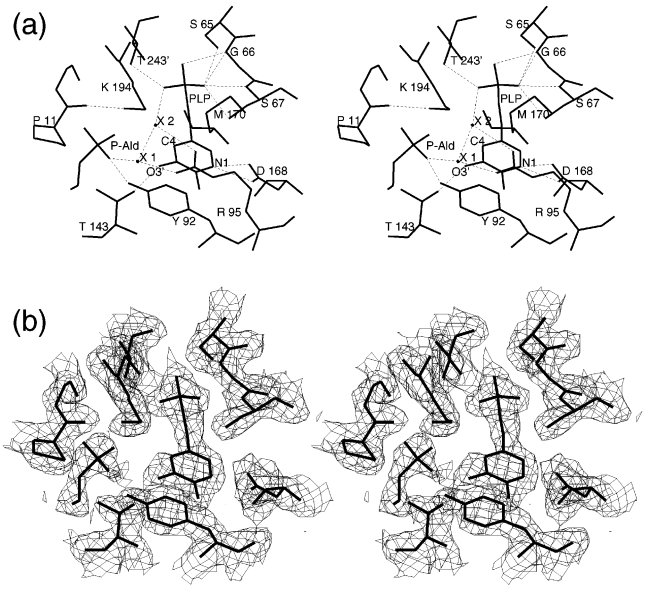

FIGURE 3: Active site of AEPT. (a) Stereoscopic presentation of the environment of PLP (—) and key interactions (- - -). The phosphate group of the cofactor is on the *re* face of the ring. In the active form, a Schiff base is formed between C4′ (missing in this structure) and Lys194 Nζ. The nitrogen atom, N1, and the hydroxyl group (labeled as O3′) interact with Asp168 and Thr143, respectively. Thr243 of a neighboring molecule (labeled T243′) interacts with the PLP phosphate group. The P-Ald is bound in a nonproductive mode for external aldamine formation. (b) Stereoscopic representation of the electron density map at the active site. There is no continuous density between Lys194 and PLP, indicating a lack of Schiff base. The density around the C4 atom is weak. The $2F_o - F_c$ coefficients and calculated phases are used. The map is contoured at the 1σ level.

Both the N- and C-termini are located within the small domain which contains a set of four highly twisted *β*-strands. Strands h, i, and n run parallel, while strand j runs antiparallel. Three α-helices (I−K) surround the strands. Helices H and I are connected by two consecutive glycine residues. Interdomain contacts are formed by the loops at cE, dF, ef, and fg in the large domain and helix I and loops nA, Ih, and ij in the small domain.

Dimers. The functional unit of the enzyme is a dimer (Figure 2a). There are three dimers per asymmetric unit. The two molecules of a dimer are related to each other by a 2-fold rotation axis located between the two large domains so that the edges of the large domains interact with each other through hydrogen bonds as well as hydrophobic interactions.

Whereas the molecules within a dimer are closely packed, the interdimer packing is loose and mainly involves the small domains.

Cofactor Binding Site. The active site of each monomer is located at the domain interface. In addition, loop residue Thr243 of a neighboring molecule interacts with the cofactor PLP (Figures 2a and 3a). An active site formed by residues from both subunits is a common feature for members of the aspartate aminotransferase family (type I of the PLP-dependent enzymes), consistent with the requirement of a dimer for biological function.

The short N-terminal *β*-strand makes a sharp turn at Pro11 and Gly12 (followed by a cis peptide bond between Gly12 and Pro13), exposing the peptidyl carbonyl of Pro11 to the active site so that it interacts with the side chain of Lys194 (Figure 3a). Three functionally important residues in this

region are located at or close to turns linking strands of the main β-sheet. They are Lys194 (between β-strands f and g), the lysine that forms the aldimine with PLP; Asp168 (on β-strand e) which interacts with the nitrogen atom of the PLP's pyridine ring; and Thr143 (between β-strand d and helix F) which forms a hydrogen bond between its Oγ atom and the O3′ atom of PLP. These three residues define the orientation of the pyridine ring of the PLP together with Tyr92 (on helix D), which forms π-electron interactions with the pyridine ring, and with Met170 (between β-strands e and f), which flanks the pyridine ring on the other side. The phosphate group of PLP and the active site Lys194 are oriented toward the *re* face of the pyridine ring, whereas the *si* face is exposed to the solvent channel. The main chain nitrogen atoms of Gly66 and Ser67 (on helix C) form an oxyanion hole for the phosphate group of the PLP, while the guanidinium group of Arg95 is bridged to the phosphate group through water molecules (shown as X1 and X2 in Figure 3a). From the second molecule of the dimer, the hydroxyl group of Thr243′ interacts with the phosphate group of PLP. Site-directed mutagenesis studies have shown that Asp168 and Lys194 are essential for AEPT catalysis (*20*).

Arg340 is equivalent to Arg386 of *E. coli* aspartate aminotransferase (AAT). It binds the α-carboxylate group of the substrate used in the second half of the reaction (*45*, *46*). Although the α-carboxylate substrate was absent in the current study, Arg340 is positioned appropriately for such interaction. Site-directed mutagenesis studies have shown that Arg340 plays an important role in L-Ala binding and catalysis (*20*).

DISCUSSION

An Elusive Schiff Base. The well-established chemistry of the aminotransferases involves the formation of a Schiff base (internal aldimine) between the active site lysine and the aldehyde group at the fourth position of the pyridine ring (*21*, *22*). When a substrate with an amino group approaches the complex, the amine group on the lysine side chain is exchanged for the amino group on the substrate, forming a second Schiff base complex (the external aldimine) and a free lysine (Figure 1). Subsequent steps of deprotonation, protonation, and hydration of the cofactor complex yield a free α-keto acid and a pyridoxamine 5′-phosphate (PMP). These steps comprise the first half of the transamination mechanism. The second half is essentially the reverse of the first, creating an amino acid from an α-keto acid, and returning to the internal aldimine.

For AEPT, the active site lysine is Lys194, and the expected product of the first half of the reaction is P-Ald. However, although Lys194 is located at the active site (Figure 3b), and projects toward the cofactor, no electron density links the Nζ atom of Lys194 to C4′ as expected for a Schiff base (atom labeling as in Figure 1). Moreover, in some of the molecules in the asymmetric unit, the C4 atom in the pyridine ring is associated with poor electron density, in contrast to the overall high quality of electron density associated with rest of the active site moieties. Instead of forming the Schiff base, the Nζ atom of Lys194 forms a hydrogen bond with the carbonyl oxygen of Pro11.

This was a surprise because the chemistry of the PLP-dependent enzymes is well-established and a Schiff base of either the internal or the external aldimine is present in all of the previously published crystal structures of aminotransferases except for the PMP-bound histidinol phosphate transaminase, a colorless protein (*25*). Both the protein solution and the crystals of AEPT were colored, orange and yellow, respectively, indicative of Schiff base formation.

Several attempts were made during refinement to see whether Schiff base formation was possible. To link the Nζ atom of Lys194 to the C4′ atom of PLP, it was necessary to tilt the pyridine ring and to move main chain atoms of residues near Lys194. However, models including forced Schiff bases were not refined successfully. Modeling the cofactor as an aldehyde (free PLP in solution) was equally unsuccessful. In the final model, the PLP molecules lack the C4′ atom, and extra residual electron density peaks observed on the edge of the pyridine rings opposite Lys194 residues were modeled as water molecules with alternating sites.

We have no substantiated explanation for this anomaly. It is tempting to speculate that the electron-rich Schiff base is susceptible to synchrotron radiation damage even at cryo temperatures. This is consistent with recent publications reporting the breaking of disulfide bonds and carboxylate groups upon prolonged exposure to synchrotron radiation (*47*, *48*). Unfortunately, the crystals diffract X-rays to a resolution lower than 3 Å on a conventional generator, and although data were collected on the home facility, the low resolution prevented the reliable assessment of the presence of a Schiff base. We have also noted that the crystals remain yellow upon flash-cooling; therefore, the breaking of the Schiff base is not the consequence of flash-cooling.

Binding of Product. The crystallization medium contained a 50-fold excess of the expected product, P-Ald. Accordingly, the solvent channels at the active sites of all six molecules contain electron density that may correspond to the product. The density is located between the two arginine residues, Arg95 and Arg340. In contrast to the well-defined electron density of the cofactor PLP, the electron density at the postulated sites of the P-Ald molecules is weak. Only two of the six molecules in the asymmetric unit could be modeled as P-Ald, and for the remaining four, only a phosphate group could be fitted. Temperature factors in a crystal structure indicate the mobility of the atoms. In the structure reported here, the crystallographic temperature factors of the PLP groups refined to an average value of 23 Å^2, while the average temperature factor of P-Ald and phosphate ions is 69 Å^2. The high mobility of the product is expected to enable the binding of the next substrate of the second half of the enzymatic reaction. The electrostatic surface shows that the active site crevice is largely positively charged (Figure 4a). The phosphate group of PLP binds at the bottom of the crevice and is obstructed from bulk solvent by Tyr92. In contrast, the phosphate ion that may represent part of P-Ald binds above PLP in a less restrictive environment (Figure 4b).

Comparison with Phosphoserine Aminotransferase. The closest homologue with a known three-dimensional structure is phosphoserine aminotransferase (PSAT) (*19*), classified as type I by Schneider, Käck, and Lindqvist (*24*). Despite the low level of sequence identity between the two proteins (17.5%), the rmsd between the α-carbon atom positions of the two proteins (monomers) is 2.0 Å for 321 superposed atom pairs. The most notable difference is that helices E and F of AEPT are missing in PSAT (Figure 2c).

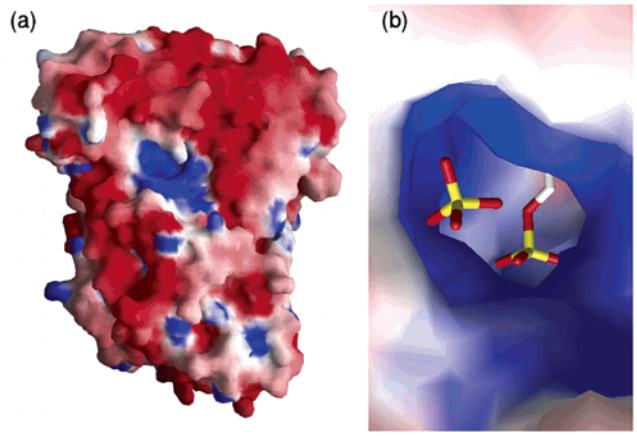

FIGURE 4: Surface potential of AEPT. (a) Dimer. The surface is colored according to electrostatic potential using blue for positively charged regions and red for negatively charged regions. The positively charged cavity corresponds to the active site of one molecule. A second active site cavity is located on the backside of the dimer and is not seen in the figure. (b) Close-up view of the active site showing a phosphate group (part of P-Ald) in the solvent channel and part of a more buried PLP. The remaining P-Ald molecule does not adopt a unique position in all six molecules in the asymmetric unit. Both compounds are negatively charged, and thus complement the positively charged active site environment.

Of the seven common structural elements derived from a comparison of representative crystal structures of five unrelated PLP-dependent enzymes (*44*), strands a and d−g and helix C are present in both AEPT and PSAT. The seventh, corresponding to a helix at positions 283−291 in AAT, however, is absent from both AEPT and PSAT.

The PLP in the PSAT structure forms a Schiff base with the inhibitor methyl glutamate (GAM). Therefore, the catalytic lysine residue of PSAT (Lys198) is free, with its Nζ atom pointing away from the C4′ atom, although remaining closer to C4′ compared with the Nζ atom of Lys194 of AEPT.

Other invariant residues in the two active sites include (residue number of PSAT in parentheses) Thr143 (153), Asp168 (174), Arg340 (335), and Thr243′ (240′), belonging to a neighboring molecule. Instead of Tyr92 that forms a $\pi-\pi$ interaction with PLP, PSAT has another aromatic residue, Trp102. A superposition of the two active sites (using the above residues) is shown in Figure 5. While the aromatic rings of Trp102 in PSAT and Tyr92 in AEPT are coplanar, the PLP pyridine ring orientations are different by 17°. The superposition shows why phosphoserine would not be a substrate of AEPT, as the site where the phosphoserine side chain should bind is blocked in AEPT by the hydroxyl group of Tyr92. Although P-Ald is located in the general vicinity of GAM (closer to the γ-carboxylate than the α-carboxylate), the pocket formed by Tyr92, Tyr329, Pro11, and Thr143 in AEPT is shaped specifically to accommodate AEP and not phosphoserine.

Another interesting feature emerging from the superposition is that the guanidinium groups of Arg95 in AEPT and Arg77 in PSAT occupy approximately the same position, yet they are located on different structural units. In both enzymes, the arginine residues may play a role in providing a counter charge to the phosphonyl or phosphoryl group of the cofactor and/or substrate.

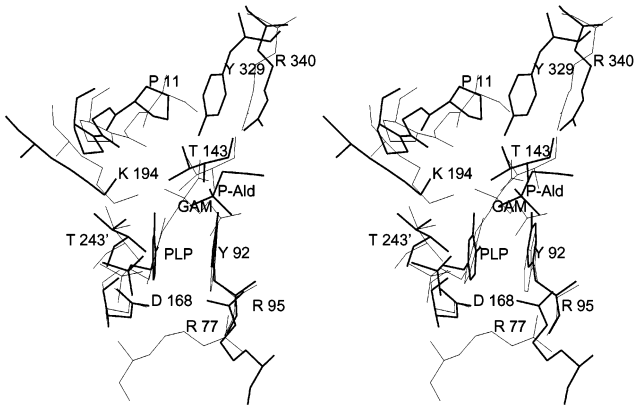

FIGURE 5: Comparison of the active site of AEPT and PSAT. The PSAT active site residues (thin lines) were superimposed on those of AEPT (thick lines) using 33 atom pairs of four invariant residues: Thr143 (153), Asp168 (174), Thr243′ (240′), and Arg340 (335). The PSAT active site includes GAM, and that of AEPT includes P-Ald.

Mechanistic Implications. The established enzymatic reaction for all aminotransferases follows a two-step, bi-bi, ping-pong mechanism (*49*). The structural implication is that the substrate and the enzyme's Lys194 approach the aromatic ring of the cofactor PLP from opposite sides. In the crystal structure, even though the electron density for the Schiff base is missing, and the electron density for the product is poor, the positions of Lys194 and the product group are consistent with such a mechanism. The PLP binds with the *re* face of the ring toward Lys194 and the *si* face toward a solvent channel, blocked from bulk solvent by Tyr92. The environment of the PLP is very crowded, restricting the size and shape of possible substrates.

One possible model of the bound substrate AEP is shown in Figure 6 as a quinoid, the planar external aldimine that comprises the first obligatory intermediate along the reaction pathway (Figure 6a). In this model (Figure 6b), the phosphonate group binds between Arg95 and Arg340, with the phosphonate oxygen atoms within hydrogen bonding distance of the hydroxyl groups of Tyr92 and Tyr329. The side chain of Lys194 is modeled in an extended conformation so that it is within striking distance of (a) the C2 atom of the AEP moiety from which a proton is to be abstracted and (b) the C4′ atom attached to the PLP pyridine ring to which a proton is to be donated. Following stereoelectronic rules, the C−H bond to be broken at tetrahedral C2 of AEP is the one orthogonal to the plane of the PLP pyridine ring (*50*). Accordingly, the C2−H1 bond will be broken. This stereospecificity has been proven by Lacoste et al. (*51*) in their ^{1}H NMR studies on the AEPT of *P. aeruginosa*.

The final product of the AEPT reaction, or the substrate of the reverse reaction, L-alanine, could not be successfully modeled in the active site as seen in the crystal structure reported here. When the model of the external aldimine is built by analogy to the model of AEP, the hydrogen atom (H) attached to the α-carbon (CA) of L-alanine occupies the same site as H1 of the AEP model (Figure 6c). The L-alanine methyl group (CB) aligns with the phosphonate group of AEP and the carboxylate group with H2 of the AEP. While there is ample space to accommodate the methyl group at the phosphonate site, the carboxylate group is too large to replace H2, due to the spatial restriction imposed by Pro11

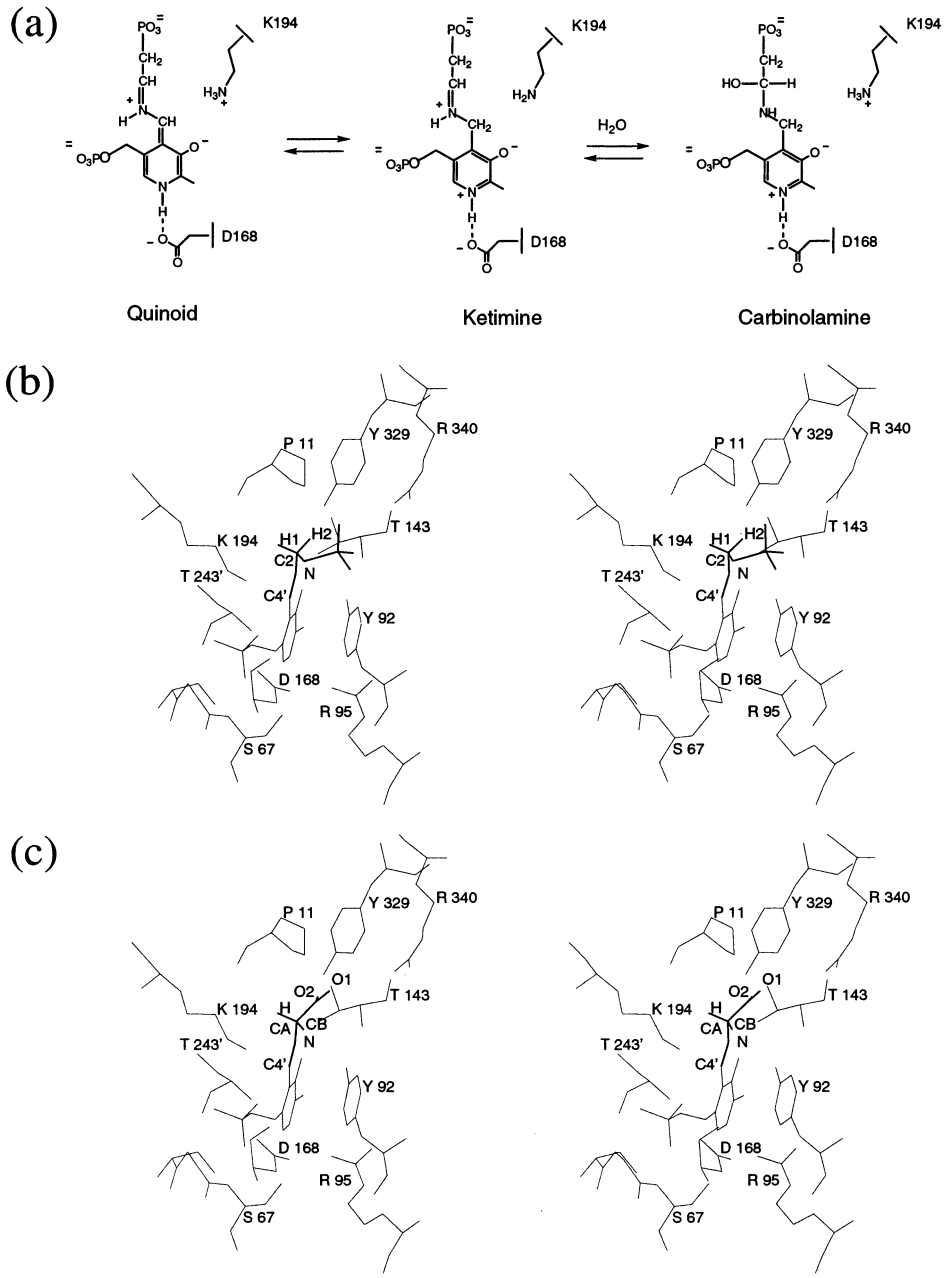

FIGURE 6: Models of external aldimines. (a) Scheme of the quinoid, ketimine, and carbinolamine tautomeric forms. (b) Modeled AEP−PLP Schiff base. H1 is the hydrogen atom to be abstracted by Lys194. H2 is surrounded by Pro11, Tyr329, and Thr143. (c) Modeled L-Ala−PLP Schiff base. Note the steric clash of the carboxylate group of Ala with residues Pro11 and Thr143, indicating that the enzyme must undergo conformational transition.

and Thr143 in the present conformation. Even when the requirement for planar quinoid is relaxed, and ketimine or carbinolamine conformations (Figure 6a) such as seen in inhibitor complexes of PSAT and 2,2-diakylglycine decarboxylase (*52*) are explored, the steric clash cannot be relieved.

Accommodation of the α-carboxylate group of L-Ala requires conformational adjustment of the surrounding residues (Tyr92, Thr143, Tyr329, and Pro11). This conformational change may be achieved by a global shift of the small domain, which would have an impact on Tyr329 and Pro11. Such shifts of the small domain are known to accompany ligand binding in other aminotransferases (*43*, *53*, *54*). It is tempting to speculate that conformational changes may be facilitated by a hinge movement centered at the Gly263-Gly264 peptide connecting helices H and I.

Kinetic experiments for assessing the substrate specificity of AEPT (*20*) showed that while the activity toward α-ketoglutaric acid (yielding L-glutamate) was low but measurable, the enzyme was inactive toward the shorter chain carboxylates: oxaloacetate (yielding L-aspartate) and phosphonopyruvate (yielding phosphonoalanine) (*26*). The only residue which could possibly be responsible for this discrimination is Tyr329, a residue unique to AEPT. A crystal structure of an enzyme−inhibitor complex will reveal the nature of domain movement and the role of Tyr329 in determining substrate specificity.

ACKNOWLEDGMENT

We thank the staff of beamline X12C at the National Synchrotron Light Source and the staff of IMCA-CAT at the Advanced Photon Source for their help during data

collection. Use of the National Synchrotron Light Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Prime Contract DE-AC02-98CH110886, by the National Science Foundation, and by National Institutes of Health Grant 1P41 RR12408-01A1. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under Contract W-31-109-Eng-38. The IMCA-CAT facility is supported by the companies of the Industrial Macromolecular Crystallographic Association, through a contract with the Illinois Institute of Technology.

REFERENCES

1. Schowanek, D., and Verstraete, W. (1990) *Appl. Environ. Microbiol. 56*, 895−903.
2. Wassef, M. K., and Hendrix, J. W. (1976) *Biochim. Biophys. Acta 486*, 172−178.
3. Araki, S., Komai, Y., and Satake, M. (1980) *J. Biochem. 87*, 503−510.
4. Hasegawa, S., Tamari, M., and Kametaka, M. (1976) *J. Biochem. 80*, 531−535.
5. Tan, S. A., and Tan, L. G., (1989) *Clin. Physiol. Biochem. 7*, 303−309.
6. Aminophosphonate metabolism: Reference Pathway from GenomeNet World Wide Web server in Japan (www.genome.ad.jp). Under KEGG Pathway/Metabolic pathway/Metabolism of Other Amino Acids.
7. Baumann, H., Tzianabos, A. O., Brisson, J. R., Kasper, D. L., and Jennings H. J. (1992) Structural elucidation of two capsular polysaccharides from one strain of *Bacteroides fragilis* using high-resolution NMR spectroscopy. *Biochemistry 31*, 4081−4089.
8. Kariotoglou, D. M., and Mastronicolis, S. K. (2001) *Aurelia aurita*, *Lipids 36*, 1255−1264.
9. Hard, K., Van Doorn, J. M., Thomas-Oates, J. E., Kamerling, J. P., and Van der Horst, D. J. (1993) *Biochemistry 32*, 766−775.
10. Rosenthal, A. F., and Pousada, M. (1968) *Biochim. Biophys. Acta 164*, 226−237.
11. Duroma, C., Lacoste, A. M., and Cassaigne, A. (1989) *Biochim. Biophys. Acta 997*, 193−198.
12. La Nauze, J. M., and Rosenberg, H. (1968) *Biochim. Biophys. Acta 165*, 438−447.
13. Jiang, W., Metcalf, W. W., Lee, K. S., and Wanner, B. L. (1995) *J. Bacteriol. 177*, 6411−6421.
14. Ternan, N. G., and Quinn, J. P. (1998) *Syst. Appl. Microbiol. 21*, 346−352.
15. Wanner, B. L. (1994) *Biodegradation 5*, 175−184.
16. Yasuta, T., Okazaki, S., Mitsui, H., Yuhashi, K. I., Ezura, H., and Minamisawa, K. (2001) *Appl. Environ. Microbiol. 67*, 4999−5009.
17. Mitchell, R. E., Frey, E. J., and Benn, M. K. (1986) *Phytochemistry 25* 2711−2715.
18. Owen, L. D., Thompson, J. F., Pitcher, R. G., and Williams, T. (1972) *J. Chem. Soc., Chem. Commun. 197*, 714.
19. Hester, G., Stark, W., Moser, M., Kallen, J., Marković-Housley, J. K. Z., and Jonsonius, J. N. (1999) *J. Mol. Biol. 286*, 829−850.
20. Kim, A. D., Baker, A. S., Dunaway-Mariano, D., Metcalf, W. W., Wanner, B. L., and Martin, B. M. (2002) *J. Bacteriol. 184*, 4134−4140.
21. von Stosch, A. G. (1996) *Biochemistry 35*, 15260−15268.
22. Malashkevich, V. N., Toney, M. D., and Jansonius, J. N. (1993) *Biochemistry 32*, 13451−13462.
23. Jansonius, J. N. (1998) *Curr. Opin. Struct. Biol. 8*, 759−769.
24. Schneider, G., Käck, H., and Lindqvist, Y. (2000) *Structure 8*, R1−R6.
25. Sivaraman, J., Li, Y., Larocque, R., Schrag, J. D., Cygler, M., and Matte, A. (2001) *J. Mol. Biol. 311*, 761−776.
26. Kim, A. (1999) Ph.D. Dissertation, pp 180−184, University of Maryland, College Park, MD.
27. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
28. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
29. Uson, I., and Sheldrick, G. M. (1999) *Curr. Opin. Struct. Biol. 9*, 643−648.
30. Sheldrick, G. M. (1998) SHELX applications to macromolecules, in *Direct methods for solving macromolecular structures* (Fortier, S., Ed.) pp 401−411, Kluwer Academic Publishers, Dordrecht, The Netherlands.
31. Xu, H., and Weeks, C. M. (2002) *Acta Crystallogr. D58*, 90−96.
32. Weeks, C. M., and Miller, R. (1999) *J. Appl. Crystallogr. 32*, 120−124.
33. Ramakrishnan, V., and Biou, V. (1997) *Methods Enzymol. 276*, 538−557.
34. Cowtan K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography 31*, 34−38.
35. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr. D49*, 129−149.
36. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol. 6*, 458−463.
37. Roussel, A., and Cambillau, C. (1991) Turbo-Frodo, in *Silicon Graphics Geometry Patners Directory*, p 86, Silicon Graphics, Mountain View, CA.
38. Brünger, A. T., et al. (1998) *Acta Crystallogr. D54*, 905−921.
39. Pannu, N. S., and Read, R. J. (1996) *Acta Crystallogr. A52*, 659−668.
40. Kleywegt, G. J., and Jones, T. A. (1998) *Acta Crystallogr. D54*, 1119−1131 (CCP4 Proceedings).
41. Cohen, G. H. (1997) *J. Appl. Crystallogr. 30*, 1160−1161.
42. Alexander, F. W., Sandmeier, E., Mehta, P. K., and Christen, P. (1994) *Eur. J. Biochem. 219*, 953−960.
43. Jäger, J., Moser, M., Sauder, U., and Jansonius, J. N. (1994) *J. Mol. Biol. 239*, 285−305.
44. Denessiouk, K. A., Deneysyuk, A. I., Lehtonen, J. V., Korpela, T., and Johnson, M. S. (1999) *Proteins 35*, 250−261.
45. Mehta, P. K., Hale, T. I., and Christen, P. (1993) *Eur. J. Biochem. 214*, 549−561.
46. Danishefsky, A. T., Onnufer, J. J., Petsko, G. A., and Ringe, D. (1991) *Biochemistry 30*, 1980−1985.
47. Weik, M., et al. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 623−628.
48. Burmeister, W. P. (2000) *Acta Crystallogr. D56*, 328−341.
49. Cleland, W. W. (1963) *Biochim. Biophys. Acta 67*, 104−137.
50. Dunathan, H. C. (1966) *Proc. Natl. Acad. Sci. U.S.A. 55*, 712−716.
51. Lacoste, A. M., Dumora, C., Balas, L., Hammerschmidt, F., and Vercauteren, J. (1993) *Eur. J. Biochem. 215*, 841−844.
52. Malashkevich, V. N., Strop, P., Keller, J. W., Jansonius, J. N., and Toney, M. D. (1999) *J. Mol. Biol. 294*, 193−200.
53. Hohenester, E., and Jansonius, J. N. (1994) *J. Mol. Biol. 236*, 963−968.
54. Okamoto, A., Nakai, Y., Hayashi, H., Hirotsu, K., and Kagamiyama, H. (1998) *J. Mol. Biol. 280*, 443−461.

BI026231V