

Pa0148 from *Pseudomonas aeruginosa* Catalyzes the Deamination of Adenine

Alissa M. Goble,[†] Zhening Zhang,[‡] J. Michael Sauder,[§] Stephen K. Burley,[§] Subramanyam Swaminathan,^{*,‡} and Frank M. Raushel^{*,†}

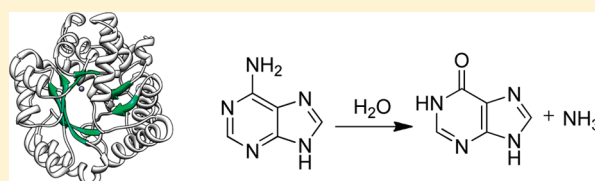
[†]Department of Chemistry, P.O. Box 30012, Texas A&M University, College Station, Texas 77843-3012, United States

[‡]Biology Department, Brookhaven National Laboratory, P.O. Box 5000, Upton, New York 11973-5000, United States

[§]Lilly Biotechnology Center, 10300 Campus Point Drive, San Diego, California 92121, United States

S Supporting Information

ABSTRACT: Four proteins from NCBI cog1816, previously annotated as adenosine deaminases, have been subjected to structural and functional characterization. Pa0148 (*Pseudomonas aeruginosa* PAO1), AAur1117 (*Arthrobacter aureescens* TC1), Sgx9403e, and Sgx9403g have been purified and their substrate profiles determined. Adenosine is not a substrate for any of these enzymes. All of these proteins will deaminate adenine to produce hypoxanthine with k_{cat}/K_m values that exceed $10^5 \text{ M}^{-1} \text{ s}^{-1}$. These enzymes will also accept 6-chloropurine, 6-methoxypurine, *N*-6-methyladenine, and 2,6-diaminopurine as alternate substrates. X-ray structures of Pa0148 and AAur1117 have been determined and reveal nearly identical distorted (β/α)₈ barrels with a single zinc ion that is characteristic of members of the amidohydrolyase superfamily. Structures of Pa0148 with adenine, 6-chloropurine, and hypoxanthine were also determined, thereby permitting identification of the residues responsible for coordinating the substrate and product.



According to NCBI, the genomes of more than 1500 bacteria have been completely sequenced. Because of the large and rapidly increasing volume of DNA sequence data, functional annotations based primarily on protein similarity scores to enzymes of known function are now routinely employed.¹ Unfortunately, the reliability of these functional annotations is directly dependent on the stringency of a given function for relatively small differences in amino acid sequence. Misannotations can occur from the utilization of similarity scores below reliable cutoff thresholds and from perturbations in catalytic function with relatively few changes in amino acid sequence.² These problems are further amplified when the functional boundaries between orthologs and paralogs are difficult to discern on the basis of sequence data alone.³ Misannotations are subsequently propagated to other databases, and the number of proteins with incorrectly identified functions has increased.² The considerable number of proteins of unknown or uncertain function suggests that a significant segment of the metabolic landscape remains undiscovered.

Adenosine deaminase is a member of the amidohydrolyase superfamily (AHS) of enzymes. This superfamily was first identified by Holm and Sander on the basis of the similarities in the three-dimensional structures of phosphotriesterase, adenosine deaminase, and urease.³ Proteins in the AHS have a distorted (β/α)₈ barrel fold with conserved metal binding residues at the C-terminal ends of β -strands 1, 4–6, and 8. These enzymes have either a mononuclear or binuclear metal center that activates a hydrolytic water for nucleophilic attack on amino acids, sugars,

nucleic acids, and organophosphate esters.⁴ Within the AHS, there are 24 clusters of orthologous groups (COG) as defined by NCBI. Three of these clusters of enzymes are known to catalyze deamination reactions: cog1001, cog0402, and cog1816. The prototypical adenine deaminase (ADE) and *N*-6-methyladenine deaminase (6-MAD) are members of cog1001.⁵ Enzymes that deaminate guanine, cytosine, *S*-adenosylhomocysteine, thio-methyladenosine, *N*-formimino-*L*-glutamate, and 8-oxoguanine are found in cog0402.⁶ According to NCBI, all bacterial proteins in cog1816 catalyze the deamination of adenosine.

The structure and mechanism of action of adenosine deaminases from several organisms have been studied.^{7–19} A sequence similarity network for cog1816 is illustrated in Figure 1.²⁰ Enzymes experimentally verified as adenosine deaminase (ADD) are found in group 5, which contains the adenosine deaminase from *Escherichia coli* K-12 (locus tag b1623). Essential residues for adenosine deaminase activity include an HxHxD motif following β -strand 1, an aspartate following β -strand 3, a glycine following β -strand 4, an HxxE motif following β -strand 5, a histidine following β -strand 6, and a pair of aspartates following β -strand 8.^{8,17,19,21,22} These residues occur in every protein of group 5. Despite an overall amino acid sequence identity of $\sim 30\%$ to the *E. coli* adenosine deaminase, proteins of group 3 are missing several residues that have been implicated in substrate binding.

Received: June 4, 2011

Revised: June 25, 2011

Published: June 28, 2011

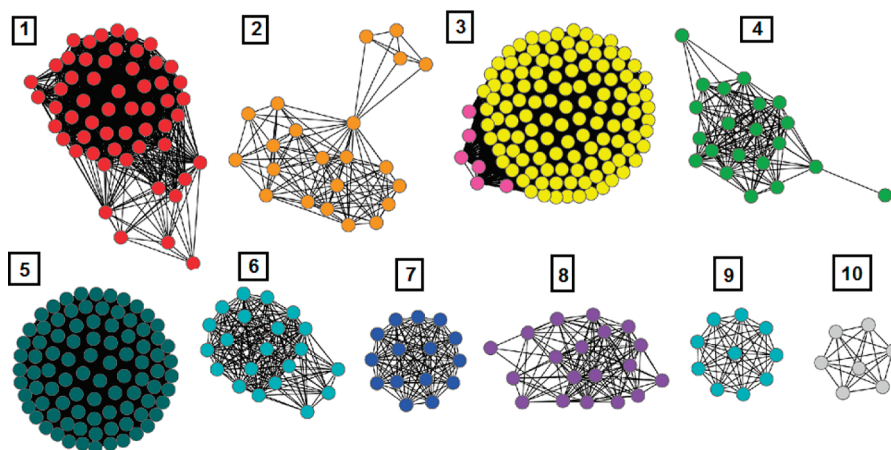


Figure 1. Sequence similarity network created using Cytoscape (<http://www.cytoscape.org>) of cog1816 from the amidohydrolase superfamily. Each node in the network represents a single sequence, and each edge (depicted as lines) represents the pairwise connection between two sequences at a BLAST E value of better than 1×10^{-70} . Lengths of edges are not significant, except for tightly clustered groups, which are more closely related than sequences with only a few connections. Group 5 contains the adenosine deaminase from *E. coli*. In group 3, the sequences colored yellow contain all of the essential residues for the deamination of adenine identified in this investigation.

These residues include aspartate residues following β -strands 1 and 3 and the glycine residue following β -strand 4. On the basis of sequence alignments between group 3 and 5 enzymes, we anticipated that some members of cog1816 do not catalyze the deamination of adenosine.

The initial target for this investigation was the functional annotation of Pa0148 from *Pseudomonas aeruginosa* PAO1. This enzyme and three other enzymes are in group 3 of cog1816. The substrate profiles for all four enzymes have been determined, and the three-dimensional structures of Pa0148 and Aaur1117 from *Arthrobacter aurescens* TC1 have been determined with substrate and product ligands bound in the active site. These enzymes do not catalyze the deamination of adenosine, but they do catalyze the deamination of adenine to hypoxanthine.

MATERIALS AND METHODS

Materials. All chemicals were purchased from Sigma-Aldrich, unless otherwise stated. *N*-6-Methyladenine was purchased from Spectrum, and 6-methoxypurine was obtained from Tokyo Chemical Industry Co. Zeatin was acquired from MP Biomedicals, and isoguanine was bought from Santa Cruz Biotechnology. 2'-Deoxy-*N*-6-methyladenosine was procured from Carbosynth.

Cloning and Purification of Pa0148, Sgx9403e, and Sgx9403g. The gene for Pa0148 (gi|15595346; Sgx9608a) was obtained from the genomic DNA of *P. aeruginosa* PAO1 (ATCC 47085D), and the genes for Sgx9403e and Sgx9403g were chemically synthesized by back-translation and codon optimized for *E. coli* expression (Codon Devices, Inc.). The latter two genes were first identified from the Sargasso Sea environmental sequencing effort through the Global Ocean Sampling expedition, and the original GI numbers (44392365 and 44590840 for Sgx9403e and Sgx9403g, respectively) are now obsolete. The sequence currently in GenBank that is most similar to that of Sgx9403e is gi|239993551, which belongs to AmacA2_3576 from *Alteromonas macleodii* ATCC 27126. With the exception of the first 16 residues of Sgx9403e that are not present in AmacA2_3576, the two proteins are 99% identical. The sequence in GenBank most similar to that of Sgx9403g is gi|78065603, which belongs to Bcep18194_A4131 from

Burkholderia sp. 383. Bcep18194_A4131 and Sgx9403g have sequences that are 99% identical. The synthesized genes were cloned into a custom TOPO-isomerase vector, pSGX3(BC), supplied by Invitrogen. Forward primers for Pa0148, Sgx9403e, and Sgx9403g were TACGAATGGCTCAACGCCTTG, AGCCTGAGCAGCATTATCAAAAATAT, and ACCCCGACCTTTAAAGACAAG, respectively, and reverse primers were CCTTCACCAGGCGGCATCCA, CTGCTGAGAGTACTTTTCAACC, and CTGCAGCTTGCCAATACGCATC, respectively. The clones encode Met-Ser-Leu followed by the polymerase chain reaction (PCR) product and end with Glu-Gly-His₆. Miniprep DNA was transformed into BL21(DE3)-Codon+RIL expression cells (Stratagene), expressed, and made into a 30% glycerol stock for large-scale fermentation.

The expression clones were cultured using high-yield selenomethionine (SeMet) medium (Orion Enterprises, Inc., Northbrook, IL). Overnight cultures (50 mL) in 250 mL baffled flasks were cultivated at 37 °C from a frozen glycerol stock for 16 h, then transferred to 2 L baffled shake flasks containing 1 L of HY-SeMet medium (100 μ g/mL kanamycin and 30 μ g/mL chloramphenicol), and grown to an OD₆₀₀ of \sim 1.0. SeMet was then added for labeling at 90 mg/L, followed by IPTG added to a final concentration of 0.4 mM. Cells were further grown at 22 °C for 18 h, then harvested using standard centrifugation for 10 min at 6000 rpm, and frozen at -80 °C.

Cells were lysed in 20 mM Tris (pH 8.0), 0.5 M NaCl, 25 mM imidazole, and 0.1% Tween 20 by sonication. The cellular debris was removed by centrifugation for 30 min (39800g). The supernatant was collected and incubated with 10 mL of a 50% slurry of Ni-NTA agarose (Qiagen) for 30 min with gentle stirring. The sample was then poured into a drip column and washed with 50 mL of wash buffer [20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10% glycerol, and 25 mM imidazole] to remove unbound proteins. The protein of interest was eluted using 25 mL of elution buffer (wash buffer with 500 mM imidazole). Fractions containing the protein were pooled and further purified by gel filtration chromatography on a GE Healthcare HiLoad 16/60 Superdex 200 prep grade column pre-equilibrated with gel filtration buffer [10 mM HEPES (pH 7.5), 150 mM NaCl, 10% glycerol, and 5 mM DTT]. Fractions containing the

Table 1. Data Collection and Refinement Statistics

	3OU8	3PAO	3PBM	3PAN	3RYS
Data Collection Statistics					
wavelength (Å)	0.9795	0.9795	0.9795	0.9795	0.9795
resolution (Å)	50–2.51	50–2.40	50–2.60	50–2.63	50–2.60
outer shell resolution (Å)	2.65–2.51	2.49–2.40	2.69–2.60	2.72–2.63	2.69–2.60
space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P4 ₁ 2 ₁ 2
cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	44.4, 74.0, 177.9	44.6, 74.0, 177.0	44.7, 74.1, 178.2	44.7, 74.2, 178.0	124.1, 124.1, 89.7
α , β , γ (deg)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
no. of molecules per asymmetric unit	2	2	2	2	2
redundancy (overall/outermost shell)	13.9 (14.0)	7.9 (7.0)	6.6 (6.3)	14.3 (14.2)	18.3 (18.7)
<i>I</i> / σ (<i>I</i>) (overall/outermost shell)	16.2 (9.8)	18.2 (5.1)	9.28 (3.32)	19.83 (8.7)	19.83 (8.7)
<i>R</i> _{merge} (overall/outermost shell)	0.126 (0.223)	0.179 (0.44)	0.176 (0.473)	0.148 (0.333)	0.149 (0.528)
completeness (%) (overall/outermost shell)	100.0 (100.0)	100.0 (100.0)	99.4 (96.1)	99.9 (99.4)	100.0 (100.0)
no. of reflections	20922	24150	19042	198396	22144
Refinement Statistics					
resolution range (Å)	50–2.51	50–2.40	50–2.60	50–2.63	50–2.60
no. of reflections	20862	21273	18038	17938	21346
completeness (work + test) (%)	99.95	99.49	94.13	97.57	96.56
<i>R</i> _{factor} (%)	19.04	18.12	17.26	19.24	20
<i>R</i> _{free} (%)	25.43	24.03	23.2	25.28	25.61
no. of protein atoms	5069	5064	5064	5064	5100
no. of water molecules	177	176	133	148	85
no. of ligand atoms	2	22	22	22	22
overall mean <i>B</i> value (Å ²)	18.44	23.13	20.75	18.57	21.11
rmsd for bonds (Å)	0.006	0.007	0.01	0.01	0.004
rmsd for angles (deg)	0.871	0.991	1.27	0.701	0.767
Ramachandran plot analysis (%)					
most favored region (additionally allowed)	95.02 (3.53)	95.83 (2.89)	94.54 (4.17)	95.83 (2.73)	97.25 (1.83)
disallowed region	1.44	1.28	1.28	1.44	0.92

target protein were combined and concentrated by centrifugation in an Amicon Ultra-15 10000 molecular weight cutoff centrifugal filter unit. The final yields and concentrations were 31 mg (6 mg/mL) of Pa0148, 92 mg (13 mg/mL) of Sgx9403e, and 114 mg (13.5 mg/mL) of Sgx9403g per liter of medium. Electrospray mass spectroscopy was used to obtain an accurate mass of the purified protein and establish the extent of selenomethionine labeling (Pa0148, 37430 Da and 4 SeMet; Sgx9403e, 39016 Da and 8 SeMet; and Sgx9403g, 39001 Da and 2 SeMet).

The expression plasmids are available through the PSI Material Repository as NYSGXRC clone IDs 9608a2BCt1p1, 9403e1BCt11p1, and 9403g1BCt10p1. Associated experimental information is available in the Protein Expression Purification Crystallization Database as TargetIDs “NYSGXRC-9608a”, “NYSGXRC-9403e”, and “NYSGXRC-9403g”.

Cloning and Purification of AAur1117 from *A. aureus* TC1. AAur1117 was cloned from the genomic DNA of *A. aureus* TC1. The PCR product was amplified using the primer pair 5'-AGAGGATCCGTGGAACTTTTGGCGAGAAAAC-TACC-3' and 5'-AGAGAAGCTTTTAGGCGGGCGTAACG-GAGGC-3'. The restriction sites for *Bam*HI and *Hind*III were inserted into the forward and reverse primers, respectively. The PCR product was purified with a PCR cleanup system (Promega), digested with *Bam*HI and *Hind*III, and ligated into a pET-30a(+) vector, which was previously digested with *Bam*HI

and *Hind*III. The cloned gene fragment was sequenced to verify the fidelity of the PCR amplification.

The gene for AAur1117 inserted into a pET-30a(+) vector was transformed into BL21(DE3) cells (Novagen). A single colony was used to inoculate a 5 mL overnight culture of LB medium containing 50 μ g/mL kanamycin. Each overnight culture was used to inoculate 1.0 L of LB medium containing 50 μ g/mL kanamycin. Cultures (1 L) were grown at 37 °C and induced with 50 μ M isopropyl D-thiogalactopyranoside (IPTG) when an OD₆₀₀ of 0.6 was reached. At the time of induction, the temperature was lowered to 20 °C and the sample shaken for 18 h before the cells were harvested at 8000 rpm for 10 min. The cells were resuspended in 50 mM HEPES buffer (pH 7.5) containing 0.1 mg/mL phenylmethanesulfonyl fluoride and 0.5 mg/mL deoxyribonuclease I from bovine pancreas. Cells were lysed by sonication. Soluble protein was separated from the cell debris by centrifugation and fractionated by ammonium sulfate precipitation. The precipitated protein (35–70% ammonium sulfate saturation) was resuspended in 50 mM HEPES buffer (pH 7.5) and loaded onto a High Load 26/60 Superdex 200 prep grade gel filtration column (GE Healthcare). Fractions containing AAur1117 were loaded onto a 6 mL ResourceQ column and eluted with a gradient of NaCl.

Crystallization and Determination of Structures of Pa0148 and AAur1117. Pa0148 and AAur1117 were each concentrated

to ~ 10 mg/mL and subjected to sitting-drop vapor diffusion by mixing 1 μ L of protein solution and 1 μ L of reservoir from Index and Crystal screen of Hampton Research. Adenine was added to the AAur1117 protein solution to a final concentration of 10 mM before crystallization trials were initiated. Pa0148 formed thin platelike crystals overnight from 0.2 M magnesium chloride, 0.1 M HEPES (pH 7.5), and 25% (w/v) polyethylene glycol (PEG) 3350. AAur1117 formed rod-shaped crystals over the course of 3 months from 0.2 M magnesium chloride hexahydrate, 0.1 M Tris hydrochloride (pH 8.5), and 30% (w/v) PEG 4000. All crystals were cryoprotected by addition of 20% glycerol (v/v) to the mother liquor and flash-frozen by direct immersion in liquid nitrogen. Pa0148 was also soaked individually in 10 mM adenine, hypoxanthine, or 6-chloropurine added to mother liquor with 20% glycerol and flash-frozen in liquid nitrogen for the collection of X-ray data.

For Pa0148 and AAur1117, X-ray diffraction data at the selenium absorption edge ($\lambda = 0.9795$ Å) were recorded at beamlines X25 and X12C, respectively, at the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. Diffraction data were processed with HKL2000 or MOSFLM/SCALA.^{23–25} Pa0148 crystallized in orthorhombic space group $P2_12_12_1$ with two molecules per asymmetric unit and diffracted to ~ 2.5 Å resolution. The structure of Pa0148 was determined by AUTOSOL,²⁶ followed by autobuilding in ARP/wARP.²⁷ Atomic models for both structures were subsequently manually adjusted using COOT.²⁸ Structures of Pa0148 bound to adenine, hypoxanthine, or 6-chloropurine were determined via molecular replacement using PHASER²⁹ with the structure of Pa0148 as a search model. AAur1117 crystallized in tetragonal space group $P4_12_12$ with one molecule in the asymmetric unit and diffracted to ~ 2.6 Å resolution. The structure of AAur1117 was determined via molecular replacement using the structure of Pa0148 as a search model. All five structures were refined to convergence with PHENIX.³⁰ Data collection and refinement statistics are summarized in Table 1.

Activity Screens. Pa0148, AAur1117, Sgx9403e, and Sgx9403g (10 nM) were incubated for 16 h with 2,6-diaminopurine, 6-mercaptopurine, 6-methoxypurine, 6-methylthiopurine, 6-chloropurine, 6-methylpurine, *N*-6-isopentyladenine, isoguanine, *N*-6-methyladenine, zeatin, adenine, adenosine, 2'-deoxyadenosine, 3'-deoxyadenosine, 5'-deoxyadenosine, 2',5'-dideoxyadenosine, AMP, ADP, ATP, *S*-adenosylhomocysteine, *S*-adenosylmethionine, *N*-6-methyl-2'-deoxyadenosine, 3',5'-cyclic AMP, 5'-thiomethyladenosine, 5'-methyldeoxycytidine, cytidine, 2'-deoxycytidine, cytosine, 5-hydroxymethylcytosine, 2-chloroadenine, or 2-dimethylaminoadenine. The substrate concentration was 80 μ M in each case. Enzymatic activity was monitored through changes in absorbance between 240 and 300 nm on a SPECTRAMax Plus spectrophotometer (Molecular Devices).

Measurement of Kinetic Constants. Assays were conducted with substrate concentrations of 3–200 μ M. Deamination of adenine and 2,6-diaminopurine was monitored by following the decrease in absorbance at 262 and 282 nm, respectively. Dechlorination of 6-chloropurine was monitored by following the increase in absorbance at 250 nm. Formation of hypoxanthine from *N*-6-methyladenine and 6-methoxypurine was monitored at 270 nm. Difference extinction coefficients were calculated by subtracting the extinction coefficient of the product from the extinction coefficient of the substrate for adenine ($\Delta\epsilon_{262} = 4600$ M⁻¹ cm⁻¹), *N*-6-methyladenine ($\Delta\epsilon_{270} = 15000$ M⁻¹ cm⁻¹), 2,6-diaminopurine ($\Delta\epsilon_{282} = 1600$ M⁻¹ cm⁻¹), 6-chloropurine

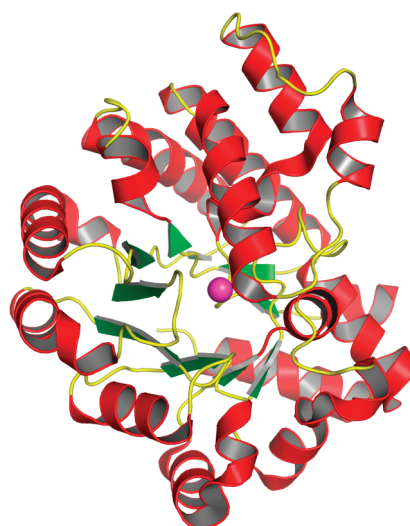


Figure 2. Three-dimensional structure of Pa0148 (PDB entry 3OU8). The eight β -strands that form the core (β/α)₈ barrel are colored green and the α -helices red. The single zinc cation is colored pink.

($\Delta\epsilon_{250} = 6100$ M⁻¹ cm⁻¹), and 6-methoxypurine ($\Delta\epsilon_{270} = 1900$ M⁻¹ cm⁻¹). Production of hypoxanthine was confirmed by mass spectrometry for the reactions of adenine, 6-chloropurine, and 6-methoxypurine. Deamination of 2,6-diaminopurine was confirmed using a coupled assay with guanine deaminase to produce the doubly deaminated product, xanthine. Production of methanol from 6-methoxypurine was detected using alcohol dehydrogenase (4 mg/mL) and monitoring the conversion of NAD⁺ to NADH at 340 nm.³¹

Metal Analysis. The metal content of the proteins was determined by ICP-MS.³² Protein samples for ICP-MS were digested with HNO₃ by refluxing for ~ 45 min to prevent protein precipitation during the measurement. The protein concentration was adjusted to ~ 1.0 μ M with 1% (v/v) HNO₃.

Data Analysis. Sequence alignments were created using ClustalW at <http://www.compbio.dundee.ac.uk/JalviewWS/services/ClustalWS>. Steady state kinetic data were analyzed using Softmax Pro, version 5.4. Kinetic parameters were determined by fitting the data to eq 1 using the nonlinear least-squares fitting program in SigmaPlot 9.0.

$$v/E_t = k_{cat}A/(K_m + A) \quad (1)$$

where A is the substrate concentration, K_m is the Michaelis constant, v is the velocity of the reaction, and k_{cat} is the turnover number.

RESULTS

Protein Purification. The four enzymes that are the focus of this investigation belong to group 3 of cog1816 (Figure 1). Pa0148 and AAur1117 were expressed in *E. coli* and purified to homogeneity. Pa0148 and AAur1117 contained, on average, 0.5 and 0.9 equiv of Zn²⁺ per monomer, respectively. The genes encoding Sgx9403e and Sgx9403g were synthesized and expressed in *E. coli*, and the resulting proteins were purified to homogeneity. Sgx9403e contained 0.6 equiv of Mn²⁺ per monomer, and Sgx9403g contained 0.8 equiv of Mn²⁺ per monomer. The role of the divalent cation in the active site of adenine deaminase is to facilitate the activation of the water molecule for

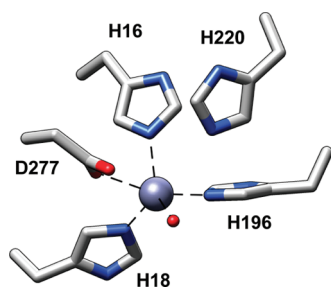


Figure 3. Active site of Pa0148 showing the residues that coordinate the divalent metal ion. The divalent cation is presented as a gray sphere, and the water molecule is colored red.

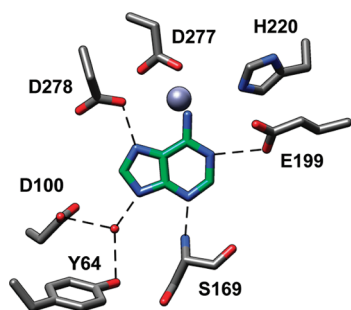


Figure 4. Active site of Pa0148 showing the residues and water molecule (red) that interact with the substrate adenine.

the hydrolytic deamination reaction.⁴ It is assumed that protein without a divalent cation in the active site is inactive. No attempt was made to determine the dissociation constant for the protein–metal complex.

Three-Dimensional Structures of Pa0148 and AAur1117. Pa0148 and AAur1117 were crystallized and their three-dimensional structures determined to resolution limits of ~ 2.5 and ~ 2.6 Å, respectively. A ribbon diagram illustrating the overall protein fold for Pa0148 is presented in Figure 2. The sequence identity of these two proteins is 53%, and the rmsd for the equivalent α -carbon atoms is 0.67 Å. The two proteins fold into a distorted (β/α)₈ barrel with a single divalent cation bound in the active site. The ligand-free structure of Pa0148 (PDB entry 3OU8) reveals those residues responsible for the coordination of zinc in the active site and the residues that stabilize the position of a metal-bound water molecule. Metal binding residues for Pa0148 include an HxH motif from the C-terminus of β -strand 1 (His-16 and His-18), a histidine from the C-terminus of β -strand 5 (His-196), and an aspartate from the C-terminus of β -strand 8 (Asp-277). The presumed hydrolytic water molecule is located 3.6 Å from the zinc ion. This water molecule is 3.6 Å from the side chain of His-220 and 3.0 Å from the side chain of Asp-277. The metal coordination scheme for Pa0148 is presented in Figure 3.

Cocrystal structures of Pa0148 bound to adenine (PDB entry 3PAO), hypoxanthine (PDB entry 3PAN), and 6-chloropurine (PDB entry 3PBM) were obtained by soaking preformed Pa0148 apoprotein crystals. Binding of adenine to Pa0148 is mediated by Tyr-64 and Asp-100 (Figure 4). These two residues hydrogen bond to a water molecule that is, in turn, hydrogen bonded to N9 of the purine ring. N3 of the purine ring hydrogen bonds to the backbone amide group of Ser-169. There are additional intermolecular interactions between Asp-278 and N7 and between

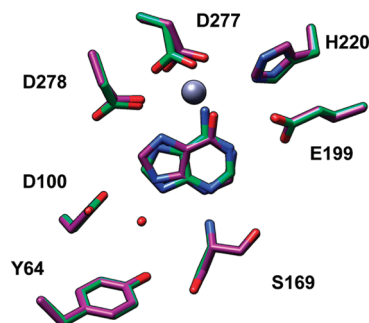


Figure 5. Active site of Pa0148 showing a comparison of the binding of the substrate adenine (PDB entry 3PAO) and the product hypoxanthine (PDB entry 3PAN). The divalent cation is presented as a gray sphere, and the water molecule is colored red.

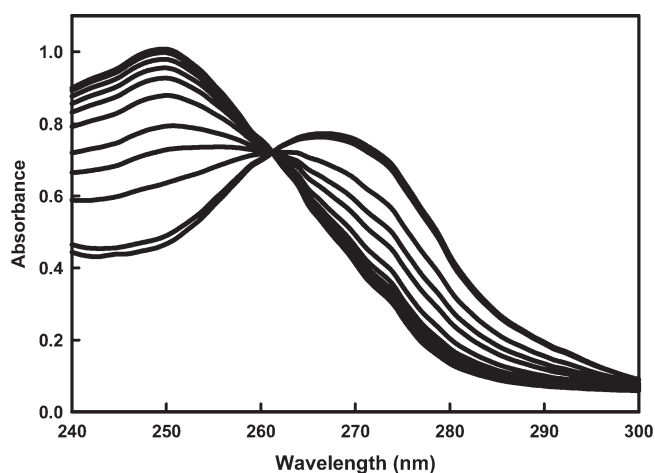


Figure 6. Time course for the change in absorbance as Pa0148 dechlorinates 6-chloropurine to hypoxanthine. The decrease in absorbance at 267 nm is coupled with an increase in absorbance at 250 nm as 6-chloropurine is converted to hypoxanthine.

Glu-169 and N1 of the purine ring. The putative hydrolytic water molecule is not observed in the active site of Pa0148 when either the substrate or products are bound in the active site. The structure of Pa0148 with hypoxanthine bound in the active site is nearly identical to the structure with adenine (rmsd = 0.11 Å) except that the substrate is rotated slightly. A similar rotation is apparent in the structure of 6-chloropurine bound to Pa0148; the rmsd between the two structures is 0.1 Å. An overlay of the active sites showing the relative positions of adenine and hypoxanthine bound to Pa0148 is presented in Figure 5. No conformational changes to the protein are apparent upon the binding of either substrate or products. The cocrystal structure of AAur1117 in a complex with adenine (PDB entry 3RYS) demonstrates a binding mode similar to those of 6-chloropurine and hypoxanthine in the active site of Pa0148.

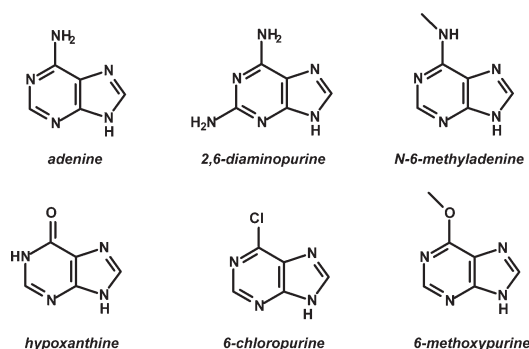
Substrate Specificity. The substrate profiles for Pa0148, AAur1117, Sgx9403e, and Sgx9403g were determined by monitoring changes in absorbance between 240 and 300 nm after the addition of enzyme to a library of modified purines and other nucleic acid derivatives. The best substrates for these four enzymes are adenine, 6-chloropurine, *N*-6-methyladenine, 6-methoxypurine, and 2,6-diaminopurine. All four enzymes were unable to

Table 2. Kinetic Constants for Pa0148, AAur1117, Sgx9403e, and Sgx9403g^a

enzyme	adenine			6-chloropurine			N-6-methyladenine			2,6-diaminopurine			6-methoxypurine		
	k_{cat} (s ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (M ⁻¹ s ⁻¹)	k_{cat} (s ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (M ⁻¹ s ⁻¹)	k_{cat} (s ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (M ⁻¹ s ⁻¹)	k_{cat} (s ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (M ⁻¹ s ⁻¹)	k_{cat} (s ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (M ⁻¹ s ⁻¹)
Pa0148	36	25	1.4×10^6	1.9	260	7.3×10^3	1.28	88	1.5×10^4	17	140	1.2×10^5	7.1	270	2.6×10^4
AAur1117	61	200	3.1×10^5	—	—	9.6×10^3	—	—	1.3×10^3	6.9	290	2.4×10^4	1.9	160	1.2×10^4
Sgx9403e	62	200	3.1×10^6	0.84	56	1.5×10^4	0.48	90	5.3×10^3	1.8	70	2.8×10^4	0.52	130	4.0×10^3
Sgx9403g	56	170	3.3×10^5	0.42	26	1.6×10^4	—	—	1.6×10^3	—	—	1.9×10^4	2.2	140	1.7×10^4

^a Standard errors for the kinetic constants are less than 20%.

Scheme 1



deaminate adenosine. The upper limit for the deamination of adenosine, relative to adenine, by Pa0148 is 0.01%. The change in absorbance as 6-chloropurine is converted to hypoxanthine in the presence of Pa0148 is shown in Figure 6. The absorbance decreases at 267 nm and increases at 250 nm. Production of hypoxanthine was confirmed by mass spectrometry for the reactions of adenine, 6-chloropurine, and 6-methoxypurine. A coupled assay with alcohol dehydrogenase was used to detect the formation of methanol from 6-methoxypurine by monitoring the conversion of NAD⁺ to NADH at 340 nm. The deamination of 2,6-diaminopurine was confirmed using a coupled assay with guanine deaminase for the ultimate formation of xanthine via guanine. The kinetic constants for the reactions of adenine, 6-chloropurine, N-6-methyladenine, 6-methoxypurine, and 2,6-diaminopurine are listed in Table 2. Adenine was the best overall substrate for all four enzymes examined, with $k_{\text{cat}}/K_{\text{m}}$ values exceeding $10^5 \text{ M}^{-1} \text{ s}^{-1}$.

DISCUSSION

Substrate Specificity. We have identified a large group of bacterial enzymes from cog1816 within the amidohydrolase superfamily of enzymes that are currently incorrectly annotated as adenosine deaminases. The best substrate for this group of enzymes identified thus far is adenine. Additional substrates include 2,6-diaminopurine, 6-chloropurine, 6-methoxypurine, and N-6-methyladenine (Scheme 1). These enzymes show no activity with adenosine. The ability to accept an amino substituent attached to C2 (for example, 2,6-diaminopurine) is similar to that of the authentic adenosine deaminases from cog1816, which will accept a methyl or amino group at this position, but not a halide substituent.³³ The 2-amino group is apparently able to fit into a pocket within the active site that contains two water molecules in the crystal structure of Pa0148 and AAur1117.

Hydrogen bonding interactions between the waters and the amino group help to stabilize the binding of 2,6-diaminopurine but apparently do not allow for substitution of a halide in this position. 2,6-Diaminopurine has been used in the treatment of cancer³⁴ and at one time was proposed to be an intermediate in the conversion of guanine to adenine in mammals.³⁵ More recently, 2,6-diaminopurine has been shown to bind to purine riboswitches.^{36,37}

The binding pocket near the N6 amino group of adenine in the crystal structure of Pa0148 is large enough to accommodate an additional methyl group, allowing for the binding and turnover of N-6-methyladenine and 6-methoxypurine. This observation is in agreement with the reported substrate specificity of the mammalian adenosine deaminase³⁸ and that of *E. coli* adenosine deaminase from cog1816 (data not shown), both of which accept the N-6-methyl and 6-methoxy derivatives of adenosine as substrates. N-6-Methyladenine is a naturally occurring base found in 1–2% of the total adenine content of bacterial genomes.³⁹ Methylation of adenine serves to protect host DNA from endonucleases that target foreign DNA and regulates the interaction between DNA and some DNA-binding proteins.^{39,40}

Deamination of N-6-methyladenine has been observed in cell free extracts from *E. coli*,⁴¹ and an enzyme has been identified with this catalytic activity from *Bacillus halodurans*.⁵ However, the N-6-methyladenine deaminase (6-MAD) from *B. halodurans* (Bh0637) is found in cog1001 from the amidohydrolase superfamily. The 6-MAD enzymes from cog1001 have a significantly higher activity with N-6-methyladenine than with adenine. With the group 3 enzymes from cog1816, adenine is deaminated ~2 orders of magnitude faster than N-6-methyladenine.

Sequence Analysis. The four enzymes analyzed in this investigation have sequences that are more than 50% identical (Figure 7). Compared to adenosine deaminase from *E. coli* (locus tag b1623) from group 5 of cog1816, the level of sequence identity is less than 30%. All of the residues that coordinate the divalent cation in the active site are conserved among the group 5 and group 3 enzymes, including the HxH motif following β -strand 1, a histidine at the C-terminus of β -strand 5, and an aspartate at the C-terminus of β -strand 8. These residues are colored red in Figure 7. Additional conserved residues include an aspartate that follows the invariant aspartate that is coordinated to the divalent cation found at the C-terminal end of β -strand 8. This residue hydrogen bonds to N7 of the adenine ring. Also conserved among the group 3 and group 5 enzymes of cog1816 is the glutamate from the HxxE motif at the C-terminus of β -strand 5. This residue is responsible for the protonation of N1 during the deamination reaction. The other functionally important residue is the histidine at the C-terminus of β -strand 6 that appears to form a hydrogen bond with the putative hydrolytic water and

may participate in proton transfer reactions. These residues are colored blue in Figure 7.

Pa0148 and the other members of group 3 of cog1816 share a four-residue deletion between β -strands 1 and 2. In addition, group 3 proteins possess a conserved tyrosine repeat between β -strands 1 and 2. This pair of tyrosine residues in the crystal structures of Pa0148 and AAur1117 occurs within the last of three α -helices connecting β -strands 1 and 2. The second tyrosine, together with an aspartate from the C-terminal end of β -strand 4, is responsible for positioning a water molecule to form a hydrogen bond with N3 of adenine. Adenosine is apparently unable to bind in the active site of group 3 proteins because the first tyrosine occludes ribose binding (see Figure 8).¹⁷ A further impediment to ribose binding is an aspartate residue at the C-terminus of β -strand 2 that occupies the binding site of the 5'-hydroxyl group of adenosine for the group 5 adenosine deaminases. In Pa0148, this residue is Asp-168. Adenine binding residues are colored green in Figure 7. The four-residue deletion, the two conserved tyrosines following β -strand 1, the aspartate following β -strand 2, and the aspartate from β -strand 4, are fully conserved in Pa0148 and approximately 120 other proteins from group 3 of cog1816. We conclude that these enzymes are adenine deaminases, not adenosine deaminases, as currently annotated. These proteins are listed in Table S1 of the Supporting Information. These adenine deaminases are similar in amino acid sequence to a previously described fungal adenine deaminase.⁴²

Comparison to Adenine Deaminase from cog1001. The prototypical adenine deaminase from *E. coli* (locus tag b3665) is also a member of the amidohydrolase superfamily but belongs to cog1001. The *E. coli* adenine deaminase has a binuclear metal center with a high sensitivity to iron and will apparently accept no substrate other than adenine.⁴³ The active site of the *E. coli* adenine deaminase resembles the active sites of urease and phosphotriesterase. It is unclear why some bacteria possess a mononuclear adenine deaminase and others from the same class of bacteria possess a binuclear adenine deaminase. The adenine deaminases from cog1001 and cog1816 have nearly equal representation among the bacterial genomes. Although there is overlap between the types of bacteria that possess a specific adenine deaminase, each species of bacteria appears to possess only one adenine deaminase.

Strategy for Functional Annotation. The COG database was used to construct a sequence similarity network of cog1816, which identified enzymes belonging to group 3 as significantly different from experimentally verified adenosine deaminases in group 5. This divergence and the absence of key residues in sequence alignments between proteins in groups 3 and 5 lead to the prediction that proteins within group 3 would have a unique substrate profile, other than the deamination of adenosine. After the structure for Pa0148 had been determined, modeling of adenosine in the active site of this protein demonstrated that the ribose binding region was obscured by residues found only in proteins of group 3. This observation prompted the successful search for the true substrate for Pa0148 and related proteins, using adenine and modified purines. The subsequent determination of the structure of Pa0148 in the presence of adenine and hypoxanthine illustrated how these compounds are recognized within the active site. A similar inspection of the other major groups within cog1816 indicates that they almost certainly catalyze the deamination of substrates other than adenosine.

■ ASSOCIATED CONTENT

S Supporting Information. A list of proteins from cog1816 that are predicted to catalyze the deamination of adenine (Table S1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ Accession Codes

The X-ray coordinates and structure factors have been deposited in the Protein Data Bank as entries 3OU8, 3PAN, 3PAO, 3PBM, and 3RYS.

■ AUTHOR INFORMATION

Corresponding Author

*F.M.R.: telephone, (979) 845-3373; fax, (979) 845-9452; e-mail, raushel@tamu.edu. S.S.: telephone, (631) 344-3187; fax, (631) 344-3407; e-mail, swami@bnl.gov.

Funding Sources

This work was supported in part by the National Institutes of Health (GM 71790 and GM 74945) and the Robert A. Welch Foundation (A-840).

■ ACKNOWLEDGMENT

We gratefully acknowledge data collection support from the National Synchrotron Light Source. We also thank the NYSGXRC protein production team for preparation of Pa0148, Sgx9403e, and Sgx9403g.

■ ABBREVIATIONS

ADE, adenine deaminase; ADD, adenosine deaminase; COG, cluster of orthologous groups; IPTG, isopropyl β -galactoside; DTT, dithiothreitol; AHS, amidohydrolase superfamily; ICP-MS, inductively coupled plasma mass spectrometry; 6-MAD, N-6-methyladenine deaminase; rmsd, root-mean-square deviation.

■ REFERENCES

- (1) Pegg, S. C., Brown, S. D., Ojha, S., Seffernick, J., Meng, E. C., Morris, J. H., Chang, P. J., Huang, C. C., Ferrin, T. E., and Babbitt, P. C. (2006) Leveraging enzyme structure-function relationships for functional inference and experimental design: The structure-function linkage database. *Biochemistry* 45, 2545–2555.
- (2) Schnoes, A. M., Brown, S. D., Dodevski, I., and Babbitt, P. C. (2009) Annotation error in public databases: Misannotation of molecular function in enzyme superfamilies. *PLoS Comput. Biol.* 5, e1000605.
- (3) Holm, L., and Sander, C. (1997) An evolutionary treasure: Unification of a broad set of amidohydrolases related to urease. *Proteins* 28, 72–82.
- (4) Seibert, C. M., and Raushel, F. M. (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* 44, 6383–6391.
- (5) Kamat, S. S., Fan, H., Sauder, J. M., Burley, S. K., Shoichet, B. K., Sali, A., and Raushel, F. M. (2011) Enzymatic deamination of the epigenetic base N-6-methyladenine. *J. Am. Chem. Soc.* 133, 2080–2083.
- (6) Hall, R. S., Agarwal, R., Hitchcock, D., Sauder, J. M., Burley, S. K., Swaminathan, S., and Raushel, F. M. (2010) Discovery and structure determination of the orphan enzyme isoxanthopterin deaminase. *Biochemistry* 49, 4374–4382.
- (7) Wang, Z., and Quioco, F. A. (1998) Complexes of adenosine deaminase with two potent inhibitors: X-ray structures in four independent molecules at pH of maximum activity. *Biochemistry* 37, 8314–8324.

- (8) Ho, M. C., Cassera, M. B., Madrid, D. C., Ting, L. M., Tyler, P. C., Kim, K., Almo, S. C., and Schramm, V. L. (2009) Structural and metabolic specificity of methylthioformycin for malarial adenosine deaminases. *Biochemistry* 48, 9618–9626.
- (9) Baer, H. P., Drummond, G. I., and Gillis, J. (1968) Studies on the specificity and mechanism of action of adenosine deaminase. *Arch. Biochem. Biophys.* 123, 172–178.
- (10) Kinoshita, T., Nakanishi, I., Terasaka, T., Kuno, M., Seki, N., Warizaya, M., Matsumura, H., Inoue, T., Takano, K., Adachi, H., Mori, Y., and Fujii, T. (2005) Structural basis of compound recognition by adenosine deaminase. *Biochemistry* 44, 10562–10569.
- (11) Kinoshita, T., Tada, T., and Nakanishi, I. (2008) Conformational change of adenosine deaminase during ligand-exchange in a crystal. *Biochem. Biophys. Res. Commun.* 373, 53–57.
- (12) Koch, A. L., and Vallee, G. (1959) The properties of adenosine deaminase and adenosine nucleoside phosphorylase in extracts of *Escherichia coli*. *J. Biol. Chem.* 234, 1213–1218.
- (13) Lin, J., Westler, W. M., Cleland, W. W., Markley, J. L., and Frey, P. A. (1998) Fractionation factors and activation energies for exchange of the low barrier hydrogen bonding proton in peptidyl trifluoromethyl ketone complexes of chymotrypsin. *Proc. Natl. Acad. Sci. U.S.A.* 95, 14664–14668.
- (14) Gleeson, M. P., Burton, N. A., and Hillier, I. H. (2003) Prediction of the potency of inhibitors of adenosine deaminase by QM/MM calculations. *Chem. Commun.*, 2180–2181.
- (15) Sadat Hayatshahi, S. H., Abdolmaleki, P., Ghiasi, M., and Safarian, S. (2007) QSARs and activity predicting models for competitive inhibitors of adenosine deaminase. *FEBS Lett.* 581, 506–514.
- (16) Wilson, D. K., Rudolph, F. B., and Quioco, F. A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: Understanding catalysis and immunodeficiency mutations. *Science* 252, 1278–1284.
- (17) Larson, E. T., Deng, W., Krumm, B. E., Napuli, A., Mueller, N., Van Voorhis, W. C., Buckner, F. S., Fan, E., Lauricella, A., DeTitta, G., Luft, J., Zucker, F., Hol, W. G., Verlinde, C. L., and Merritt, E. A. (2008) Structures of substrate- and inhibitor-bound adenosine deaminase from a human malaria parasite show a dramatic conformational change and shed light on drug selectivity. *J. Mol. Biol.* 381, 975–988.
- (18) Weiss, P. M., Cook, P. F., Hermes, J. D., and Cleland, W. W. (1987) Evidence from nitrogen-15 and solvent deuterium isotope effects on the chemical mechanism of adenosine deaminase. *Biochemistry* 26, 7378–7384.
- (19) Sideraki, V., Mohamedali, K. A., Wilson, D. K., Chang, Z., Kellems, R. E., Quioco, F. A., and Rudolph, F. B. (1996) Probing the functional role of two conserved active site aspartates in mouse adenosine deaminase. *Biochemistry* 35, 7862–7872.
- (20) Atkinson, H. J., Morris, J. H., Ferrin, T. E., and Babbitt, P. C. (2009) Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. *PLoS One* 4, e4345.
- (21) Mohamedali, K. A., Kurz, L. C., and Rudolph, F. B. (1996) Site-directed mutagenesis of active site glutamate-217 in mouse adenosine deaminase. *Biochemistry* 35, 1672–1680.
- (22) Sideraki, V., Wilson, D. K., Kurz, L. C., Quioco, F. A., and Rudolph, F. B. (1996) Site-directed mutagenesis of histidine 238 in mouse adenosine deaminase: Substitution of histidine 238 does not impede hydroxylate formation. *Biochemistry* 35, 15019–15028.
- (23) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
- (24) Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography*, Vol. 26.
- (25) Collaborative Computational Project Number 4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* 50, 760–763.
- (26) Hattne, J., and Lamzin, V. S. (2008) Pattern-recognition-based detection of planar objects in three-dimensional electron-density maps. *Acta Crystallogr. D* 64, 834–842.
- (27) Langer, G., Cohen, S. X., Lamzin, V. S., and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* 3, 1171–1179.
- (28) Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* 60, 2126–2132.
- (29) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658–674.
- (30) Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) PHENIX: Building new software for automated crystallographic structure determination. *Acta Crystallogr. D* 58, 1948–1954.
- (31) Bergmeyer, H. U., Ed. (1974) *Methods of Enzymatic Analysis*, Academic Press, New York.
- (32) Hall, R. S., Xiang, D. F., Xu, C., and Raushel, F. M. (2007) N-Acetyl-D-glucosamine-6-phosphate deacetylase: Substrate activation via a single divalent metal ion. *Biochemistry* 46, 7942–7952.
- (33) Yokozeki, K., and Tsuji, T. (2000) A novel enzymatic method for the production of purine-2'-deoxyribonucleosides. *J. Mol. Catal. B: Enzym.* 10, 207–213.
- (34) Burchenal, J. H., Karnofsky, D. A., Kingsley-Pillers, E. M., Southam, C. M., Myers, W. P., Escher, G. C., Craver, L. F., Dargeon, H. W., and Rhoads, C. P. (1951) The effects of the folic acid antagonists and 2,6-diaminopurine on neoplastic disease, with special reference to acute leukemia. *Cancer* 4, 549–569.
- (35) Bendich, A., Furst, S. S., and Brown, G. B. (1950) On the role of 2,6-diaminopurine in the biosynthesis of nucleic acid guanine. *J. Biol. Chem.* 185, 423–433.
- (36) Wickiser, J. K., Cheah, M. T., Breaker, R. R., and Crothers, D. M. (2005) The kinetics of ligand binding by an adenine-sensing riboswitch. *Biochemistry* 44, 13404–13414.
- (37) Gilbert, S. D., Mediatore, S. J., and Batey, R. T. (2006) Modified pyrimidines specifically bind the purine riboswitch. *J. Am. Chem. Soc.* 128, 14214–14215.
- (38) Orsi, B. A., McFerran, N., Hill, A., and Bingham, A. (1972) Kinetics and the mechanism of action of adenosine aminohydrolase. *Biochemistry* 11, 3386–3392.
- (39) Ratel, D., Ravanat, J. L., Berger, F., and Wion, D. (2006) N6-methyladenine: The other methylated base of DNA. *BioEssays* 28, 309–315.
- (40) Wion, D., and Casadesus, J. (2006) N6-methyl-adenine: An epigenetic signal for DNA-protein interactions. *Nat. Rev. Microbiol.* 4, 183–192.
- (41) Remy, C. N. (1961) Metabolism of 6-methylaminopurine: Synthesis and demethylation by *Escherichia coli*. *J. Biol. Chem.* 236, 2999–3005.
- (42) Ribard, C., Rochet, M., Labedan, B., Daignan-Fornier, B., Alzari, P., Scazzocchio, C., and Oestreich, N. (2003) Sub-families of α/β barrel enzymes: A new adenine deaminase family. *J. Mol. Biol.* 334, 1117–1131.
- (43) Kamat, S. S., Bagaria, A., Kumaran, D., Holmes-Hampton, G. P., Fan, H., Sali, A., Sauder, J. M., Burley, S. K., Lindahl, P. A., Swaminathan, S., and Raushel, F. M. (2011) Catalytic Mechanism and Three-Dimensional Structure of Adenine Deaminase. *Biochemistry* 50, 1917–1927.