Published in final edited form as:
Biochemistry. 2013 September 17; 52(37): . doi:10.1021/bi400750a.

Deamination of 6-Aminodeoxyfutalosine in Menaquinone Biosynthesis by Distantly Related Enzymes

Alissa M. Goble[§], **Rafael Toro**[Ø], **Xu Li**[∂], **Argentina Ornelas**[§], **Hao Fan**[φ,η], **Subramaniam Eswaramoorthy**[‡], **Yury Patskovsky**[Ø], **Brandan Hillerich**[Ø], **Ron Seidel**[Ø], **Andrej Sali**[η], **Brian K. Shoichet**[φ,η], **Steven C. Almo**[Ø], **Subramanyam Swaminathan**[‡], **Martin E. Tanner**[∂], and **Frank M. Raushel**[§,*]

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

[φ]Department of Pharmaceutical Chemistry, University of California, San Francisco, 1700 Fourth Street, San Francisco, California 94158

[η]California Institute for Quantitative Biosciences, University of California, San Francisco, 1700 Fourth Street, San Francisco, California 94158

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

[∂]Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, British Columbia V6T 1Z1, Canada

[Ø]Department of Biochemistry, Einstein College of Medicine, Bronx, NY

Abstract

Proteins of unknown function belonging to cog1816 and cog0402 were characterized. Sav2595 from *Steptomyces avermitilis* MA-4680, Acel0264 from *Acidothermus cellulolyticus* 11B, Nis0429 from *Nitratiruptor* sp. SB155-2 and Dr0824 from *Deinococcus radiodurans* R1 were cloned, purified, and their substrate profiles determined. These enzymes were previously incorrectly annotated as adenosine deaminases or chlorohydrolases. It was shown here that these enzymes actually deaminate 6-aminodeoxyfutalosine. The deamination of 6-aminodeoxyfutalosine is part of an alternative menaquinone biosynthetic pathway that involves the formation of futalosine. 6-Aminodeoxyfutalosine is deaminated by these enzymes with catalytic efficiencies greater than 10^5 M^{-1} s^{-1}, K_m values of 0.9 to 6.0 μM and k_{cat} values of 1.2 to 8.6 s^{-1}. Adenosine, 2′-deoxyadenosine, thiomethyladenosine, and S-adenosylhomocysteine are deaminated at least an order of magnitude slower than 6-aminodeoxyfutalosine. The crystal structure of Nis0429 was determined and the substrate, 6-aminodeoxyfutalosine, was positioned in the active site, based on the presence of adventitiously bound benzoic acid. In this model Ser-145 interacts with the carboxylate moiety of the substrate. The structure of Dr0824 was also determined, but a collapsed active site pocket prevented docking of substrates. A computational model of Sav2595 was built based on the crystal structure of adenosine deaminase and substrates were docked. The model predicted a conserved arginine after β-strand 1 to be partially responsible for the substrate specificity of Sav2595.

Menaquinone can be biosynthesized through either the shikimate pathway or the futalosine pathway. In *Escherichia coli*, the well-studied shikimate pathway utilizes seven enzymes

Corresponding Author: (F.M.R.) Address: Department of Chemistry, P.O. Box 30012, Texas A&M University, College Station, TX 77843-3012., Tel: (979) 845-3373. raushel@tamu.edu.

(*menA* to *menG*) to convert chorismate into menaquinone (1). Recently, menaquinone producing organisms were identified that lack *menF*, *menD*, *menC*, *menE*, and *menB* orthologs (2, 3). These organisms, including *Streptomyces coelicolor* A3(2), *Helicobacter pylori*, *Campylobacter jejuni*, and *Acidothermus cellulolyticus*, lack *men* gene homologs and were suspected of employing a different pathway for the biosynthesis of menaquinone (1, 4). Using a combination of bioinformatics and gene disruption experiments, the *mqn* genes involved in the futalosine pathway were identified (Figure 1) (1). *Mqn* genes are found scattered throughout the genomes, with the exception of *A. cellulolyticus*, which has two clusters of menaquinone related genes (4). The smaller cluster contains *MqnC* and *MqnB* orthologs and the other cluster is made up of several genes, including *MqnA*, another *MqnC* and an adenosine deaminase ortholog (Acel0264) (Figure 2).

Adenosine deaminase is a member of the amidohydrolase superfamily (AHS) of enzymes: a superfamily first identified by Holm and Sander based on the similarities in the three-dimensional structures of phosphotriesterase, adenosine deaminase and urease (5). The common fold within the AHS is a distorted $(\beta/\alpha)_8$-barrel with conserved metal binding residues at the C-terminal ends of β-strands 1, 4, 5, 6, and 8. Enzymes in the AHS have either a mononuclear or binuclear metal center that functions to activate a water molecule for nucleophilic attack on amino acids, sugars, nucleic acids and organophosphate esters (6). The AHS is composed of 24 clusters of orthologous groups (COG). Enzymes that are known to deaminate aromatic bases are found in cog1001, cog0402 and cog1816. Members of cog1001 include the prototypical adenine deaminase (ADE) and *N*-6-methyladenine deaminase (7). Enzymes that deaminate guanine, cytosine, *S*-adenosylhomocysteine, thiomethyl adenosine, *N*-formimino-L-glutamate, and 8-oxoguanine are found in cog0402 (8). According to annotations reported by NCBI, bacterial enzymes in cog1816 function in the deamination of adenosine. The functional diversity of cog1816 has recently been expanded to include enzymes that deaminate adenine and cytokinins (9, 10). The mechanism, structure and biological function of adenosine deaminase enzymes from several organisms have been investigated (11–23). A sequence similarity network for cog1816 is illustrated in Figure 3 at a BLAST E-value of 1×10^{-70}, which corresponds to a sequence identity within each group of ~30% (24). In general, members of a group share a common function, and the substrate profiles differ between groups. Members of group 5 of cog1816 are prototypical adenosine deaminases, including the adenosine deaminase from *E. coli* K-12. Group 3 of cog1816 includes adenine deaminases with a single divalent cation and enzymes that catalyze the deamination of cytokinins.

The initial targets for this investigation were the functional annotation of Acel0264 from *A. cellulolyticus* and Sav2595 from *Streptomyces avermitilis* MA-4680. These enzymes belong to groups 11 and 13 of cog1816, respectively. The substrate profile for these two enzymes has been established and a three-dimensional model has been proposed for how the best substrate, 6-aminodeoxyfutalosine (AFL), binds in the active site. Based on the absence of *men* genes in the genomes of *Deinococcus radiodurans* R1 and *Nitratiruptor* sp. SB155-2, two distantly related enzymes from cog0402, Dr0824 from *D. radiodurans* R1 and Nis0429 from *N.* sp. SB155-2, were identified as 6-aminodeoxyfutalosine deaminases. The three dimensional structures of Dr0824 and Nis0429 were determined by X-ray crystallography. A sequence similarity network of cog0402, which includes guanine deaminase, cytosine deaminase and other deaminases, is shown in Figure 4.

MATERIALS and METHODS

Materials

All chemicals were purchased from Sigma-Aldrich, unless otherwise stated. 6-Aminodeoxyfutalosine (AFL) was synthesized as previously described (25). 1-(6-

Amino-9H-purin-9-yl)-1-deoxy-*N*-ethyl-β-D-ribofuranuronamide (NECA) was purchased from Ascent Scientific.

Cloning and Purification of Sav2595 from *Streptomyces avermitilis* MA-4680 and Dr0824 from *Deinococcus radiodurans* R1

The genes for Sav2595 and Dr0824 were cloned from the genomic DNA of *Streptomyces avermitilis* MA-4680 and *Deinococcus radiodurans* R1, respectively. The PCR products were amplified using the primer pairs 5′-AGAGCCATGGTGACCGAGCACTTCGACGC -3′ and 5′-AGAGAAGCTTCAGGGCGCGAGCCATGCC-3′, and pair 5′-GGGAATTACCATATGCGCTTTTCTGCCGTCAGCCG -3′ and 5′-CCCAAGCTTTCATCACAAGTCGCGGCTCAACTCCCAACG -3′, respectively. The restriction sites for *Nco*I and *Hin*dIII for Sav2595, and *Nde*I and *Hin*dIII for Dr0824 were inserted into the forward and reverse primers, respectively. The PCR products were purified with a PCR clean-up system (Qiagen), digested with their respective restriction enzymes and ligated into a pET-30a(+) vector, which was previously digested with the same two restriction enzymes. The cloned gene fragments were sequenced to verify fidelity of the PCR amplification. Dr0824 contains a C-terminal polyhistidine tag, whereas Sav2595 has an N-terminal polyhistidine tag.

Sav2595-pET-30a(+) and Dr0824-pET30a(+) constructs were transformed into BL21(DE3) cells (Novagen). A single colony was used to inoculate a 5 mL overnight culture of LB medium containing 50 μg/mL kanamycin. Each overnight culture was used to inoculate 1.0 L of LB medium containing 50 μg/mL kanamycin. One liter cultures were grown at 37 °C, supplemented with 1.0 mM $ZnCl_2$ and induced with 0.5 mM isopropyl D-thiogalactopyranoside (IPTG) when an OD_{600} of 0.6 was reached. Cultures of Dr0824 were induced with 1.0 mM IPTG. At the time of induction, the temperature was lowered to 20 °C and allowed to shake for 18 hours before the cells were harvested by centrifugation at 8,000 rpm for 10 minutes. The cells were resuspended in 20 mM HEPES buffer, pH 7.5, containing 20 mM imidazole, 120 mM ammonium sulfate, 10% glycerol, 0.1 mg/mL phenylmethanesulfonyl fluoride (PMSF) and 0.5 mg/mL deoxyribonuclease I from bovine pancreas. Cells were lysed by sonication and the soluble protein was separated from the cell debris by centrifugation. Soluble protein from cultures of Sav2595 was loaded onto a HisTrap column (GE Healthcare) and eluted with a linear gradient of a solution containing 20 mM HEPES, pH 7.5, imidazole, 500 mM imidazole, and 500 mM ammonium sulfate. Soluble protein from cultures of Dr0824 was partitioned by ammonium sulfate precipitation at 30, 50 and 70% saturation. The precipitated protein (30–50% ammonium sulfate saturation) was re-suspended in 50 mM HEPES buffer (pH 7.7) and applied to a High Load 26/60 Superdex 200 preparative grade gel filtration column (GE Healthcare). Fractions containing protein were pooled and passed through an anion exchange column.

Cloning and Purification of Acel0264 from *Acidothermus cellulolyticus*

The gene for Acel0264 was amplified from *Acidothermus cellulolyticus* Strain 11B genomic DNA using 5′-TTAAGAAGGAGATATACCATGGTGATGACACCCCACGATCCCGTC -3′ as the forward primer and 5′-GATTGGAAGTAGAGGTTCTCTGCCAGCGACTCTCCCGCCGCTG -3′ as the reverse primer. PCR was performed using KOD Hot Start DNA Polymerse (Novagen). The amplified fragment was cloned into the C-terminal TEV cleavable C-terminal StrepII-6x-His-tag containing vector, CHS23, by ligation-independent cloning (26).

The gene for Acel0264 inserted in a CHS23 vector was transformed into Rosetta2 (DE3) cells (Novagen). A single colony was used to inoculate a 5 mL overnight culture of LB medium containing 50 μg/mL kanamycin and 25 μg/mL chloramphenicol. Each overnight

culture was used to inoculate 1.0 L of LB medium containing 50 μg/mL kanamycin and 25 μg/mL chloramphenicol. One liter cultures were grown at 37 °C. At an OD_{600} of 0.6, 1.0 mM $ZnCl_2$ was added, followed by induction with 50 μM IPTG. At the time of induction, the temperature was lowered to 20 °C and allowed to shake for 18 hours before the cells were harvested by centrifugation at 8,000 rpm for 10 minutes. The cells were resuspended in 20 mM HEPES buffer, pH 6.5, with 120 mM ammonium sulfate and 20 mM imidazole. Cells were lysed by sonication in the presence of 0.1 mg/mL PMSF and 0.5 mg/mL deoxyribonuclease I from bovine pancreas. Soluble protein was separated from the cell debris by centrifugation. The protein was then loaded onto a HisTrap HP column (GE Healthcare) and eluted with a gradient of buffer containing 20 mM HEPES, 500 mM imidazole and 500 mM ammonium sulfate, pH 6.5.

Cloning and Purification of Nis0429 from *Nitratiruptor sp*. SB155-2

The gene encoding Nis0429 was chemically synthesized and cloned into the plasmid pUC57 (GenScript). Expression constructs were generated by PCR amplification using the 5′-TTAAGAAGGAGATATACCATGCGTATCATTAAGCCGTTCGCC-3′ as the forward primer and the 5′-GATTGGAAGTAGAGGTTCTCTGCTTCGCGGACGTGTTCTTCGCC-3′ as the reverse primer. PCR was performed using KOD Hot Start DNA Polymerase (Novagen). The conditions were: 2 minutes at 95 °C, followed by 40 cycles of 30 seconds at 95 °C, 30 seconds at 66 °C, and 30 seconds at 72 °C. The amplified fragment was cloned into the C-terminal TEV cleavable C-terminal StrepII-6x-His-tag containing vector, CHS30, by ligation-independent cloning (26).

Expression vectors were transformed into *E. coli* BL21(DE3) containing the pRIL plasmid (Stratagene) and used to inoculate a 25 mL M9 minimal media culture containing 25 μg/mL kanamycin and 34 μg/mL chloramphenicol. The culture was allowed to grow overnight at 37 °C in a shaking incubator. A 10 mL aliquot of the overnight culture was used to inoculate 2 L of M9 SeMET high-yield medium (Shanghai Medicilon) supplemented with 150 mM 2, 2′-bipyridyl, 1.0 mM $ZnCl_2$, and 1.0 mM $MnCl_2$. The culture was placed in a LEX48 airlift fermenter and incubated at 37 °C until the OD600 reached 1.2. Cells were induced by addition of 0.5 mM IPTG and allowed to grow for 16 hours at 22 °C overnight. The culture was harvested and pelleted by centrifugation.

Cells were re-suspended in lysis buffer containing 20 mM HEPES, pH 7.5, 500 mM NaCl, 20 mM imidazole, and 10% glycerol and lysed by sonication. Lysates were clarified by centrifugation at 35,000g for 30 minutes. Proteins were purified on an AKTAxpress FPLC (GE Healthcare). Clarified lysates were loaded onto a 5 mL Strep-Tactin column (IBA), washed with 5 column volumes of lysis buffer, and eluted with 20 mM HEPES, pH 7.5, 500 mM NaCl, 20 mM imidazole, 10% glycerol, and 2.5 mM desthiobiotin. The eluent was loaded onto a 1 mL Histrap HP column (GE Healthcare), washed with 10 column volumes of lysis buffer, and eluted with 20 mM HEPES, pH 7.5, 500 mM NaCl, 500 mM imidazole, and 10% glycerol. The purified sample was applied to a HiLoad S200 16/60 HR gel filtration column which was equilibrated in 20 mM HEPES, pH7.5, 150 mM NaCl, 10% glycerol, and 5 mM DTT. Peak fractions were collected and protein was analyzed by SDS-PAGE. Samples were concentrated to 18 mg/mL using Amicon Ultra centrifugal filters (Millipore), snap frozen in liquid nitrogen, and stored at −80 °C.

Initial Activity Screens

Sav2595, Acel0264, Dr0824, and Nis0429 (5 IM) were incubated for 16 hours with AFL, adenine, adenosine, 2′-deoxyadenosine, 3′-deoxyadenosine, 5′-deoxyadenosine, 2′-5′-dideoxyadenosine, AMP, ADP, ATP, *S*-adenosylhomocysteine, *S*-adenosyl methionine,

N-6-methyl-2′-deoxyadenosine, 5′-thiomethyladenosine, formaminopyrimidine-adenine, and NECA. The substrate concentration was 80 μM. Enzymatic activity was monitored by changes in absorbance between 240–300 nm with a SPECTRAmax Plus spectrophotometer (Molecular Devices).

Measurement of Kinetic Constants

Assays were conducted with substrate concentrations of 4–200 μM. The deamination of AFL, NECA, adenosine, AMP, 2′-deoxyadenosine, 3′-deoxyadenosine, 5′-deoxyadenosine, SAH and 5′-methylthioadenosine was monitored by following the decrease in absorbance at 263 nm. Differential extinction coefficients were calculated by subtracting the extinction coefficient of the product from that of the substrate for AFL, adenosine, NECA, 5′-methylthioadenosine, 5′-deoxyadenosine, SAH and AMP ($\Delta\varepsilon_{263}$ = 6900 M^{-1} cm^{-1}).

Metal Analysis

Metal content of the proteins was determined by ICP-MS (27). Protein samples for ICP-MS were digested with HNO_3 by refluxing for ~45 minutes to prevent protein precipitation during the measurement. Protein concentration was adjusted to ~1.0 μM with 1% (v/v) HNO_3.

Data Analysis

Sequence alignments were created using Clustal (28). Steady state kinetic data were analyzed using Softmax Pro, version 5.4 (Molecular Devices). Kinetic parameters were determined by fitting the data to equation 1 using the nonlinear least-squares fitting program in SigmaPlot 11.0 (Systat Software Inc.). In this equation, 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.

$$v/\mathrm{E_t} = k_{\mathrm{cat}}\,\mathrm{A}/(K_{\mathrm{m}}+\mathrm{A}) \quad (1)$$

Homology Modeling and Ligand Docking

The X-ray structure of adenosine deaminase from *Bos taurus* with a transition-state analogue bound in the active site (PDB ID: 1KRM, 26% and 25% sequence identity to Acel0264 and Sav2595, respectively) was used as the template to build homology models for Acel0264 and Sav2595. The sequence alignment was computed with Multiple Sequence Comparison by Log-Expectation (MUSCLE) (29) followed by generation of 500 homology models using the standard “automodel” class in MODELLER 9v12 (30). The model with the best DOPE (31) score was selected and this initial model was refined with the “loopmodel” class in MODELLER and the side chain prediction protocol in PLOP (32). After refinement, the two homology models of Acel0264 and Sav2595 were targeted by docking screens using DOCK 3.6 (33, 34) with a virtual library (35–37) that contains NECA, AFL, as well as 57680 different high-energy intermediates (HEI) of 6440 KEGG (Kyoto Encyclopedia of Genes and Genomes) molecules (38, 39). In docking screens, a computed pose was selected if the O^- moiety of the HEI portion of the molecule was found within 4 Å of the metal ion in the active site. The selected molecules were ranked by an energy function consisting of protein-ligand van der Waals interactions, protein-ligand electrostatic interactions, and a correction for ligand desolvation. The top 500 ranked molecules were inspected visually to ensure the compatibility of the pose with the amidohydrolase reaction mechanism.

Mutation of Sav2595 and Nis0429

Single-site mutations of Sav2595 were constructed using the standard QuikChange PCR protocol according to the manufacturer's instructions. The mutants, R87A and R87M, were expressed and purified using the protocol established for wild-type Sav2595. A single-site mutation was made to Nis0429 through standard QuikChange PCR protocols to create the S145A mutant.

Structure Determination of Dr0824

Crystals were grown by mixing 1 μL protein with 1 μL reservoir solution via the sitting drop vapor diffusion method. Protein was in HEPES buffer at pH 7.5 at a concentration of approximately 14 mg/mL. Reservoir solution contained 30% PEG MME500, 0.1 M BisTris, pH 6.5, and 50 mM calcium chloride. Harvested crystals were mounted in cryo-loops and immediately flash-cooled in liquid nitrogen. Diffraction from these crystals was consistent with the monoclinic space group C2 and data extending to 1.78 Å resolution collected on the X12C beam-line of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. Data were processed, integrated, and scaled using HKL2000 (40); data collection parameters are given in Table 1. The structure was determined by single anomalous dispersion (SAD) using Se-Met substituted crystals. Se-Met crystals were obtained under the same conditions that yielded native crystals. Se atoms were located using SOLVE and the phases were refined with SHARP (41, 42). The initial protein model was built by an automated procedure using ARP/wARP software (43). The model was completed manually using 'O' and water molecules added based on the Fourier difference map (44). The structure was refined with CNS 1.1 and the refinement statistics are presented in Table 1 (45). The final model contains a zinc ion in addition to protein atoms and 170 water molecules (PDB ID: 2IMR).

Structure Determination of Nis0429

Se-Met derivatized protein was crystallized by the sitting-drop vapor diffusion method, as described below. 0.5 μL of the protein solution (18 mg/mL) was mixed with an equal volume of a precipitant solution and equilibrated at room temperature (294 K) against the same precipitant solution in clear tape-sealed 96-well INTELLI-plates. Crystallization was performed using a TECAN crystallization robot. A single crystal appeared approximately 11 months after setting up crystallization drops in 0.1 M MES/NaOH, pH 6.0, 1 M potassium sodium tartrate. This crystal was flash-cooled in liquid nitrogen prior to data collection.

Two complete X-ray diffraction data sets were collected from the same crystal at 100 K on the beamline X29A at NSLS using wavelengths of 1.075 Å (the resolution cutoff was 1.35 Å) and 0.979 Å (the resolution cutoff was 1.5 Å), respectively. All data were processed and scaled with HKL2000 (40). The crystal structure was determined by selenium-SAD applying anomalous X-ray diffraction data (collected at a wavelength of 0.979 Å) for phasing and SHELXD software (the CCP4 program package suite (46, 47). The structure was further refined using a higher resolution "native" data (resolution 1.35 Å) and the program REFMAC (48). The model was rebuilt and fixed manually using COOT visualization and refinement software (49). The data collection and refinement statistics are listed in Table 1. The coordinates and structure factors were deposited to the Protein Data Bank (PDB ID: 3V7P).

Structure Determination of the S145A mutant of Nis0429

The purified S145A protein was concentrated to about 38 mg/mL and crystallized as described above for the wild-type Nis0429 using a PHOENIX crystallization platform and four different MCSG crystallization screens. Protein crystals appeared in two to three weeks.

The crystals were directly flash frozen in liquid nitrogen. The 1.08 Å resolution X-ray diffraction data were collected on the beamline X29A (NSLS) at a wavelength of 1.075 Å and a temperature of 100 K. The data were processed and scaled with HKL2000 (40). The crystal of the S145S mutant of Nis0429 was essentially isomorphous to that of the wild-type protein (PDB ID: 3V7P), and thus the coordinates of the latter were used to build and subsequently refine the structure of the mutant using the program REFMAC (48) and a COOT visualization and refinement software (49). The final refinement statistics altogether with data collection parameters are listed in Table 1. The coordinates and structure factors were deposited in the Protein Data Bank (PDB ID: 4M51).

RESULTS

Protein Isolation

The enzymes studied in this investigation belong to groups 11 and 13 of cog1816 and groups 7 and 41 of cog0402 (Figures 3 and 4). Sav2595, Acel0264, Dr0824, and Nis0429 were expressed in *E. coli* and purified to homogeneity. Despite supplementing the growth medium with 1.0 mM $ZnCl_2$, the metal content of all proteins was low. Sav2595 contained an average of 0.1 equivalent of Zn^{2+} per subunit. Acel0264 contained an average of 0.4 equivalent of Zn^{2+} and 0.1 equivalent of Ni^{2+} per monomer. Both proteins were incubated overnight with $ZnCl_2$, but no increase in catalytic activity was observed after incubation. Dr0824 contained 0.8 equivalents of Mn^{2+}. Nis0429 contained 0.5 equivalents of Zn^{2+} and 0.1 equivalent each of Fe^{2+} and Mn^{2+}.

Substrate Specificity

The substrate profiles of Sav2595, Acel0264, Dr0824, and Nis0429 were initially determined by monitoring changes in absorbance between 240–300 nm using a small library of modified adenosine derivatives. The best substrate for these four enzymes is AFL. Less efficient substrates are NECA, adenosine, 5′-methylthioadenosine, *S*-adenosylhomocysteine, 2′-deoxyadenosine, 3′-deoxyadenosine, 5′-deoxyadenosine and AMP. The kinetic constants are presented in Table 2. The values of k_{cat}/K_m with AFL for Sav2595, Acel0264, Dr0824 and Nis0429 are greater than 10^5 M^{-1} s^{-1}.

Mutational Analysis

Single-site mutants of Sav2595 and Nis0429 were constructed to help identify residues that may interact with the side chain carboxylate of AFL. The R87A and R87M variants of Sav2595 were expressed in *E. coli* and subsequently purified. The S145A mutant of Nis0429 was constructed, expressed in *E. coli,* and purified. The mutant proteins were assayed, and the kinetic constants for the deamination of AFL and adenosine were determined. The results for the Sav2595 mutants are presented in Table 3. The Nis0429 S145A mutant has a k_{cat} of 4.5 s^{-1}, a K_M of 12 μM and a k_{cat}/K_M of 3.8×10^5 M^{-1} s^{-1}.

Structure of Nis0429

The crystal structure of Nis0429 (PDB ID: 3V7P) was determined at 1.35 Å resolution, revealing a slightly distorted TIM-barrel domain with a metal binding site and a much smaller β-barrel domain positioned between two short α-helices. The latter is comprised of both N-terminal (residues 1–54) and C-terminal (starting around residue 354) sequences. The likely biological assembly of Nis0429 is a C2-symmetrical homodimer (Figure 5), in which both domains are involved in formation of the dimer interface and maintenance of the active site architecture. Besides protein atoms and water molecules, the refined structure contains a few tightly bound ligands (listed in Table 1). The metal is coordinated to His-61, His-62, His-206 and water 608, which was modeled and refined as Fe^{2+} on the basis of

coordination geometry and crystallographic refinement (Figure 6). However, the possibility exists that there is a mixture of metals in the active site, although refinement parameters and coordination geometry are more consistent with the presence of iron and thus the metal was modeled and refined as such. The difference between the metal analysis of the purified protein and the metal in the active site may be related to the crystallization conditions.

In addition to two small unknown ligands (refined as formic acid and glycine) inside the active site, strong electron density has revealed the presence of a larger molecule that was modeled and refined as benzoic acid (Figure 6). Based on the location of this particular ligand and using the coordinates of previously reported structures of adenosine deaminases, one AFL molecule was manually modeled into the active site of Nis0429. In the crystal structure, the carboxylic acid moiety of benzoic acid is coordinated by the side chain hydroxyl of Ser-145 from β-strand 3 and a water molecule. The same interactions can be mimicked by the 5′-side chain of AFL (Figure 7). The analogous residue position in all members of group 7 from cog0402 is either serine or threonine. Comparison of Nis0429 (PDB ID: 3V7P) with structures of adenosine deaminases revealed striking similarities between their highly conserved (>90%) metal and nucleoside binding sites (magenta, Figure 8). By contrast, the hydrophobic portion of the binding site is represented by mainly non-polar residues specific for Nis0429 and its closest homologues (yellow, Figure 8).

Structure of S145 Mutant of Nis0429

The crystals of wild-type Nis0429 and the S145A mutant were essentially isomorphous (Table 1) and their structures are very similar with one important difference: the orientation of the benzoic acid ligand bound inside the active site next to a residue 145 (Figure 9). In the wild-type protein the carboxyl moiety of the ligand is hydrogen-bonded to the Ser-145 side-chain hydroxyl group and a Gly-144 amide nitrogen, whereas in the mutant structure the same group is hydrogen-bonded to the side chain from Arg-89. Another interesting feature of the Ser-145 mutant is the presence of the metal ion (refined as Fe^{2+}) in the active site. The high resolution structure revealed a slightly distorted tetrahedral metal ion coordination with varied distances between bonded atoms. The distances are as follows: 1.63 Å (Fe - water A962); 1.96 Å (Fe - His-61); 2.12 Å (Fe - His-63) and 2.33 Å (Fe – His-206). The side chain of Asp-306 is positioned in two alternate conformations, one of which may support a weak coordination bond to the iron at a distance of 2.52 Å.

Structure of Dr0824

The crystal structure of Dr0824 was determined by the SAD method and refined to 1.78 Å resolution (Table 1). Dr0824 consists of two domains, a distorted TIM-barrel domain and a small domain with a β-sheet flanked by a-helices on both sides. The small domain is made up of both the N-terminal (34 through 91) and C-terminal (399 through 413) residues. Residues 1 – 33, 47 – 49, and 395 – 396 have poor or undefined electron density. Missing residues 47 – 49 and 395 – 396 are located at the interface of the intra-molecular domains.

A metal ion is located at the C-terminal side of the β-sheet of the distorted TIM-barrel domain, which is the catalytic domain. The metal ion is coordinated with His-97 and His-99 of β-strand 1, His-238 of β-strand 6, and a water molecule (O554). Water molecule O554 also forms a hydrogen bond with OD1 of Asp-352 and is expected to be the nucleophilic center. Both N- and C-terminal sides of the distorted β-barrel of this catalytic domain are covered by α-helices and the active site is closed from outside interactions.

The crystallographic data support the conclusion that the biological assembly for this protein is a monomer rather than a homodimer. Excluding the unique N-terminal sequence fragment (~50 residues), the Dr0824 structure shows a high degree of similarity to Nis0429 (Figure

10) with RMSD between Cα atoms along the aligned length (residues 54–274) of about 2.1 Å. The only large difference between these two structures is the conformation of the two segments near the active site pocket (residues 102–128 and 253–283, colored blue in the 2IMR structure in Figure 10. The observed differences may be the result of crystal packing, loops flexibility, or the presence of ligands bound in the substrate-binding site of Nis0429.

Docking Poses of AFL in Sav2595 and Acel0264

In docking screens, the chiral tetrahedral intermediate for the deamination reaction that was generated by the attack of hydroxide on the *re*-face of the adenine moiety of AFL was ranked 51 and 30 among the 57,680 high-energy intermediates (HEI) in the virtual substrate library for Acel0264 and Sav2595, respectively. In the modeled active site of Sav2595 from cog1816, the metal ion is coordinated by three histidine residues (His-32, His-34, and His-215) and one aspartate residue (Asp-296). This catalytic machinery is conserved in Dr0824 and Nis0429 from cog0402. The modeled active site of Sav2595 shows the two hydroxyl groups from the ribose moiety of AFL forming hydrogen bonds with the side chain carboxylate group from Asp-154. This interaction is also identified in the docking of AFL to Nis0429 (Glu-141). The N3 nitrogen from the adenine moiety of AFL forms a hydrogen bond with the backbone amine group from Gly-188, but this interaction is replaced by a histidine residue in Nis0429 (His-179). The AFL specificity of Sav2595 can be explained by the presence of an arginine residue (Arg-87) that interacts with the carboxylate group in the 3-ketobenzoic acid moiety of AFL (Figure 11). In Nis0429, Ser-145 is predicted to hydrogen bond to the 3-ketobenzoic acid moiety (Figure 8).

DISCUSSION

Utilization of Bioinformatics, Structural Biology, and Molecular Enzymology to Determine Substrate Specificity

A group of bacterial deaminases from cog1816 and cog0402 within the amidohydrolase superfamily of enzymes has been characterized. These enzymes are currently incorrectly annotated as adenosine deaminases and chlorohydrolases. The most likely physiological substrate is 6-aminodeoxyfutalosine (AFL). AFL is an intermediate in the biosynthesis of menaquinone in certain bacteria that utilize the futalosine pathway (4). Further evidence supporting a role in the futalosine pathway is the location of Acel0264 in a cluster of genes related to the biosynthesis of menaquinone (Figure 2). The gene for Acel0264 is flanked by the genes for Acel0261 and Acel0263. These two genes are orthologs to *MqnA* and *MqnC*, respectively. Other members of group 11 from cog1816 have only *UbiE*, a methyltransferase in menaquinone biosynthesis, and *cog-FixC*, a geranylgeranyl reductase, as recurring neighbors. Members of group 13 of cog1816 and groups 7 and 41 of cog0402 show no genomic hints related to menaquinone biosynthesis, except for the lack of *men* gene homologs and the presence of *mqn* homologs.

Sequence Comparison and Structural Analysis

The four enzymes in group 11 of cog1816 share sequence identity to one another of greater than 50%, but are less than 30% identical to the well-characterized adenosine deaminase from *E. coli* (locus tag: b1623) from group 5 of cog1816. The same degree of similarity is observed between members of group 13. All of the residues that coordinate the divalent metal ion in the active site are conserved among the group 5, 11 and 13 enzymes, which include the HxH motif following β-strand 1, the histidine at the C-terminus of β-strand 5 and the aspartate at the C-terminus of β-strand 8. Another mutually conserved aspartate following β-strand 8 is responsible for hydrogen bonding to N7 of the adenine ring. Also conserved among these deaminases from cog1816 is the glutamate from the HxxE motif at the C-terminus of β-strand 5. A functionally important and conserved histidine at the C-

terminus of β-strand 6 forms a hydrogen bond with the putative hydrolytic water and may participate in proton transfer reactions.

Unlike other members of cog1816, members of group 11 and 13 have a conserved arginine following β-strand 1. Using adenosine deaminase (PDB ID: 1KRM) as a structural template, a model was constructed in an attempt to understand the substrate specificity of Acel0264 and Sav2595. An arginine residue was identified that could possibly recognize the carboxylate group of the 5′-side chain of AFL. The predicted location of Arg-87 in Sav2595 is on a turn between two α-helices after β-strand 1. This arginine was mutated to alanine or methionine in Sav2595, resulting in slower rates for the deamination of AFL, with essentially no effect on K_M (Table 3). The R87A and R87M mutants could be saturated with adenosine, whereas the wild-type enzyme was not. From these results, we propose that Arg-87 is likely near the substrate binding site (Figure 11) because changing this residue leads to the loss of substrate specificity. The remainder of the large substrate binding pocket is made up of hydrophobic residues and an aspartate or glutamate to interact with the 2′-hydroxyl group.

Members of group 7 of cog0402 share a sequence identity of 40% or more within the group, but much less than 30% with atrazine chlorohydrolase, and less than 25% compared to guanine deaminases from group 2. Dr0824 has a sequence identity of 25% or less with members of group 7 of cog0402, and has very low sequence similarity to well-characterized members of cog0402 and groups 11 and 13 of cog1816. Active site residues involved in metal binding are conserved among group 7 and 41 members, including an HxH motif after β-strand 1, another histidine after β-strand 5 and an aspartate after β-strand 8. A histidine after β-strand 6 is associated with binding a hydrolytic water molecule. A catalytically essential glutamate is conserved after β-strand 5. Additional shared substrate recognition elements include a glutamate after β-strand 1 involved in binding of the 3′-hydroxyl group of the ribose moiety and a histidine after β-strand 4 responsible for hydrogen bonding to N3 of the purine ring system of a nucleoside substrate.

The active site structure of Nis0429 (PDB ID: 3V7P) is shown in Figure 6 and the complete structure is shown in Figure 7 with AFL manually docked in the active site. The placement of AFL in the active site of Nis0429 was guided by the position of benzoic acid in the structure of this protein and the binding of adenosine in the active site of adenosine deaminases. This binding mode predicts an interaction with Ser-145 after β-strand 3 to bind the side chain of AFL (Figure 8). A serine or threonine residue is conserved in this position in all members of group 7 of cog0402, but not in group 41 or groups 11 and 13 of cog1816. It is noteworthy that the S145A mutation of Nis0429 has resulted in more than a 10-fold higher K_M value than that of the wild-type enzyme. These data support the idea that Ser-145 is directly involved in binding of AFL.

We have demonstrated that Acel0264, Sav2595, Nis0429 and Dr0824 are not adenosine deaminases or chlorohydrolases, as currently annotated. Based on sequence similarity and conservation of putative essential residues, all known members of groups 11 and 13 of cog1816 and groups 7 and 41 of cog0402 are probable 6-aminodeoxyfutalosine deaminases. It is quite likely that these enzymes evolved from distinct deaminases, which may explain the variation in observed enzyme architecture. On the basis of our data, we propose a list of proteins believed to be 6-aminodeoxyfutalosine deaminase enzymes (Table 4). The presence of an AFL deaminase in certain organisms is surprising to some degree because *Campylobacter jejuni* has been reported to have a methylthioadenosine nucleosidase (MTAN) that will hydrolyze AFL directly to the MqnB product (Figure 1) but will not hydrolyze futalosine (25). *C. jejuni* MTAN is the only close relative for futalosine hydrolase (MqnB) activity in the genome. It is unclear why organisms would have an AFL deaminase,

if they do not also have a futalosine hydrolase. Perhaps the AFL deaminase gene is a remnant of the original futalosine pathway in these organisms.

Comparison to Adenosine Deaminase from cog1816

The prototypical adeonosine deaminase from *E. coli* (locus tag: b1623) is also a member of the amidohydrolase superfamily and cog1816. This enzyme has a mononuclear metal center and will deaminate adenosine, N-6-methyladenosine, and 2′-deoxyadenosine as substrates with comparable rate constants. With the high conservation of metal-binding and catalytically essential residues between Acel0264, Sav2595, and b1623, it is very likely that these enzymes will employ the same mechanism for deamination. Dr0824 and Nis0429 have a slightly modified nucleoside binding site compared to b1623, but the structures and the biochemical data clearly support the notion that they function through an activated water molecule to facilitate the deamination reaction.

Acidothermus celluloyticus 11B has a protein from cog0402 (Acel1042) of unknown function, and no other proteins in cog1816 or cog1001, and thus adenosine or adenine deaminases are not likely to be found in this organism. The genome of *Streptomyces avermitilis* MA-4680 contains five proteins from cog1816 and four proteins from cog0402. Of the five proteins from cog1816, Sav5577 is a confirmed adenine deaminase belonging to group 3, Sav1165 and Sav3358 are believed to be adenosine deaminases from group 1, and Sav4906 from group 17 has unknown function. The genes from cog0402 include an isoxanthopterin deaminase (Sav2017) and Sav4899, Sav6652, and Sav3494 of unknown function. *Deinococcus radiodurans* R1 has an adenine deaminase from cog1001 (DrA0270) and a guanine deaminase from cog0402 (DrA0180) in addition to Dr0824. Nis0429 is the sole protein in *Nitratiruptor* sp. SB155-2 from cog0402, cog1816 or cog1001.

Strategy for Annotation and Characterization

Acidothermus celluloyticus 11B, *Streptomyces avermitilis* MA-4680, *Nitratiruptor* sp. SB155-2, and *Deinococcus radiodurans* R1 were identified as lacking the *men* genes necessary for menaquinone biosynthesis. The genes involved in an alternative menaquinone biosynthetic pathway were identified in clusters in the *Acidothermus celluloyticus* 11B genome and we have confirmed that Acel0264 was capable of deaminating 6-aminodeoxyfutalosine. The genomes of *Nitratiruptor* sp. SB155-2, *Streptomyces avermitilis* MA-4680, and *Deinococcus radiodurans* R1 lacked the *men* genes associated with the shikimate pathway, but they do possess the *mqn* genes of the futalosine pathway. A sequence similarity network was constructed and used to identify additional proteins with similar sequences to Acel0264, Sav2595, Dr0824, and Nis0429. Substrate profiles for these proteins were determined using a small focused library and a model for substrate binding was proposed. The high resolution crystal structure of Nis0429 with smaller ligands in the active site helped us to predict the binding mode of AFL as a substrate.

Acknowledgments

FUNDING

This work was supported in part by the Robert A. Welch Foundation (A-840) and the NIH (GM 71790). This work was made possible by the Center for Synchrotron Biosciences grant, P30-EB-009998, from the National Institute of Biomedical Imaging and Bioengineering (NIBIB). Use of the National Synchrotron Light Source, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-98CH10886.

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
NECA	1-(6-amino-9H-purin-9-yl)-1-deoxy-N-ethyl-β-D-ribofuranuronamide
SAH	S-adenosylhomocysteine
AFL	6-aminodeoxyfutalosine

References

1. Hiratsuka T, Furihata K, Ishikawa J, Yamashita H, Itoh N, Seto H, Dairi T. An alternative menaquinone biosynthetic pathway operating in microorganisms. Science. 2008; 321:1670–1673. [PubMed: 18801996]
2. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, Harris DE, Quail MA, Kieser H, Harper D, Bateman A, Brown S, Chandra G, Chen CW, Collins M, Cronin A, Fraser A, Goble A, Hidalgo J, Hornsby T, Howarth S, Huang CH, Kieser T, Larke L, Murphy L, Oliver K, O'Neil S, Rabbinowitsch E, Rajandream MA, Rutherford K, Rutter S, Seeger K, Saunders D, Sharp S, Squares R, Squares S, Taylor K, Warren T, Wietzorrek A, Woodward J, Barrell BG, Parkhill J, Hopwood DA. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). Nature. 2002; 417:141–147. [PubMed: 12000953]
3. Borodina I, Krabben P, Nielsen J. Genome-scale analysis of Streptomyces coelicolor A3(2) metabolism. Genome research. 2005; 15:820–829. [PubMed: 15930493]
4. Arakawa C, Kuratsu M, Furihata K, Hiratsuka T, Itoh N, Seto H, Dairi T. Diversity of the early step of the futalosine pathway. Antimicrobial agents and chemotherapy. 2011; 55:913–916. [PubMed: 21098241]
5. Holm L, Sander C. An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. Proteins. 1997; 28:72–82. [PubMed: 9144792]
6. Seibert CM, Raushel FM. Structural and catalytic diversity within the amidohydrolase superfamily. Biochemistry. 2005; 44:6383–6391. [PubMed: 15850372]
7. Kamat SS, Fan H, Sauder JM, Burley SK, Shoichet BK, Sali A, Raushel FM. Enzymatic deamination of the epigenetic base N-6-methyladenine. Journal of the American Chemical Society. 2011; 133:2080–2083. [PubMed: 21275375]
8. Hall RS, Agarwal R, Hitchcock D, Sauder JM, Burley SK, Swaminathan S, Raushel FM. Discovery and structure determination of the orphan enzyme isoxanthopterin deaminase. Biochemistry. 2010; 49:4374–4382. [PubMed: 20415463]
9. Goble AM, Zhang Z, Sauder JM, Burley SK, Swaminathan S, Raushel FM. Pa0148 from Pseudomonas aeruginosa catalyzes the deamination of adenine. Biochemistry. 2011; 50:6589–6597. [PubMed: 21710971]
10. Goble AM, Fan H, Sali A, Raushel FM. Discovery of a cytokinin deaminase. ACS chemical biology. 2011; 6:1036–1040. [PubMed: 21823622]
11. Wang Z, Quiocho FA. Complexes of adenosine deaminase with two potent inhibitors: X-ray structures in four independent molecules at pH of maximum activity. Biochemistry. 1998; 37:8314–8324. [PubMed: 9622483]

12. Ho MC, Cassera MB, Madrid DC, Ting LM, Tyler PC, Kim K, Almo SC, Schramm VL. Structural and metabolic specificity of methylthiocoformycin for malarial adenosine deaminases. Biochemistry. 2009; 48:9618–9626. [PubMed: 19728741]
13. Baer HP, Drummond GI, Gillis J. Studies on the specificity and mechanism of action of adenosine deaminase. Arch Biochem Biophys. 1968; 123:172–178. [PubMed: 5689048]
14. Kinoshita T, Nakanishi I, Terasaka T, Kuno M, Seki N, Warizaya M, Matsumura H, Inoue T, Takano K, Adachi H, Mori Y, Fujii T. Structural basis of compound recognition by adenosine deaminase. Biochemistry. 2005; 44:10562–10569. [PubMed: 16060665]
15. Kinoshita T, Tada T, Nakanishi I. Conformational change of adenosine deaminase during ligand-exchange in a crystal. Biochem Biophys Res Commun. 2008; 373:53–57. [PubMed: 18549808]
16. Koch AL, Vallee G. The properties of adenosine deaminase and adenosine nucleoside phosphorylase in extracts of Escherichia coli. J Biol Chem. 1959; 234:1213–1218. [PubMed: 13654350]
17. Lin J, Westler WM, Cleland WW, Markley JL, Frey PA. 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. 1998; 95:14664–14668. [PubMed: 9843946]
18. Gleeson MP, Burton NA, Hillier IH. Prediction of the potency of inhibitors of adenosine deaminase by QM/MM calculations. Chem Commun (Camb). 2003:2180–2181. [PubMed: 13678190]
19. Sadat Hayatshahi SH, Abdolmaleki P, Ghiasi M, Safarian S. QSARs and activity predicting models for competitive inhibitors of adenosine deaminase. FEBS Lett. 2007; 581:506–514. [PubMed: 17250831]
20. Wilson DK, Rudolph FB, Quiocho FA. Atomic structure of adenosine deaminase complexed with a transition-state analog: understanding catalysis and immunodeficiency mutations. Science. 1991; 252:1278–1284. [PubMed: 1925539]
21. Larson ET, Deng W, Krumm BE, Napuli A, Mueller N, Van Voorhis WC, Buckner FS, Fan E, Lauricella A, DeTitta G, Luft J, Zucker F, Hol WG, Verlinde CL, Merritt EA. 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. 2008; 381:975–988. [PubMed: 18602399]
22. Weiss PM, Cook PF, Hermes JD, Cleland WW. Evidence from nitrogen-15 and solvent deuterium isotope effects on the chemical mechanism of adenosine deaminase. Biochemistry. 1987; 26:7378–7384. [PubMed: 3427079]
23. Sideraki V, Mohamedali KA, Wilson DK, Chang Z, Kellems RE, Quiocho FA, Rudolph FB. Probing the functional role of two conserved active site aspartates in mouse adenosine deaminase. Biochemistry. 1996; 35:7862–7872. [PubMed: 8672487]
24. Atkinson HJ, Morris JH, Ferrin TE, Babbitt PC. Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. PLoS One. 2009; 4:e4345. [PubMed: 19190775]
25. Li X, Apel D, Gaynor EC, Tanner ME. 5′-methylthioadenosine nucleosidase is implicated in playing a key role in a modified futalosine pathway for menaquinone biosynthesis in Campylobacter jejuni. J Biol Chem. 2011; 286:19392–19398. [PubMed: 21489995]
26. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic acids research. 1990; 18:6069–6074. [PubMed: 2235490]
27. Hall RS, Xiang DF, Xu C, Raushel FM. N-Acetyl-D-glucosamine-6-phosphate deacetylase: substrate activation via a single divalent metal ion. Biochemistry. 2007; 46:7942–7952. [PubMed: 17567047]
28. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009; 25:1189–1191. [PubMed: 19151095]
29. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32:1792–1797. [PubMed: 15034147]

30. Sali A, Blundell TL. Comparative Protein Modeling by Satisfaction of Spatial Restraints. J Mol Biol. 1993; 234:779–815. [PubMed: 8254673]
31. Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein Science. 2006; 15:2507–2524. [PubMed: 17075131]
32. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R. Novel procedure for modeling ligand/receptor induced fit effects. J Med Chem. 2006; 49:534–553. [PubMed: 16420040]
33. Lorber DM, Shoichet BK. Hierarchical docking of databases of multiple ligand conformations. Current Topics in Medicinal Chemistry. 2005; 5:739–749. [PubMed: 16101414]
34. Mysinger MM, Shoichet BK. Rapid Context-Dependent Ligand Desolvation in Molecular Docking. J Chem Inf Model. 2010; 50:1561–1573. [PubMed: 20735049]
35. Hermann JC, Ghanem E, Li YC, Raushel FM, Irwin JJ, Shoichet BK. Predicting substrates by docking high-energy intermediates to enzyme structures. J Am Chem Soc. 2006; 128:15882–15891. [PubMed: 17147401]
36. Hermann JC, Marti-Arbona R, Fedorov AA, Fedorov E, Almo SC, Shoichet BK, Raushel FM. Structure-based activity prediction for an enzyme of unknown function. Nature. 2007; 448:775–779. [PubMed: 17603473]
37. Fan H, Hitchcock D, Seidel RD, Hillerich B, Lin H, Almo SC, Sali A, Shoichet BK, Raushel FM. Assignment of pterin deaminase activity to an enzyme of unknown function guided by homology modeling and docking. J Am Chem Soc. 2013; 135:795–803. [PubMed: 23256477]
38. Kanehisa M, Goto S. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 2000; 28:27–30. [PubMed: 10592173]
39. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M. From genomics to chemical genomics: new developments in KEGG. Nucleic Acids Res. 2006; 34:D354–D357. [PubMed: 16381885]
40. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
41. De-La-Fortelle E, Bricogne G. Maximum-likelihood heavy atom parameter refinement in the MIR and MAD methods. Methods Enzymol. 1997; 276:472–493.
42. Terwilliger TC, Berendzen J. Automated structure solution for MIR and MAD. Acta Cryst D. 1997; 55:849–861. [PubMed: 10089316]
43. Perrakis A, Morris R, Lamzin VS. Automated protein model building combined with iterative structure refinement. Nature Struct Biol. 1999; 6:458–463. [PubMed: 10331874]
44. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M. Improved methods in building protein models in electron density map and the location of errors in these models. Acta Crystallogr. 1991; A47:110–119.
45. Brunger AT, Adams PD, Clore GM, Delano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Somonsom T, Warren GL. Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr D. 1998; 54:905–921. [PubMed: 9757107]
46. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. Overview of the CCP4 suite and current developments. Acta crystallographica Section D, Biological crystallography. 2011; 67:235–242.
47. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. Acta Cryst. 1994; D50:760–763.
48. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica Section D, Biological crystallography. 1997; 53:240–255.
49. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta crystallographica Section D, Biological crystallography. 2004; 60:2126–2132.

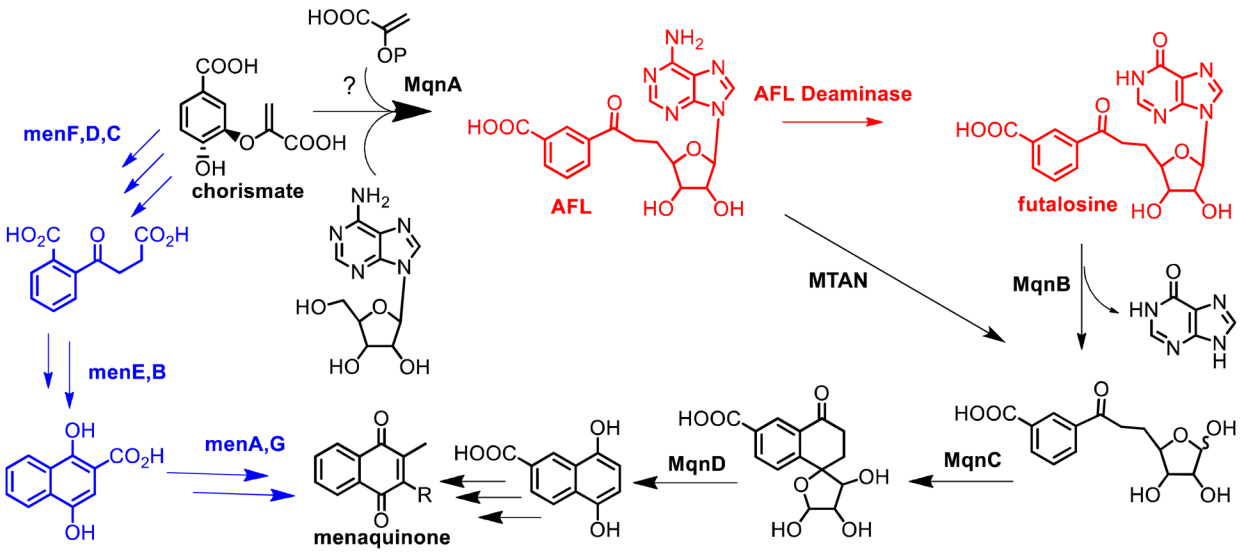

Figure 1.

Biosynthetic pathways for the assembly of menaquinone. The deamination of AFL is shown in red. The pathway for menaquinone biosynthesis in *E. coli* is shown in blue.

Figure 2.

Genomic context for Acel0264. *UbiD* (Acel0255) is believed to catalyze a decarboxylation reaction in the late stages of menaquinone biosynthesis. *UbiA* (Acel0256) functions as a 4-hydroxybenzoate polyprenyltransferase. *UbiX* (Acel0257) is a 3-octaprenyl-4-hydroxybenzoate carboxy-lyase. Acel0261 and Acel0263 are *MqnA* and *MqnC* homologs, respectively. Acel0264 catalyzes the deamination of 6-aminodeoxyfutalosine to futalosine and has been characterized in this work. *UbiE* (Acel0265) is a methyltransferase involved in menaquinone biosynthesis. *cog-FixC* (Acel0266) serves as a geranylgeranyl reductase in the biosynthesis of menaquinone. Genes depicted in grey have unknown function or function in an unrelated pathway.

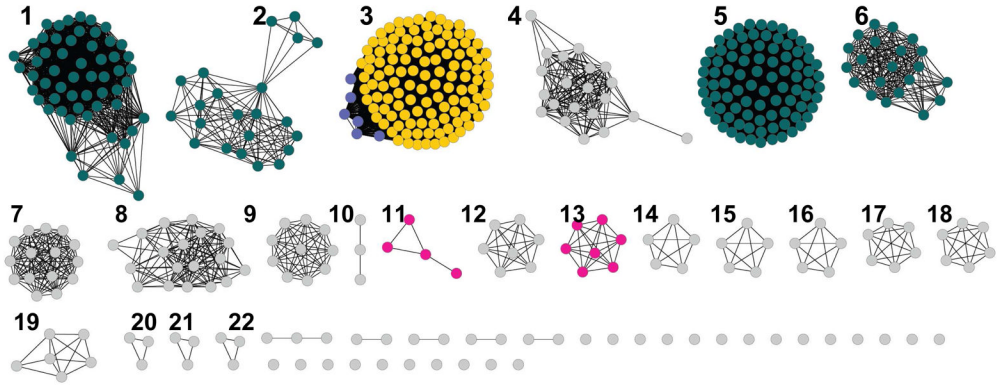

Figure 3.

Sequence similarity network created using Cytoscape (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*. Additional groups containing adenosine deaminase enzymes are shown in green. Group 3 contains adenine deaminase enzymes, with the exception of the blue outliers, which are cytokinin deaminases. In groups 11 and 13, the sequences depicted in pink contain all of the essential residues for the deamination of AFL identified in this investigation.

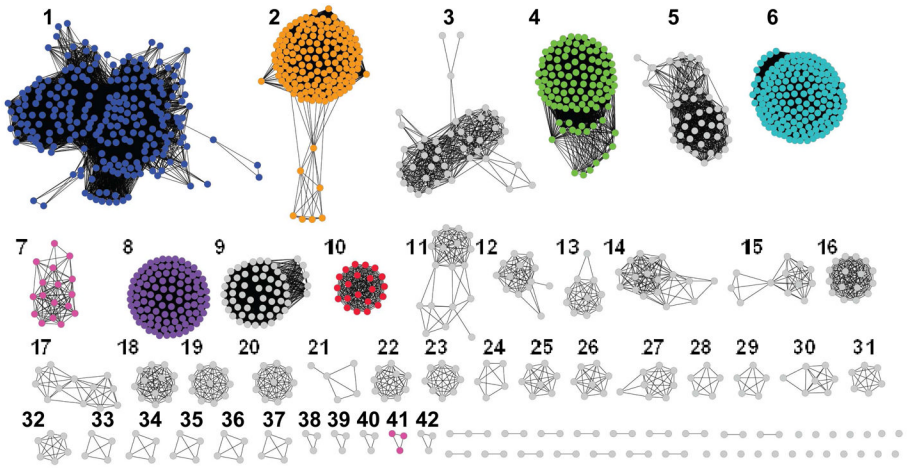

Figure 4.

Sequence similarity network of cog0402. Groups 7 and 41 (pink) contain genes with all the essential residues to be AFL deaminases. Other groups with known function include group 1 (5′-modified adenosine deaminases), group 2 (guanine deaminases), group 4 (isoxanthopterin deaminases or 8-oxoguanine deaminases), group 6 (cytosine deaminases), group 8 (N-formimino-L-glutamate deiminases) and group 10 (pterin deaminase). Groups with no known function are shown in grey.

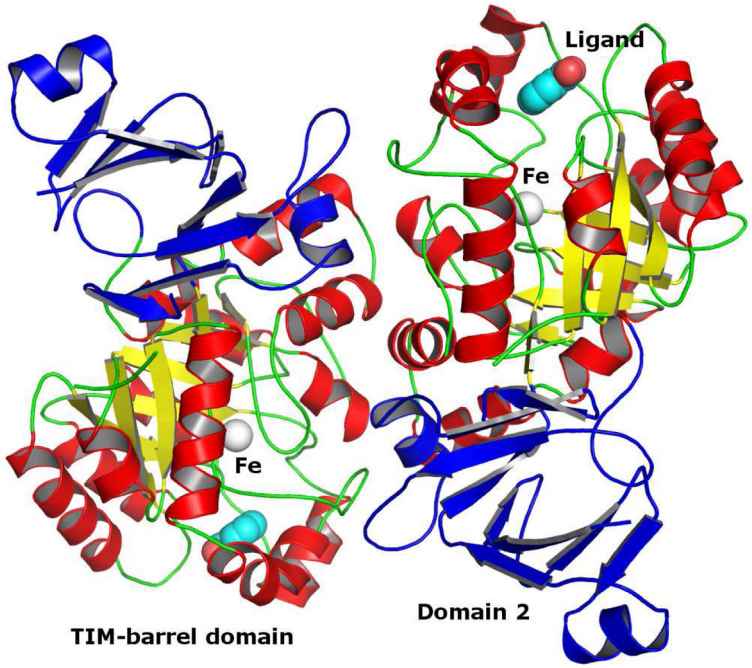

Figure 5.
Cartoon model of Nis0429 C2-symmetrical homo-dimer (based on the structure 3V7P). Each monomer consists of two domains, a TIM-barrel domain (helices are red, beta-sheets are yellow and loops are green) and a smaller domain 2 (a beta-barrel flanked with two short alpha-helices, all are blue colored). Iron and benzoic acid ligands are shown as spheres. Both domains contribute to interactions between amino acid residues along the dimer interface.

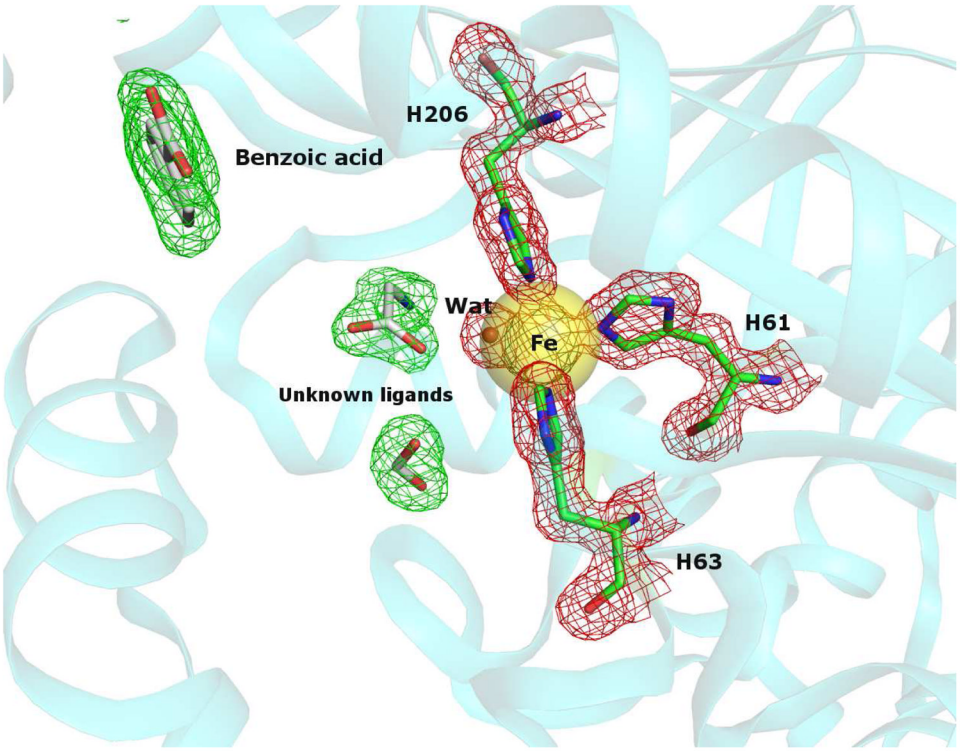

Figure 6.

The $2F_oF_c$ electron density map (1σ cutoff) is shown around amino acid residues involved in metal coordination (red) inside of the active site of Nis0429 (PDB ID: 3V7P) and around small unknown ligands (green). The residues are shown as stick models, and iron is a yellow sphere. The unknown ligands were modeled and refined as benzoic acid, glycine, and formate. They occupy the suspected substrate-binding site of Nis0429. The overall protein structure in the background is represented by a cartoon model.

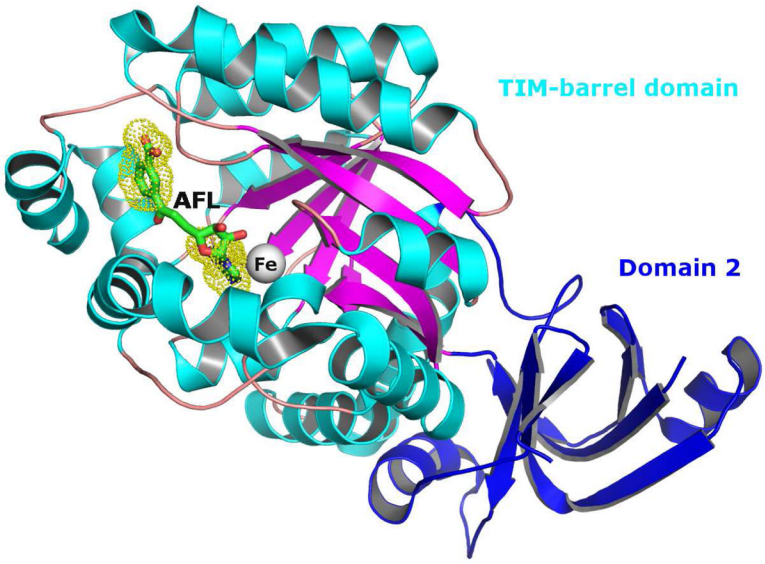

Figure 7.

The cartoon model of Nis0429 with a bound substrate (6-aminodeoxyfutalosine, **AFL,** the stick model). The substrate was modeled using the location of a few smaller ligands detected inside of the substrate-binding site in the structure 3V7P (the surface of these small ligands is marked by yellow dots). The iron atom is drawn as a sphere.

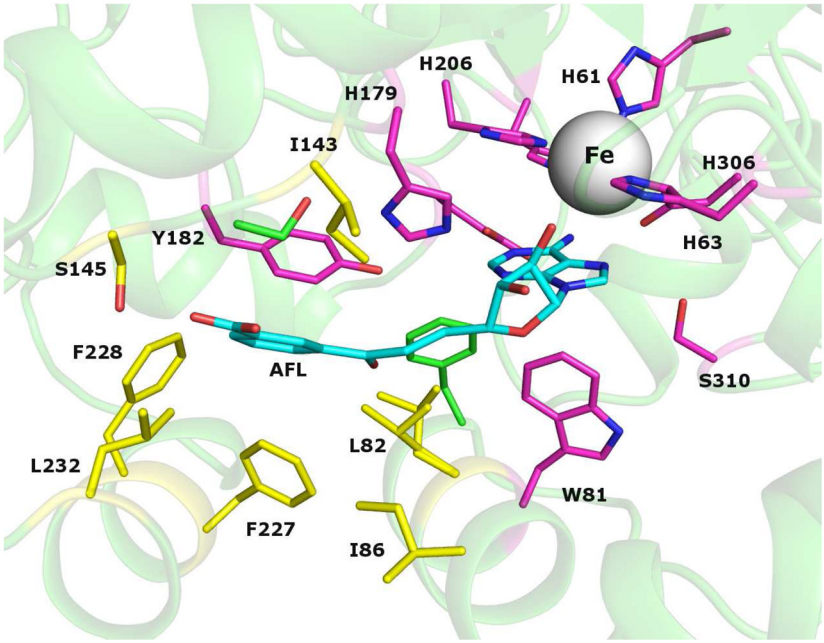

Figure 8.

The amino acid residues in the active site of Nis0429 (PDB ID: 3V7P) are shown with a modeled AFL presented as sticks. The residues that are conserved among “typical” adenosine deaminases have carbon atoms colored in magenta. The residues of the proposed hydrophobic binding site are unique to Nis0429 and its homologues and have carbon atoms drawn in yellow. The carbon atoms of AFL are colored in cyan. The iron atom is shown as a grey sphere. The protein in the background is presented by a cartoon model.

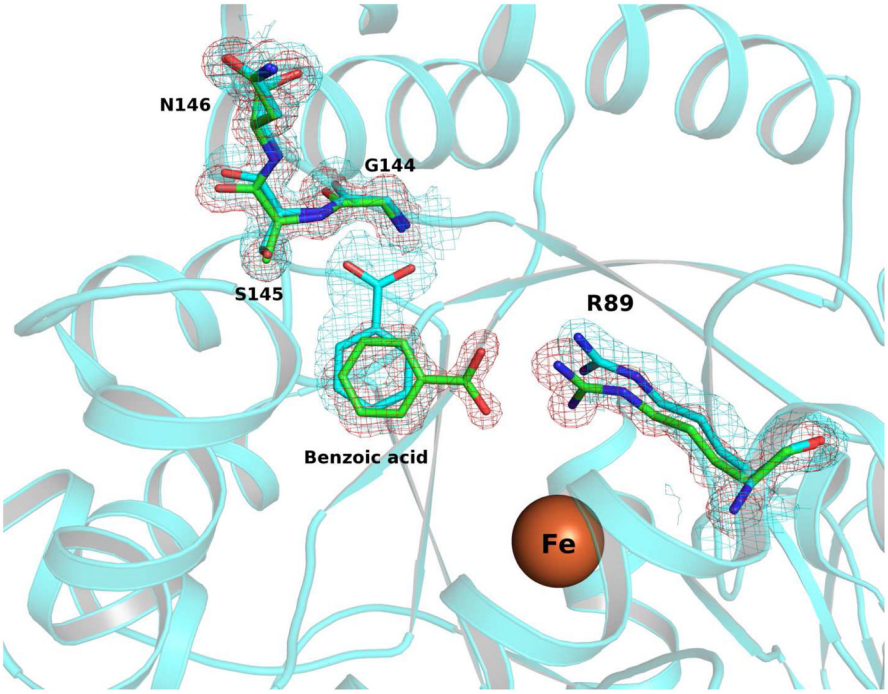

Figure 9.
Superimposition of wild type Nis0429 (PDB ID: 3V7P) and the S145A mutant (PDB ID: 4M51). For the wild type enzyme the carbon atoms are cyan, oxygen atoms are red and nitrogen atoms are blue. For the S145A mutant the carbon atoms are green with selected residues drawn as sticks. Iron in drawn as a brown sphere. The 2FoFc electron density maps (sigma cutoff 1.0) are shown around the benzoic acid ligands and interacting residues.

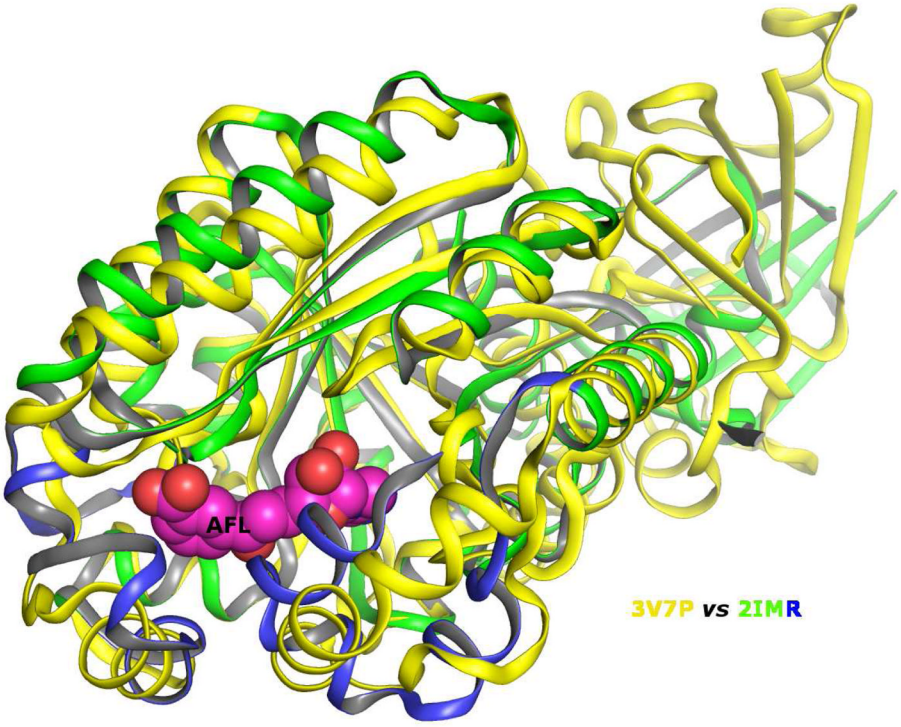

Figure 10.

Superposition of two structures, 3V7P (yellow) and 2IMR (green and blue), is presented as a cartoon model. Two sequence segments (residues 102–128 and 253–283) in the structure 2IMR are blue. Except for the small domain 2, these are the only parts in the Dr0824 structure that are significantly different form their counterparts in the Nis0429 structure. AFL (spheres) cannot be fitted inside of the active site of the structure 2IMR due to multiple clashes with residues residing on segments colored in blue.

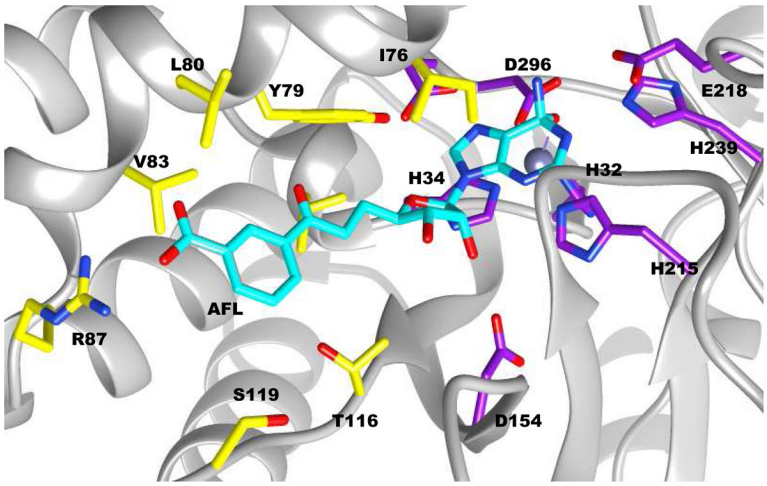

Figure 11.

Model of Sav2595 with AFL docked in the active site. The HEI of AFL is shown in blue. The hydrophobic pocket that accommodates the AFL side chain and substrate recognizing arginine are shown in yellow. Residues in purple are conserved among adenosine deaminases. The zinc ion is shown as a grey sphere.

Table 1

Data Collection and Refinement Statistics (data for high resolution shells)

PDB identifier	**2IMR**	**3V7P**	**4M51**
Space group	C 1 2 1	C 1 2 1	C 1 2 1
Unit cell dimension (Å)			
-a	113.83	91.515	92.163
-b	71.30	75.169	75.788
-c	51.34	75.646	76.291
Cell angles (degrees)			
-alpha	90.00	90.00	90.00
-beta	115.20	120.72	120.99
-gamma	90.00	90.00	90.00
Molecules per ASU	1	1	1
Solvent content	40.24	47.01	47.03
Matthew's Coefficient	2.06	2.32	2.33
Ligands	Zn^{2+}	benzoate tartrate unknown Fe^{2+}	benzoate sulfate HEPES Fe^{2+}
X-ray source	NSLS X12C	NSLS X29A	NSLS X29A
Wavelength	0.9795	1.075	1.075
Method of structure solution	SAD	SAD	molecular replacement
Resolution	29.30–1.78 (1.84–1.78)	70.00–1.35 (1.38–1.35)	33.91–1.08 (1.10–1.08)
Resolution/refinement	29.30–1.78	50.00–1.35	33.91–1.08
Completeness (%)	95.3 (64.7)	99.6 (94.3)	95.1 (91.1)
I/sigma (I)	19.20 (3.30)	9.00 (1.30)	10.20 (2.50)
R_{sym}	0.048 (0.364)	0.051 (0.730)	0.061 (0.640)
R_{work} (R_{free})	20.1 (23.0)	12.8 (16.1)	11.8 (13.5)
R_{free} reflections (%)	1321 (5.0%)	2889 (3.0%)	5505 (3.0%)
Average B factor	24.11	24.77	16.48
RMSD			
-bonds lengths	0.005	0.012	0.012
-bond angles	1.310	1.389	1.614
Number of solvent molecules	170	226	609
Ramachandran plot statistics			
Most favored regions	97.1%	96.9%	96.4%
Allowed regions	2.6%	2.9%	3.4%
Outliers	0.27% (His301)	0.25% (His257)	0.25%(His257)

Table 2

Kinetic Constants for Sav2595, Acel0264, Dr0824, and Nis0429[a]

	Sav2595			Acel0264			Dr0824			Nis0429		
	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
AFL	3.8	4.8	7.9×10^5	6.8	6.0	1.1×10^6	8.6	1.8	4.8×10^6	1.2	0.9	1.3×10^6
NECA	-	-	34	0.007	61	110	-	-	1.0×10^4	-	-	110
5′-methylthioadenosine	0.40	94	4.3×10^3	-	-	28	-	-	3.4×10^4	0.31	110	2.8×10^3
adenosine	-	-	2.2×10^4	-	-	20	-	-	1.8×10^4	-	-	11
S-adenosylhomocysteine	-	-	310	-	-	24	-	-	7.0×10^3	-	-	79
5′-deoxyadenosine	-	-	2.9×10^3	-	-	47	-	-	7.4×10^4	0.017	42	410
3′-deoxyadenosine	-	-	240	-	-	44	-	-	240	-	-	6
2′-deoxyadenosine	-	-	1.1×10^3	-	-	51	0.72	37	1.9×10^4	-	-	5
AMP	-	-	32	-	-	45	-	-	21	-	-	3

[a]Standard errors on the kinetic constants are less than 15%. Assays were conducted at pH 7.5 and at 30 °C. Sav2595 is from group 13 of cog1816; Acel0264 is from group 11 of cog1816; Dr0824 is from group 41 of cog0402; and Nis0429 is from group 7 of cog0402.

Table 3

Kinetic Constants for Sav2595 mutants[a]

	AFL			adenosine		
	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
WT	3.8	4.8	8×10^5	-	-	2.2×10^4
R87A	0.094	3.4	2.7×10^4	0.106	8.3	1.3×10^4
R87M	0.42	8	5.3×10^4	0.41	10	4.1×10^4

[a]Standard errors on the kinetic constants are less than 15%. Assays were conducted at pH 7.5 and at 30 °C.

Table 4

6-Aminodeoxyfutalosine deaminase enzymes

Organism	Locus Tag	GI	cog
Acidothermus cellulolyticus 11B	Acel_0264	117927473	1816
Arcobacter butzleri RM018	Abu_0489	157736750	0402
Campylobacter concisus 13826	CCC13826_0371	157164865	0402
Campylobacter curvus 525.92	CCV52592_0362	154174508	0402
Campylobacter fetus subsp. fetus 82–40	CFF8240_1504	118475030	0402
Campylobacter hominis ATCC BAA_381	CHAB381_1212	154147820	0402
Campylobacter jejuni NCTC 11168	Cj0067	15791459	0402
Campylobacter jejuni RM1221	CJE0064	57237077	0402
Campylobacter jejuni subsp. 81–176	CJJ81176_0105	121612374	0402
Campylobacter jejuni subsp. doylei 269.97	JJD26997_0077	153952082	0402
Campylobacter jejuni subsp. jejuni 81116	C8J_0060	157414381	0402
Campylobacter lari RM1200	Cla_0229	222823270	0402
Deinococcus deserti VCD115	Deide_06659	226355533	0402
Deinococcus geothermalis DSM 11300	Dgeo_0576	94984683	0402
Deinococcus radiodurans R1	Dr0824	15805850	0402
Frankia alni ANC14a	FRAAL1027	111220494	1816
Frankia sp. Ccl3	Francci3_0534	86739247	1816
Frankia sp. EAN1pec	Franean1_6098	158317840	1816
Frankia sp. EAN1pec	Franean1_2196	158314030	1816
Helicobacter hepaticus ATCC 51449	HH1764	32267263	0402
Nautilia profundicola AmH	NAMH_1355	224373376	0402
Nitratiruptor sp. SB155-2	NIS_0429	152990178	0402
Nocardioides sp. JS614	Noca_1355	119715591	1816
Salinospora arenicola CNS-205	Sare_1766	159037391	1816
Salinospora tropica CNB-440	Strop_1779	145594323	1816
Streptomyces avermitilis MA-4680	Sav_2595	29829137	1816
Streptomyces coelicolor A3(2)	Sco5662	21224013	1816
Streptomyces griseus subsp. griseus NBRC 13359	SGR_1841	182435634	1816
Sulfurimonas denitrificans DSM 1251	Suden_0686	78776885	0402
Sulfurovum sp. NBC37-1	SUN_1037	152992630	0402
Wolinella succinogenes DSM 1740	WS1477	34557817	0402