4374 *Biochemistry* **2010,** *49,* 4374–4382
DOI: 10.1021/bi100252s

BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Discovery and Structure Determination of the Orphan Enzyme Isoxanthopterin Deaminase[†,‡]

Richard S. Hall,[§] Rakhi Agarwal,[⊥] Daniel Hitchcock,[∥] 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 77842-3012,* [∥]*Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas 77843,* [⊥]*Biology Department, Brookhaven National Laboratory, Upton, New York 11973, and* [@]*Lilly Biotechnology Center, 10300 Campus Point Drive, San Diego, California 92121*

Received February 18, 2010; Revised Manuscript Received April 22, 2010

ABSTRACT: Two previously uncharacterized proteins have been identified that efficiently catalyze the deamination of isoxanthopterin and pterin 6-carboxylate. The genes encoding these two enzymes, NYSGXRC-9339a (gi|44585104) and NYSGXRC-9236b (gi|44611670), were first identified from DNA isolated from the Sargasso Sea as part of the Global Ocean Sampling Project. The genes were synthesized, and the proteins were subsequently expressed and purified. The X-ray structure of Sgx9339a was determined at 2.7 Å resolution (Protein Data Bank entry 2PAJ). This protein folds as a distorted $(\beta/\alpha)_8$ barrel and contains a single zinc ion in the active site. These enzymes are members of the amidohydrolase superfamily and belong to cog0402 within the clusters of orthologous groups (COG). Enzymes in cog0402 have previously been shown to catalyze the deamination of guanine, cytosine, *S*-adenosylhomocysteine, and 8-oxoguanine. A small compound library of pteridines, purines, and pyrimidines was used to probe catalytic activity. The only substrates identified in this search were isoxanthopterin and pterin 6-carboxylate. The kinetic constants for the deamination of isoxanthopterin with Sgx9339a were determined to be 1.0 s^{-1}, 8.0 μM, and 1.3×10^5 M^{-1} s^{-1} (k_{cat}, K_m, and k_{cat}/K_m, respectively). The active site of Sgx9339a most closely resembles the active site for 8-oxoguanine deaminase (Protein Data Bank entry 2UZ9). A model for substrate recognition of isoxanthopterin by Sgx9339a was proposed on the basis of the binding of guanine and xanthine in the active site of guanine deaminase. Residues critical for substrate binding appear to be conserved glutamine and tyrosine residues that form hydrogen bonds with the carbonyl oxygen at C4, a conserved threonine residue that forms hydrogen bonds with N5, and another conserved threonine residue that forms hydrogen bonds with the carbonyl group at C7. These conserved active site residues were used to identify 24 other genes which are predicted to deaminate isoxanthopterin.

High-throughput DNA sequencing efforts have substantially increased the number of fully sequenced bacterial genomes and improved our understanding of the enzymes contained within these organisms. A critical assessment of the corresponding protein sequences reveals that a significant fraction of these enzymes have not been functionally annotated or are incorrectly annotated as catalyzing the wrong reactions. Misannotated functional assignments are likely due to an overreliance on automated homology-based sequence comparisons that are not substantiated by active site analysis and experimental verification. The ever-increasing number of putative enzymes for which the function(s) is unknown further suggests that many metabolic and catabolic pathways remain to be discovered. Our efforts to develop comprehensive strategies for annotating enzymes of unknown function have focused on the amidohydrolase superfamily (*1*−*10*).

The amidohydrolase superfamily (AHS)[1] was first recognized by Sander and Holm on the basis of the structural similarities of urease, phosphotriesterase, and adenosine deaminase (*11*). Enzymes within this superfamily catalyze hydrolytic reactions of amide and ester bonds contained within carbohydrates, peptides, and nucleic acids (*12*). Less common amidohydrolase-catalyzed reactions include decarboxylations, hydrations, and isomerizations (*3*, *13*−*17*). Active sites of enzymes within the AHS contain a binuclear or mononuclear metal center that activates a hydrolytic water molecule and/or the substrate for nucleophilic attack. More than 12000 unique protein sequences have been identified as members of the AHS within the genomes of the first 1000 completely sequenced bacteria deposited in the NCBI. Contained within the AHS are 24 clusters of orthologous groups (COGs).

One of the largest of these clusters with more than 1000 members is the group of proteins within cog0402. A network representation of the sequence relationships for cog0402 at a BLAST *E* value cutoff of 10^{-70} is presented in Figure 1 (*18*). All of the functionally annotated enzymes within this COG catalyze deamination reactions, including guanine deaminase (subcluster **2**),

[†]This work was supported in part by the National Institutes of Health (Grants GM 71790 and GM 74945).

[‡]The X-ray coordinates and structure factors for Sgx9339a have been deposited in the Protein Data Bank (entry 2PAJ).

*To whom correspondence should be addressed. 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.

[1]Abbreviations: AHS, amidohydrolase superfamily; COG, cluster of orthologous groups; NYSGXRC, New York SGX Research Center for Structural Genomics; PDB, Protein Data Bank; rmsd, root-mean-square deviation; VAST, vector alignment search tool.

 Published on Web 04/23/2010

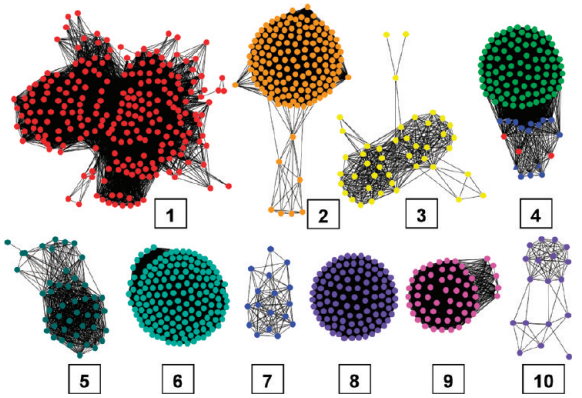


FIGURE 1: Cytoscape network representation (www.cytoscape.org) of the sequence relationships in cog0402 from the amidohydrolase superfamily. Each node in the network represents a single sequence, and each edge (depicted as lines) represents the pairwise connection between two sequences with the most significant BLAST *E* values (better than 1×10^{-70}). Lengths of edges are not meaningful except that sequences in tightly clustered groups are relatively more similar to each other than sequences with few connections. In subcluster **4**, the green-colored nodes catalyze the deamination of 8-oxoguanine and possess a conserved Cys-Ser dyad C-terminal to β-strand 4. In subcluster **4**, the blue-colored nodes possess a Thr-Thr dyad and catalyze the deamination of isoxanthopterin. In subcluster **4**, the red-colored nodes have a Thr-Ser dyad. Additional information is presented in the text.

cytosine deaminase (subcluster **6**), *S*-adenosylhomocysteine/thiomethyl adenosine deaminase (subcluster **1**), and *N*-formimino-L-glutamate deiminase (subcluster **8**). In addition, we have recently demonstrated that the majority of enzymes found within subcluster **4** catalyze the deamination of 8-oxoguanine to urate (*19*). The catalytic functions for the other subclusters within cog0402 remain uncharacterized.

The structural determinants of 8-oxoguanine deaminase in subcluster **4** are a conserved glutamine and tyrosine which occur in the loops at the C-termini of β-strands 1 and 2, respectively, and a conserved Cys-Ser dyad found at the C-terminus of β-strand 4 within a $(\beta/\alpha)_8$ barrel structural fold. In the subcluster of enzymes that catalyze the deamination of guanine, the Cys-Ser dyad is replaced with an Arg-Phe pair. Curiously, at the periphery of subcluster **4** is a group of sequences (colored blue in Figure 1) with a Thr-Thr dyad. This paper describes the discovery of function for this group of enzymes.

Two enzymes interrogated for this investigation were prepared for crystallization as part of the NIH-sponsored Protein Structure Initiative (PSI) by the New York SGX Research Center for Structural Genomics (NYSGXRC). The two targets, Sgx9236b and Sgx9339a, were derived from environmental DNA sequences isolated from bacteria in the Sargasso Sea under the Global Ocean Sampling Project. The closest homologue to Sgx9339a (88% identical) is Bxe_A2016 from *Burkholderia xenovorans* LB400, and the closest homologue of Sgx9236b (95% identical) is Bphyt_7136 from *Burkholderia phytofirmans* PsJN. Both of these proteins are currently annotated by NCBI as ATZ/TRZ-like amidohydrolases. The three-dimensional structure of Sgx9339a was determined at 2.7 Å resolution. Using a combination of compound library screening and structural comparisons, we have determined that both of these enzymes catalyze the deamination of isoxanthopterin to 7-oxylumazine.

MATERIALS AND METHODS

Materials. All chemicals and buffers were obtained from Sigma-Aldrich unless specified otherwise. 8-Oxoguanine was purchased from Toronto Research Chemicals Inc. Kinetic assays were performed in 96-well microtiter plates with a SPECTRAmax 384 Plus spectrophotometer from Molecular Devices. Compounds tested for enzymatic activity with Sgx9339a and Sgx9236b included isoxanthopterin, pterin 6-carboxylate, D-*erythro*-neopterin, 6-biopterin, 6,7-dimethylpterin, pterin, 8-oxoguanine, guanine, acycloguanosine, isocytosine, cytosine, folate, 2,4-diamino-6-hydroxymethylpteridine, GTP, GDP, GMP, guanosine, 3′,5′-cGMP, 2′-dGTP, 7-methylguanosine, 7-methylguanine, 2′-deoxyguanosine, melamine, ammeline, 5-azacytosine, isoguanine, creatinine, thiamine, cytidine, 2-deoxycytidine, 5-fluorocytosine, 5-methylcytosine, 3-oxauracil, CTP, CDP, CMP, cytosine arabinoside, cytidine-5′-diphosphocholine, dCTP, dCTP, dCMP, 3′,5′-cCMP, 3′-CMP, 2′,3′-cCMP, 5-methyl-2′-deoxycytidine, adenine, 7-methyladenine, 2′-deoxyadenosine, ATP, ADP, AMP, 3′,5′-cAMP, and thiomethyl adenosine.

Protein Expression and Purification. The sequences of Sgx9339a (gi|44585104) and Sgx9236b (gi|44611670) were obtained from the Global Ocean Sampling Project. The target selection processes for structure determination are described by Pieper et al. (*6*). Codon-optimized synthetic DNA (Codon Devices, Inc.) was used as a template and cloned into a custom TOPO-isomerase vector, pSGX3(BC), supplied by Invitrogen. Forward and reverse primers for Sgx9339a were ACCACGTATGATACCCAGCC and CTACCACAACTTCGCGCAAGAGC and for Sgx9236b were CAGACCGAAGCGCTGTTAATTA and CATCCGCTAACAACTGTTTAACCC, respectively. The clone encodes Met-Ser-Leu followed by the PCR product and then Glu-Gly-His_6. Miniprep DNA was transformed into BL21(DE3)-Codon+RIL expression cells (Stratagene), expressed, and made into a 30% glycerol stock for large-scale fermentation.

Expression clones of Sgx9339a and Sgx9236b were cultured using ZYP-5052 autoinduction media (W. Studier, Brookhaven National Laboratory). Overnight cultures were prepared from glycerol stocks that were derived from fresh transformants in LB medium. The following morning, overnight cultures were used to inoculate ZYP-5052 autoinduction medium. A 5−10% inoculum was achieved via addition of 12−25 mL of overnight culture to 250 mL of ZYP-5052 medium containing 90 μg/mL kanamycin and 30 μg/mL chloramphenicol in a 2 L baffled shake flask. Cultures were grown at 37 °C for 5 h with 225 rpm agitation. The temperature was then switched to 22 °C, and the cells were grown overnight for 18 h to allow for induction of the intended protein. Selenomethionine-labeled protein was produced for Sgx9339a and Sgx9236b using High Yield SeMet medium (Orion Enterprises, Inc., Northbrook, IL). Cells were 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 (39800*g*). 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,

Table 1: Data Collection and Refinement Statistics

Data Collection	
wavelength (Å)	0.9792
resolution (Å)	50−2.7
space group	$P3_121$
unit cell dimensions	a = 88.9 Å, b = 88.9 Å, c = 161.9 Å, α = β = 90°, γ = 120°
diffraction protocol	single-wavelength anomalous dispersion
no. of molecules per asymmetric unit	1
overall redundancy (outermost shell)	19 (10.5)
$I/\sigma(I)$	11.1 (2.5)
outermost shell (Å)	2.8−2.7
overall R_{merge} (outermost shell)	0.07 (0.30)
overall completeness (%) (outermost shell)	94.4 (63)
no. of reflections	19968
Refinement	
resolution range (Å)	50−2.7
no. of reflections	19173
completeness (working + test) (%)	91.6
R_{factor}	0.23
R_{free}	0.28
no. of protein atoms	3194
no. of zinc ions	1
no. of water molecules	85
B value (Å^2)	
from Wilson plot	39.2
mean B value (overall)	58.4
rmsd for bonds (Å)	0.009
rmsd for angles (deg)	1.6
Ramachandran plot analysis	
most favored region (additionally allowed) (%)	82.5 (16.1)
generously allowed region (disallowed) (%)	1.1 (0.3)

10% glycerol, and 5 mM DTT]. Fractions containing the protein of interest were combined and concentrated to 10 mg/mL by centrifugation in an Amicon Ultra-15 10000 Da MWCO centrifugal filter unit. Electrospray mass spectrometry was used to obtain an accurate mass of the protein purified to confirm its identity. The expression plasmids (NYSGXRC clones 9339a2BCt7p1 and 9236b2BCt11p1) are available through the PSI Material Repository (http://psimr.asu.edu/), and the DNA sequences and experimental details are available in the Protein Expression Purification and Crystallization Database (PepcDB) as TargetID "NYSGXRC-9339a" and "NYSGXRC-9236b".

Crystallization, Data Collection, and Structure Determination. Diffraction quality crystals of the selenomethionine form of Sgx9339a were obtained in trigonal space group $P3_121$ at room temperature via vapor diffusion against a reservoir solution containing 0.1 M MES (pH 6.5) and 12% (w/v) PEG 20000 (1 μL of reservoir solution with 1 μL of protein at 20 mg/mL). Crystals were frozen by direct immersion in liquid nitrogen using the mother liquor with 15% (v/v) glycerol added. X-ray diffraction data were collected under standard cryogenic conditions at beamline X29 (National Synchrotron Light Source, Brookhaven National Laboratory) and were reduced, merged, and scaled using HKL2000 (*20*). The Matthews coefficient was calculated to be 3.45 Å^3/Da assuming one molecule per asymmetric unit. Data collection and refinement statistics are listed in Table 1. The presence of zinc was confirmed by absorption edge scanning at the NSLS beamline.

All 13 internal (excluding N-terminal) selenium atoms were located using ShelxD (*21*) and used for phase refinement with SHARP (*22*) and density modification with SOLOMON (*23*). ARP/wARP (*24*) automatically built ~50% of the protein atomic model, and the remainder of the structure was built manually with O (*25*) and subsequently refined with CNS (*26*) at 2.7 Å resolution.

Measurement of Enzymatic Activity. The assays for the detection of ammonia for the enzymatic deamination of pteridines, purines, and pyrimidines were conducted with a 96-well UV−vis microplate spectrophotometer at 30 °C. Compounds in this putative substrate library were tested at a concentration of 1.0 mM, except for the pteridine derivatives and 8-oxoguanine, which were assayed at 0.1 mM because of solubility limitations. The coupling system contained 7.4 mM α-ketoglutarate, 0.4 mM NADH, 6 units of glutamate dehydrogenase, and 100 mM HEPES (pH 8.5) in a final volume of 0.25 mL. For direct product assays, extinction coefficients for the deamination of isoxanthopterin ($\Delta\varepsilon_{350}$ = 7654 M^{-1} cm^{-1}) and pterin 6-carboxylate ($\Delta\varepsilon_{264}$ = 8080 M^{-1} cm^{-1}) were determined from changes in the UV−vis absorption spectra after the deamination reaction was complete using substrate concentrations from 0.004 to 0.116 mM. Mass spectrometric analysis of product formation was performed by the Laboratory for Biological Mass Spectrometry (LBMS) in the Department of Chemistry at Texas A&M University.

Data Analysis. Sequence alignments were created using lustalW at http://align.genome.jp/sit-bin/clustalw. Alignments were formatted using the Multiple Align Show program at http://www.bioinformatics.org/sms/. Steady-state kinetic data were analyzed using Softmax Pro version 4.7.1. For the determination of the kinetic parameters (k_{cat}, K_m, and k_{cat}/K_m), the data were fit to eq 1 using the nonlinear least-squares fitting program SigmaPlot 10

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

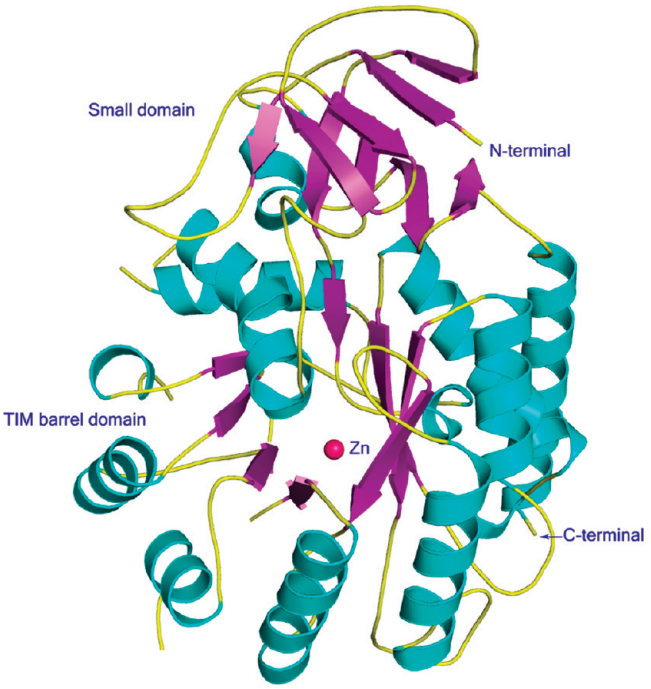


FIGURE 2: Ribbon representation of the overall structure of Sgx9339a (PDB entry 2PAJ). The α-helices, β-strands, and random coil loops are colored cyan, violet, and yellow, respectively. The zinc ion is colored pink. The small domain is made up of N-terminal residues 10−70 and residues 413−456. The ribbon diagram was made using PYMOL.

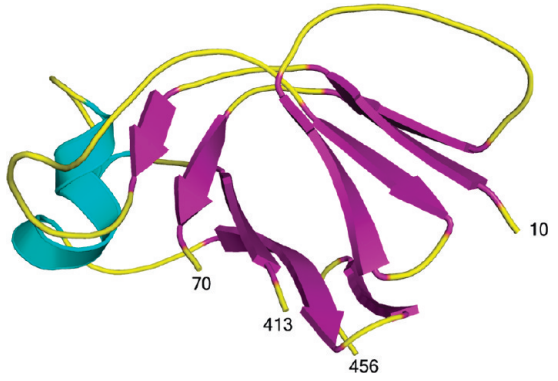


FIGURE 3: Ribbon representation of the small domain that extends from residue 10 to 70 and from residue 413 to 456.

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

RESULTS

Three-Dimensional Structure of Sgx9339a. The X-ray structure of Sgx9339a was determined to a resolution of 2.7 Å, and a ribbon representation of an individual subunit is presented in Figure 2. The protein folds into two distinct domains. The larger of the two domains consists of a distorted $(\alpha/\beta)_8$ TIM barrel, whereas the smaller domain is made up of N-terminal residues 10−70 and residues 413−456, together making a β-sheet sandwich as illustrated in Figure 3. Interestingly, this amidohydrolase has a higher molecular weight than many of the other structurally characterized members of this superfamily and has an extra helical tail that is preceded by a loop comprised of residues 457−463. The helix (residues 464−484) is approximated

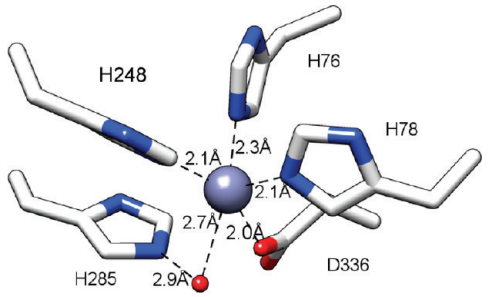
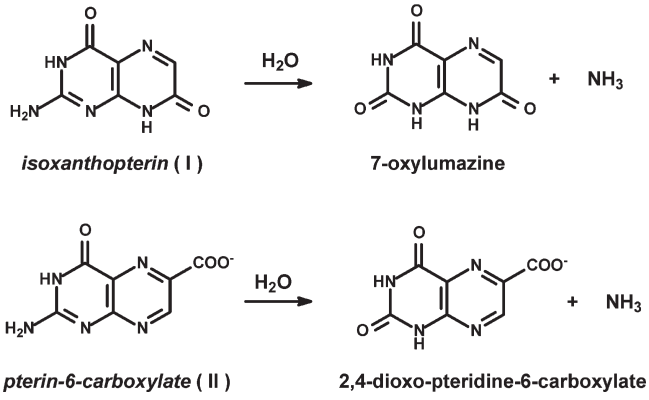


FIGURE 4: Active site interactions of Sgx9339a with the bound zinc and water molecule.

Scheme 1



to the large domain by salt bridge interactions between glutamate and arginine residues and hydophobic interactions at the interface. The opposite surface of the helix is positively charged, displaying five arginines and one lysine. The final refined model consists of residues 10−484, one zinc atom, and 85 water molecules. Loop flexibility may have caused poor or no electron density for the loops consisting of residues 89−105, 251−263, 315−319, and 362−380. These residues were not modeled in the final structure.

Active Site. The active center is located in the middle of the $(\beta/\alpha)_8$ barrel and is exposed to the solvent as illustrated in Figure 4. There is a single zinc ion in the active site. This metal ion is coordinated to two histidine residues (His-76 and His-78) at the C-terminus of β-strand 1 and another histidine at the C-terminus of β-strand 5 (His-248). The remaining protein ligand to the metal center is an aspartate that follows β-strand 8 (Asp-336). The water molecule responsible for nucleophilic attack of the substrate occupies the final coordination site. This water molecule is hydrogen bonded to the side chain of His-285 (2.9 Å) positioned near the C-terminus of β-strand 6. The geometry of the mononuclear metal center is identical to that previously observed with guanine deaminase (PDB entry 2UZ9), cytosine deaminase (*27*), and adenosine deaminase (*28*). However, the catalytically important glutamate (Glu-251) found within the HxxE motif at the C-terminus of β-strand 5 is not observed, presumably due to conformational flexibility within the crystal.

Substrate Specificity. Sgx9339a and Sgx9236b were assayed with a small compound library of pteridines, purines, and pyrimidines for deaminase activity. At a fixed enzyme concentration of ~100 nM, significant formation of ammonia could only be detected with isoxanthopterin (**I**) and pterin-6-carboxylate (**II**) as substrates. The reactions catalyzed by these enzymes are shown in Scheme 1.

The reaction products for the deamination of isoxanthopterin and pterin 6-carboxylate were verified by UV−vis spectroscopy and mass spectrometry. The absorbance spectrum of isoxanthopterin (**I**) is presented in Figure 5. At pH 8.5, there are absorbance maxima at 281 and 333 nm (*29*). The spectrum changes after the addition of Sgx9339a, and the absorbance maxima shift to 274 and 326 nm, which is consistent with the formation of 7-oxylumazine (*29*). The ESI negative ion mass spectrum for M − H of *m*/*z* 178.04 from isoxanthopterin changed to *m*/*z* 179.02 for 7-oxylumazine upon addition of enzyme. The maximum absorbance difference between isoxanthopterin and 7-oxylumazine was determined to be at 350 nm with a $\Delta\varepsilon$ of 7650 M^{-1} cm^{-1}. The kinetic constants for the deamination of isoxanthopterin with Sxg9339a and Sgx9236b were determined, and the results are listed in Table 2.

The absorbance spectrum for pterin 6-carboxylate at pH 8.5 is presented in Figure 5. In this spectrum, there are absorbance maxima at 264 and 361 nm. After the addition of Sgx9339a,

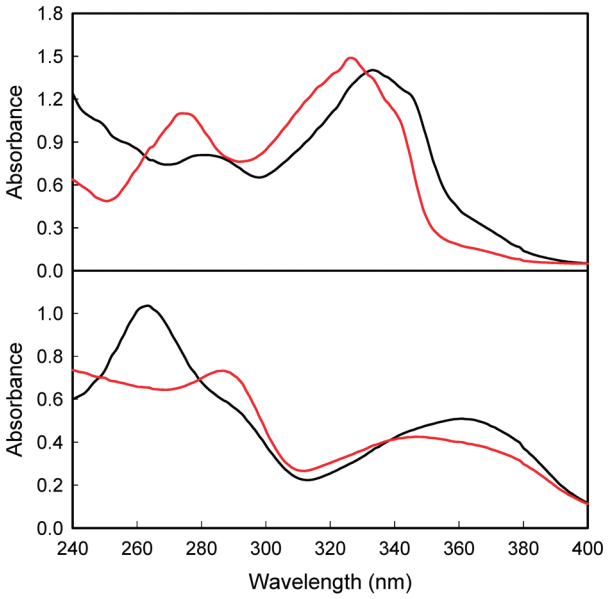


FIGURE 5: UV−vis absorption spectra of isoxanthopterin and pterin 6-carboxylate before and after the addition of Sgx9339a: (top) isoxanthopterin before (black) and after (red) the addition of Sgx9339a at pH 8.5 and (bottom) pterin 6-carboxylate before (black) and after (red) the addition of Sgx9339a. Additional information is provided in the text.

Table 2: Kinetic Parameters for Isoxanthopterin and Pterin 6-Carboxylate Deamination with Sgx9339a and Sgx9236b

enzyme	substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
Sgx9339a	isoxanthopterin	1.05 ± 0.03	0.008 ± 0.001	(1.3 ± 0.2) × 10^5
	pterin 6-carboxylate	1.01 ± 0.10	0.14 ± 0.02	(7.2 ± 0.2) × 10^3
Sgx9236b	isoxanthopterin	0.91 ± 0.06	0.05 ± 0.01	(1.8 ± 0.2) × 10^4
	pterin 6-carboxylate	0.28 ± 0.05	0.09 ± 0.03	(3.1 ± 0.4) × 10^3

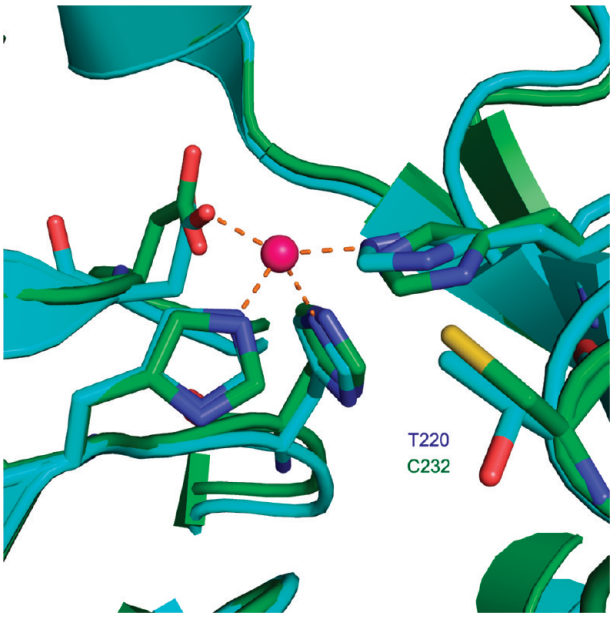


FIGURE 7: Active site overlay of Sgx9339a (PDB entry 2PAJ) shown with green carbons and guanine deaminase (PDB entry 2UZ9) shown with gray carbons.

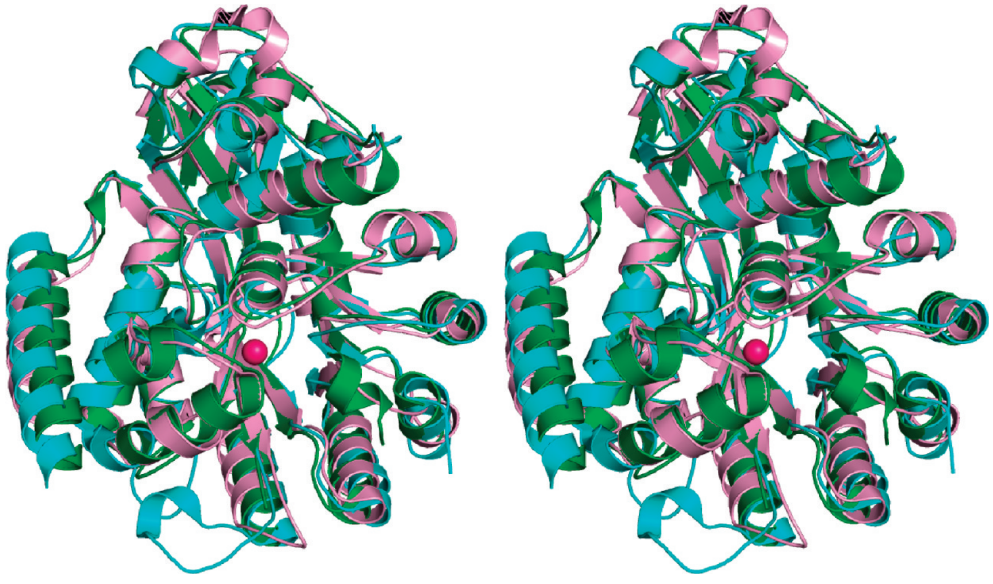


FIGURE 6: Stereoview of the structural alignment of Sgx9339a (cyan, PDB entry 2PAJ) with the structural neighbors 8-oxoguanine deaminase (green, PDB entry 3HPA) and *S*-adenosylhomocysteine deaminase (pink, PDB entry 1P1M) and zinc ions (pink).

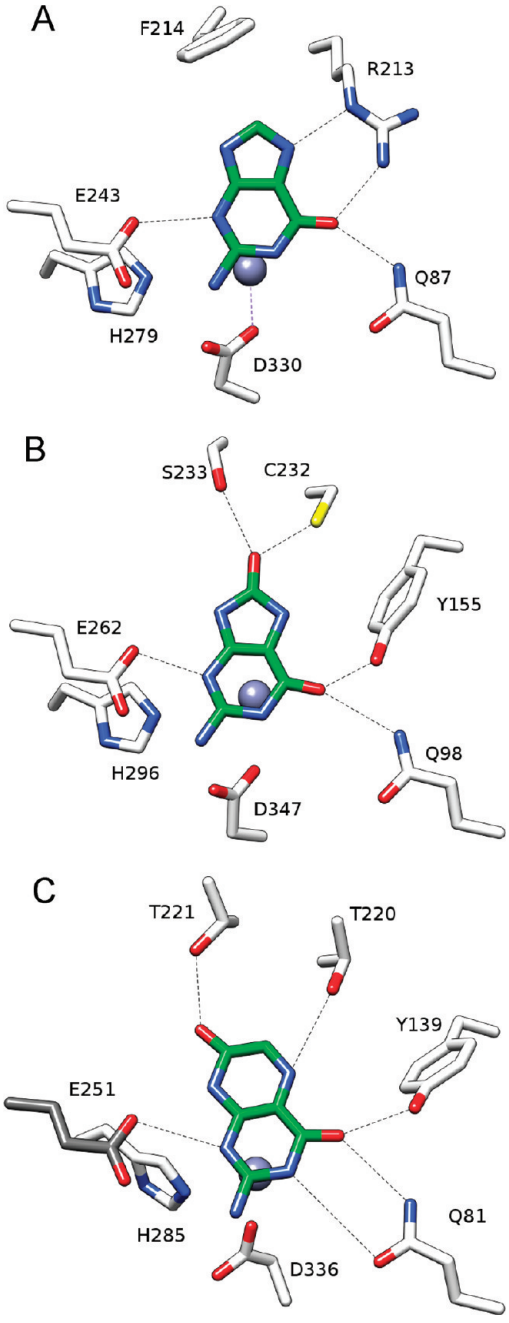
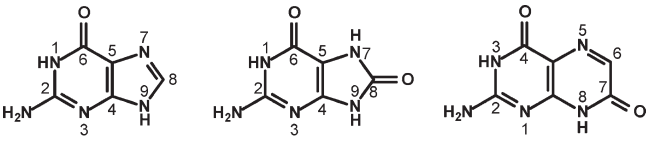


FIGURE 8: Models for substrate binding in the active sites of guanine deaminase, 8-oxoguanine deaminase, and isoxanthopterin. (A) Guanine was placed in the active site of human guanine deaminase (PDB entry 2UZ9) via a simple substitution of this compound for the product xanthine. (B) 8-Oxoguanine was positioned in the active site of Sgx9236e (PDB entry 3HPA) using a structural overlay of this protein with the active site of guanine deaminase (PDB entry 2UZ9) while overlapping the common atoms of 8-oxoguanine and xanthine. (C) Isoxanthopterin was positioned in the active site of Sgx9339a using a structural overlay of this protein with the active site of guanine deaminase while overlapping the common atoms of isoxanthopterin and xanthine. Since residues 251−263 are disordered in 2PAJ, the important glutamate (Glu-251) from the HxxE motif at the C-terminal end of β-strand 5 was inserted into the active site on the basis of the orientation of Glu-243 in guanine deaminase (PDB entry 2UZ9).

the spectrum changes to one with absorbance maxima at 287 and 347 nm. The ESI negative ion mass spectrum for M − H of m/z 206.03 from pterin 6-carboxylate changed to m/z 207.02 for 2,4-dioxopteridine 6-carboxylate after addition of enzyme. Direct kinetic assays were conducted at a wavelength of 264 nm with a

Scheme 2



$\Delta\varepsilon$ of 8080 M^{-1} cm^{-1}. The kinetic constants for the deamination of pterin 6-carboxylate with Sgx9339a and Sgx9236b are presented in Table 2. For each enzyme, the values of k_{cat}/K_m for the deamination of isoxanthopterin are approximately 1 order of magnitude greater than for pterin 6-carboxylate.

DISCUSSION

Structural Neighbors to Sgx9339a. The vector alignment search tool (VAST) from NCBI identified two close structural hits to Sgx9339a. The most similar protein (PDB entry 3HPA) catalyzes the deamination of 8-oxoguanine to urate (*19*), and the other protein (PDB entry 1P1M) catalyzes the deamination of *S*-adenosylhomocysteine (*5*). These two proteins superimpose with PDB entry 2PAJ with the Cα atoms of residues 389 and 374 as follows: *Z* scores of 49.6 and 43.5, rmsd of 1.8 and 2.5 Å, and 35 and 25% identical sequence, respectively. A stereoview of the structural alignments for these three proteins is presented in Figure 6. As expected, 2PAJ is structurally more similar to 3HPA in terms of loop size and arrangement, protein length, active site orientation, mononuclear metal center, and the presence of the C-terminal helix. The primary difference in the active site structures is the presence of a threonine (Thr-220) in 2PAJ as opposed to the cysteine (Cys-232) in 3HPA. The structural alignment comparing the active sites contained within 2PAJ and 3HPA is presented in Figure 7. The helix−turn−helix region of 3HPA (residues 105−123), which acts as a lid for the active site, is disordered in 2PAJ (residues 89−105).

It has been shown that Sgx9339a preferentially catalyzes the deamination of isoxanthopterin to 7-oxylumazine with a k_{cat}/K_m value that exceeds 10^5 M^{-1} s^{-1}. Although the structure of this enzyme has been determined, it has thus far not been possible to obtain cocrystals that contain the substrate or product bound in the active site. How then is the substrate accommodated within the active site? One of the closest structural homologues to Sgx9339a in which a product or substrate is bound in the active site is human guanine deaminase in a complex with the product xanthine (PDB entry 2UZ9). In that structure, the carbonyl group at C6 of xanthine forms a hydrogen bond with Gln-87 and Arg-213. Arg-213 also interacts with N7. In addition to these recognition elements, Asp-330 and His-279 function together with the metal center to activate the hydrolytic water molecule, and Glu-243 serves as the ultimate proton donor to N1 during the transition from substrate to product. Shown in Figure 8A is a model of how guanine binds in the active site of human guanine deaminase.

The closest structural homologue to Sgx9339a is Sgx9236e (PDB entry 3HPA). This protein catalyzes the deamination of 8-oxoguanine (*19*). A ligand-bound structure has not been determined, but a model for how the substrate for this enzyme binds in the active site has been developed on the basis of computational docking of potential high-energy intermediates in the active site and direct structural comparisons with guanine deaminase (*19*). A model for the interaction of 8-oxoguanine in

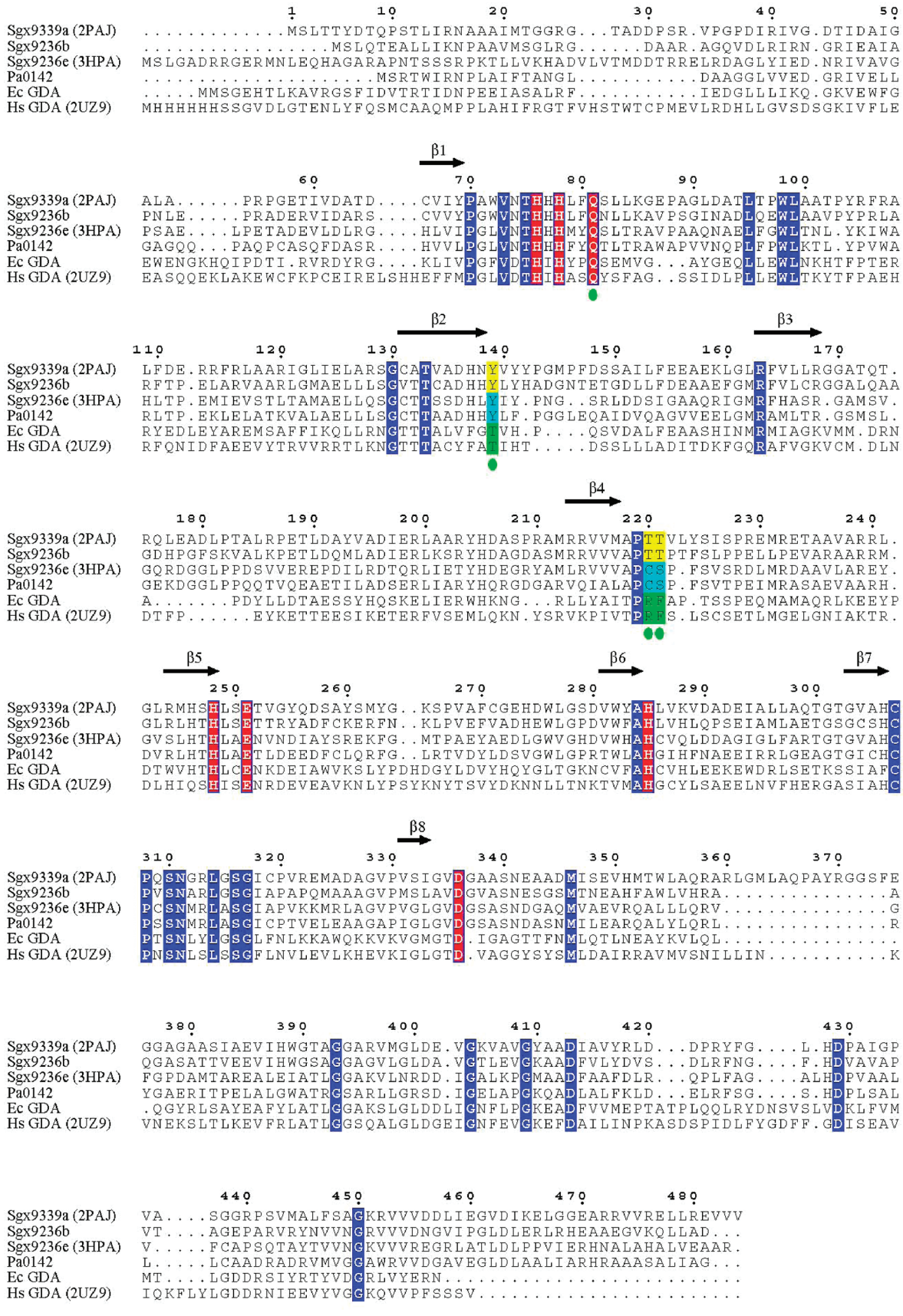


FIGURE 9: Multiple-sequence alignment of two isoxanthopterin deaminases (Sgx9339a and Sgx9236b), two 8-oxoguanine deaminases (Pa0142 and Sgx9236e), and two guanine deaminases (*E. coli* guanine deaminase and guanine deaminase from humans). The eight β-strands of the $(\beta/\alpha)_8$ barrel in Sgx9339a are denoted by the black arrows. The conserved metal binding residues and catalytic residues in all six proteins are highlighted in red. The conserved ligand binding residues in isoxanthopterin deaminase, 8-oxoguanine deaminase, and guanine deaminase are highlighted in yellow, light blue, and green, respectively. The remaining fully conserved residues in all three proteins are highlighted in dark blue.

the active site of Sgx9236e is presented in Figure 8B. In this model, a glutamine residue (Gln-98), homologous to Gln-87 in guanine deaminase, forms a hydrogen bond with the carbonyl group at C6, Tyr-155 also forms a hydrogen bond with the carbonyl group at C6, while Cys-232 and Ser-233 interact with the carbonyl group at C8. Asp-347, His-296, and Glu-262 function in concert to activate the hydrolytic water molecule and facilitate the various proton transfer reactions.

The structures of guanine deaminase (PDB entry 2UZ9) and 8-oxoguanine deaminase (PDB entry 3HPA) were used to rationalize the binding of isoxanthopterin in the active site of Sgx9339a (PDB entry 2PAJ). In this model, the conserved glutamine (Gln-81) and tyrosine (Tyr-139) form a hydrogen bond with the carbonyl group at C4. One of the two conserved threonine residues (Thr-220) forms a hydrogen bond with N5, and the other conserved threonine (Thr-221) forms a hydrogen bond with the carbonyl group at C7. The catalytic triad, consisting of the conserved aspartate (Asp-336), histidine (His-285), and glutamate (Glu-251), is poised to facilitate the deamination reaction in the same fashion as the other functionally characterized enzymes in cog0402. However, Glu-251 cannot be positioned precisely because of conformational flexibility. In this model, we have oriented this residue on the basis of the observed position of the homologous residue in guanine deaminase. The proposed interactions for the binding of isoxanthopterin in the active site of Sgx9339a are presented in Figure 8C. The numbering for the atoms contained within guanine, 8-oxoguanine, and isoxanthopterin is presented in Scheme 2.

Amino Acid Sequence Analysis. The amino acid sequences of the two isoxanthopterin deaminases, Sgx9339a and Sgx9236b, were aligned with two of the newly discovered enzymes that deaminate 8-oxoguanine (Pa0142 from *Pseudomonas aeruginosa* and Sgx9236e) and guanine deaminases from *Homo sapiens* and *Escherichia coli*.[2] The sequence alignment is shown in Figure 9. The eight β-strands from the $(\beta/\alpha)_8$ barrel structural fold in Sgx9339a are highlighted by the black arrows. The fully conserved metal binding and catalytic residues are highlighted in red. The metal binding residues include the two histidine residues from the C-terminus of β-strand 1, a histidine residue from the C-terminus of β-strand 5, and an aspartate from the C-terminus of β-strand 8. The two additional residues that are involved in the facilitation of the proton transfers and activation of the nucleophilic water molecule include a glutamate at the C-terminus of β-strand 5 (HxxE motif) and a histidine at the C-terminus of β-strand 6. The conserved glutamine residue in all six proteins that is positioned after the HxH motif at the C-terminal end of β-strand 1 is also highlighted in red. Conserved substrate binding residues found in all guanine deaminases are highlighted in green. These residues include the arginine and phenylalanine at the C-terminal end of β-strand 4 and the threonine at the C-terminal end of β-strand 2. For the two 8-oxoguanine deaminases (Pa0142 and Sgx9236e), these residues (in addition to a conserved tyrosine found at the C-terminal end of β-strand 2) are highlighted in light blue and the proposed substrate binding residues specific for the recognition of isoxanthopterin are highlighted in yellow. The remaining residues that are common to all three sets of enzymes are highlighted in dark blue.

[2]The residue numbering for Sgx9339a that appears in Figure 9 corresponds to the numbering system and amino acid sequence used in PDB entry 2PAJ.

Sequence comparisons with guanine deaminase and 8-oxoguanine deaminase were subsequently used to differentiate a small group of enzymes, which are now predicted to catalyze the deamination of isoxanthopterin. Guanine deaminases utilize an Arg-Phe dyad for the recognition of guanine, whereas in 8-oxoguanine deaminase, this is changed to a Cys-Ser dyad. In the isoxanthopterin deaminases, this changes to a Thr-Thr dyad. It is remarkable that the glutamine, which forms a hydrogen bond with the carbonyl group of the ring to be deaminated, occurs in all three enzymes and the tyrosine, which interacts with the carbonyl group or ring nitrogen of the adjacent ring, occurs in 8-oxoguanine deaminase and isoxanthopterin deaminase (see Figure 8).

A BLAST search of protein sequences in the NCBI database was used to identify the organisms containing genes anticipated to catalyze the deamination of isoxanthopterin. From this search, 23 organisms were found to share this gene (as evidenced by conservation of the Thr-Thr motif from the C-terminus of β-strand 4, the glutamine from the C-terminus of β-strand 1, and the tyrosine from the C-terminus of β-strand 2). A complete list of genes and organisms is provided in Table S1 of the Supporting Information. The novel isoxanthopterin deaminases share a sequence identity of greater than 41% and are predicted to participate in the deamination of isoxanthopterin within these organisms. Included with these 23 genes are four outliers (represented as red circles in subcluster 4 from Figure 1) which share a threonine-serine dyad at this position. These genes are Mmwyl1_4101 (gi|152998097) from *Marinomonas* sp. MWYL1, HMPREF0023_2011 (gi|226953181) from *Acinetobacter* sp. ATCC 27244, PputW619_2587 (gi|170721762) from *Pseudomonas putida* W619, and PP_3209 (gi|26989928) from *P. putida* KT2440. It appears that the Thr-Ser dyad may be an evolutionary intermediate between the Cys-Ser dyad of the 8-oxoguanine deaminases and the Thr-Thr dyad of the isoxanthopterin deaminases. It has not been determined how this small difference in amino acid sequence will affect substrate specificity.

Isoxanthopterin has been isolated from multiple biological sources such as zebrafish embryos (*30*), the firebug *Pyrrhocoris apterus* (*31*), caviar (*32*), and silkworm, where isoxanthopterin deaminase activity was measured (*33*). Isoxanthopterin deaminase activity has previously been observed in extracts of *Alcaligenes faecalis* (*34*). A search of that genome was conducted in an attempt to identify the specific deaminase that may be similar to Sgx9339a and Sgx9236b, but unfortunately, the complete sequence of *A. faecalis* is not yet publicly available. Sgx9339a and Sgx9236b are the first proteins of known sequence found to possess significant deaminase activities against the pteridine compounds, isoxanthopterin and pterin 6-carboxylate.

ACKNOWLEDGMENT

We thank Abby Sisco for help in the construction of Figure 1 and sequence analyses within cog0402. We thank the NYSGXRC protein production team for the preparation of Sgx9339a and Sgx9236b. We gratefully acknowledge data collection support for beamline X29 (NSLS).

SUPPORTING INFORMATION AVAILABLE

A supplementary table (Table S1) that identifies enzymes predicted to catalyze the deamination of isoxanthopterin. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Cummings, J. A., Fedorov, A. A., Xu, C., Brown, S., Fedorov, E., Babbitt, P. C., Almo, S. C., and Raushel, F. M. (2009) Annotating enzymes of uncertain function: The deacylation of D-amino acids by members of the amidohydrolase superfamily. *Biochemistry 48*, 6469–6481.
2. Xiang, D. F., Patskovsky, Y., Xu, C., Meyer, A. J., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2009) Functional identification of incorrectly annotated prolidases from the amidohydrolase superfamily of enzymes. *Biochemistry 48*, 3730–3742.
3. Nguyen, T. T., Brown, S., Fedorov, A. A., Fedorov, E. V., Babbitt, P. C., Almo, S. C., and Raushel, F. M. (2008) At the periphery of the amidohydrolase superfamily: Bh0493 from *Bacillus halodurans* catalyzes the isomerization of D-galacturonate to D-tagaturonate. *Biochemistry 47*, 1194–1206.
4. Marti-Arbona, R., Xu, C., Steele, S., Weeks, A., Kuty, G. F., Seibert, C. M., and Raushel, F. M. (2006) Annotating enzymes of unknown function: *N*-Formimino-L-glutamate deiminase is a member of the amidohydrolase superfamily. *Biochemistry 45*, 1997–2005.
5. Hermann, J. C., Marti-Arbona, R., Fedorov, A. A., Fedorov, E., Almo, S. C., Shoichet, B. K., and Raushel, F. M. (2007) Structure-based activity prediction for an enzyme of unknown function. *Nature 448*, 775–779.
6. Pieper, U., Chiang, R., Seffernick, J. J., Brown, S. D., Glasner, M. E., Kelly, L., Eswar, N., Sauder, J. M., Bonanno, J. B., Swaminathan, S., Burley, S. K., Zheng, X., Chance, M. R., Almo, S. C., Gerlt, J. A., Raushel, F. M., Jacobson, M. P., Babbitt, P. C., and Sali, A. (2009) Target selection and annotation for the structural genomics of the amidohydrolase and enolase superfamilies. *J. Struct. Funct. Genomics 10*, 107–125.
7. Hermann, J. C., Ghanem, E., Li, Y., Raushel, F. M., Irwin, J. J., and Shoichet, B. K. (2006) Predicting substrates by docking high-energy intermediates to enzyme structures. *J. Am. Chem. Soc. 128*, 15882–15891.
8. Irwin, J. J., Raushel, F. M., and Shoichet, B. K. (2005) Virtual screening against metalloenzymes for inhibitors and substrates. *Biochemistry 44*, 12316–12328.
9. Xiang, D. F., Xu, C., Kumaran, D., Brown, A. C., Sauder, J. M., Burley, S. K., Swaminathan, S., and Raushel, F. M. (2009) Functional annotation of two new carboxypeptidases from the amidohydrolase superfamily of enzymes. *Biochemistry 48*, 4567–4576.
10. Xiang, D. F., Kolb, P., Fedorov, A. A., Meier, M. M., Fedorov, L. V., Nguyen, T. T., Sterner, R., Almo, S. C., Shoichet, B. K., and Raushel, F. M. (2009) Functional annotation and three-dimensional structure of Dr0930 from *Deinococcus radiodurans*, a close relative of phosphotriesterase in the amidohydrolase superfamily. *Biochemistry 48*, 2237–2247.
11. Holm, L., and Sander, C. (1997) An evolutionary treasure: Unification of a broad set of amidohydrolases related to urease. *Proteins 28*, 72–82.
12. Seibert, C. M., and Raushel, F. M. (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry 44*, 6383–6391.
13. Martynowski, D., Eyobo, Y., Li, T., Yang, K., Liu, A., and Zhang, H. (2006) Crystal structure of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase: Insight into the active site and catalytic mechanism of a novel decarboxylation reaction. *Biochemistry 45*, 10412–10421.
14. Liu, A., and Zhang, H. (2006) Transition metal-catalyzed nonoxidative decarboxylation reactions. *Biochemistry 45*, 10407–10411.
15. Hara, H., Masai, E., Katayama, Y., and Fukuda, M. (2000) The 4-oxalomesaconate hydratase gene, involved in the protocatechuate 4,5-cleavage pathway, is essential to vanillate and syringate degradation in *Sphingomonas paucimobilis SYK-6*. *J. Bacteriol. 182*, 6950–6957.
16. Williams, L., Nguyen, T., Li, Y., Porter, T. N., and Raushel, F. M. (2006) Uronate isomerase: A nonhydrolytic member of the amidohydrolase superfamily with an ambivalent requirement for a divalent metal ion. *Biochemistry 45*, 7453–7462.
17. Nguyen, T. T., Fedorov, A. A., Williams, L., Fedorov, E. V., Li, Y., Xu, C., Almo, S. C., and Raushel, F. M. (2009) The mechanism of the reaction catalyzed by uronate isomerase illustrates how an isomerase may have evolved from a hydrolase within the amidohydrolase superfamily. *Biochemistry 48*, 8879–8890.
18. 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.
19. Hall, R. S., Fedorov, A. A., Marti-Arbona, R., Fedorov, E. V., Kolb, P., Sauder, J. M., Burley, S. K., Shoichet, B. K., Almo, S. C., and Raushel, F. M. (2010) The Hunt for 8-Oxoguanine Deaminase. *J. Am. Chem. Soc. 132*, 1762–1763.
20. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol. 276*, 307–326.
21. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. *Acta Crystallogr. D58*, 1772–1779.
22. De-La-Fortelle, E., and Bricogne, G. (1997) Maximum-likelihood heavy atom parameter refinement in the MIR and MAD methods. *Methods Enzymol. 276*, 472–493.
23. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50*, 760–763.
24. Perrakis, A., Morris, R., and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol. 6*, 458–463.
25. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods in building protein models in electron density map and the location of errors in these models. *Acta Crystallogr. A47*, 110–119.
26. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Somonsom, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
27. Ireton, G. C., McDermott, G., Black, M. E., and Stoddard, B. L. (2002) The structure of *Escherichia coli* cytosine deaminase. *J. Mol. Biol. 315*, 687–697.
28. Wilson, D. K., Rudolph, F. B., and Quiocho, F. A. (1991) Atomic structure of adenosine deaminase complexed with a transition-state analog: Understanding catalysis and immunodeficiency mutations. *Science 252*, 1278–1284.
29. Levy, C. C., and Mc, N. W. (1962) The biological transformation of xanthopterin by a bacterium isolated from soil. *Biochemistry 1*, 1161–1170.
30. Ziegler, I., McDonald, T., Hesslinger, C., Pelletier, I., and Boyle, P. (2000) Development of the pteridine pathway in the zebrafish, *Danio rerio*. *J. Biol. Chem. 275*, 18926–18932.
31. Bel, Y., Porcar, M., Socha, R., Nemec, V., and Ferre, J. (1997) Analysis of pteridines in *Pyrrhocoris apterus* (L.) (Heteroptera, Pyrrhocoridae) during development and in body-color mutants. *Arch. Insect Biochem. Physiol. 34*, 83–98.
32. Andondonskaja-Renz, B., and Zeitler, H. J. (1989) Determination of pteridines in royal jelly and caviar by reverse phase high-performance liquid chromatography. *J. Micronutr. Anal. 5*, 83–90.
33. Gyure, W. L. (1974) Catabolism of isoxanthopterin during the development of the silkworm, *Bombyx mori*. *Insect Biochem. 4*, 303–312.
34. McNutt, W. S., Jr. (1963) The metabolism of isoxanthopterin by *Alcaligenes faecalis*. *J. Biol. Chem. 238*, 1116–1121.