BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

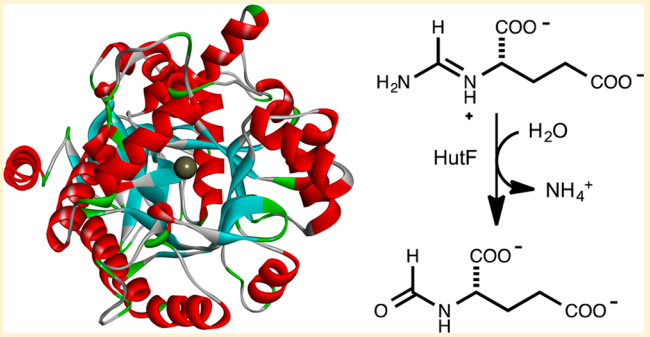
Structure of *N*-Formimino-L-glutamate Iminohydrolase from *Pseudomonas aeruginosa*

Alexander A. Fedorov,[§] Ricardo Martí-Arbona,[†] Venkatesh V. Nemmara,[†] Daniel Hitchcock,[‡] Elena V. Fedorov,[§] Steven C. Almo,*[,§] and Frank M. Raushel*[,†,‡]

[†]Department of Chemistry and [‡]Department of Biochemistry & Biophysics, Texas A&M University, College Station, Texas 77843, United States

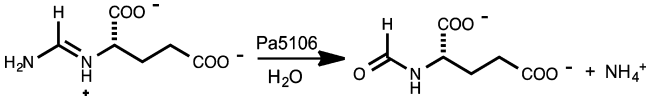
[§]Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, United States

ABSTRACT: *N*-Formimino-L-glutamate iminohydrolase (HutF), from *Pseudomonas aeruginosa* with a locus tag of Pa5106 (gil15600299), is a member of the amidohydrolase superfamily. This enzyme catalyzes the deamination of *N*-formimino-L-glutamate to *N*-formyl-L-glutamate and ammonia in the histidine degradation pathway. The crystal structure of Pa5106 was determined in the presence of the inhibitors *N*-formimino-L-aspartate and *N*-guanidino-L-glutaric acid at resolutions of 1.9 and 1.4 Å, respectively. The structure of an individual subunit is composed of two domains with the larger domain folding as a distorted $(\beta/\alpha)_8$-barrel. The $(\beta/\alpha)_8$-barrel domain is composed of eight β-strands flanked by 11 α-helices, whereas the smaller domain is made up of eight β-strands. The active site of Pa5106 contains a single zinc atom that is coordinated by His-56, His-58, His-232, and Asp-320. The nucleophilic solvent water molecule coordinates with the zinc atom at a distance of 2.0 Å and is hydrogen bonded to Asp-320 and His-269. The α-carboxylate groups of both inhibitors are hydrogen bonded to the imidazole moiety of His-206, the hydroxyl group of Tyr-121, and the side chain amide group of Gln-61. The side chain carboxylate groups of the two inhibitors are ion-paired with the guanidino groups of Arg-209 and Arg-82. Computational docking of high-energy tetrahedral intermediate forms of the substrate, *N*-formimino-L-glutamate, to the three-dimensional structure of Pa5106 suggests that this compound likely undergoes a *re*-faced nucleophilic attack at the formimino group by the metal-bound hydroxide. A catalytic mechanism of the reaction catalyzed by Pa5106 is proposed.



The amidohydrolase superfamily (AHS) was first identified from the three-dimensional structural similarities found within the active sites and global protein folds of urease, phosphotriesterase, and adenosine deaminase.[1] Members of this superfamily are found in every organism sequenced to date and are structurally characterized by a mono-, bi-, or trinuclear metal center embedded at the C-terminal end of a $(\beta/\alpha)_8$-barrel protein fold.[2,3] We have determined that Pa5106 from *Pseudomonas aeruginosa* is a member of this superfamily and catalyzes the deimination of *N*-formimino-L-glutamate to *N*-formyl-L-glutamate and ammonia.[4] This reaction is presented in Scheme 1 and is part of the pathway for the conversion of L-histidine to L-glutamate. Protein sequence analysis places Pa5106 within cog0402 from the Clusters of Orthologous Groups database.[4] The structurally and functionally characterized enzymes within cog0402 include, among others, guanine deaminase,[5] cytosine deaminase,[6] thiomethyl adenosine deaminase,[7] and 8-oxoguanine deaminase.[5] More than 1300 unique protein sequences in cog0402 have been identified within the first 1000 completely sequenced bacterial genomes compiled by NCBI.

Scheme 1



Multiple pathways for the degradation of histidine have been discovered.[8] The conversions of L-histidine to *N*-formimino-L-glutamate are similar in most organisms. In this pathway L-histidine is first converted to urocanate and ammonia by histidine ammonia-lyase (HutH). Urocanase (HutU) then transforms urocanate to L-5-imidazolone-4-propionate, which is subsequently hydrolyzed to *N*-formimino-L-glutamate by imidazolonepropionate amidohydrolase (HutI). HutI is also a member of the AHS but is localized in cog1228.[9] *N*-Formimino-L-glutamate is further transformed to glutamate by three independent paths. In the liver, L-glutamate *N*-formimidoyltransferase transfers the formimino group from *N*-formimino-L-glutamate to tetrahydrofolic acid.[10] Alternatively, in bacteria such as *Aerobacter aerogenes*, *N*-formimino-L-glutamate is hydrolyzed to L-glutamate and formamide in a

Received: October 16, 2014
Revised: December 24, 2014
Published: January 6, 2015

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897

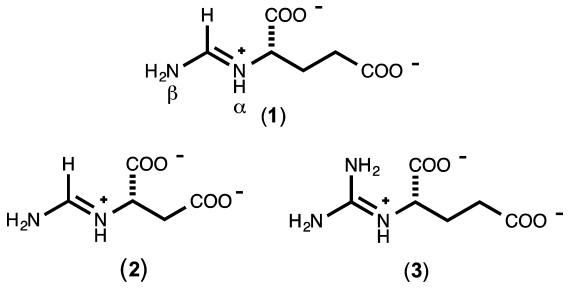
single hydrolysis reaction.[11] However, in other bacteria the conversion of *N*-formimino-L-glutamate to glutamate occurs by two consecutive hydrolytic reactions. *N*-formimino-L-glutamate is first deiminated to *N*-formyl-L-glutamate and ammonia by *N*-formimino-L-glutamate iminohydrolase (HutF), and then *N*-formyl-L-glutamate is hydrolyzed to formate and L-glutamate by *N*-formyl-L-glutamate deformylase (HutG).[12−15]

HutF has been cloned from *P. aeruginosa* PA01 and has been identified with the locus tag of Pa5106 (gi|15600299).[4] Amino acid sequence alignments of this enzyme and other members of the AHS suggest that this enzyme will bind a single divalent cation in the active site and will be ligated to His-56, His-58, His-232, and Asp-320.[16] In addition to these residues, protein transfers will likely be facilitated by Glu-235, His-269, or both. HutF from *P. aerugninosa* hydrolyzes *N*-formimino-L-glutamate with a k_{cat} of 10 s^{-1} and a Michaelis constant of 120 μM.[4] However, this enzyme exhibited no activity for the deimination of other *N*-substituted amino acids, including L-aspartate. In this paper, we present the crystal structure of HutF from *P. aeruginosa* in the presence of the inhibitors *N*-formimino-L-aspartate and *N*-guanidino-L-glutaric acid. This structure has provided insight into the mechanism for the deimination reaction and has identified conserved residues within the active site that are required for substrate recognition.

MATERIALS AND METHODS

Materials. The genomic DNA from *P. aeruginosa* was purchased from the American Type Culture Collection. The oligonucleotide synthesis and DNA sequencing reactions were performed by the Gene Technology Laboratory of Texas A&M University. The pET30a(+) expression vector and the BL21(DE3) star cells were acquired from Novagen. The *N*-formimino-L-glutamate complexes (**1**, Scheme 2) and *N*-formimino-L-aspartate (**2**) were synthesized as previously described.[4] 2-Guanidino-L-glutaric acid (**3**) was purchased from Sigma-Aldrich. All other chemicals were obtained from Sigma-Aldrich unless otherwise stated.

Scheme 2



Cloning and Purification of Pa5106. The gene encoding Pa5106 (HutF) was cloned from *P. aeruginosa* PA01 into a pET30a(+) expression vector as described previously.[4] Wild-type HutF was transformed and expressed in BL21(DE3) star cells. A single colony was grown overnight in 50 mL of LB medium containing 50 μM kanamycin and then used to inoculate 4.0 L of the same medium. Cell cultures were grown at 37 °C with a rotary shaker until an A_{600} of ∼0.6 was reached, after which induction was initiated by the addition of 1.0 mM isopropyl−thiogalactoside (IPTG). The culture was incubated overnight at 30 °C. The bacterial cells were isolated by centrifugation at 6500*g* for 15 min at 4 °C. The pellet was resuspended in 50 mM HEPES, pH 7.5, containing 0.1 mg/mL of the protease inhibitor PMSF and disrupted by sonication. The soluble protein was separated from the cell debris by centrifugation at 12000*g* for 15 min at 4 °C. Nucleic acids were precipitated by the addition of protamine sulfate to 1.5% (w/v). The protein solution was fractionated between 40% and 60% saturated ammonium sulfate. The precipitated protein from the 40−60% saturated ammonium sulfate pellet was resuspended in 50 mM HEPES, pH 7.5, and loaded onto a HiLoad 26/60 Superdex 200 prep grade gel filtration column (GE Health Care) and eluted with the same buffer. Fractions containing Pa5106 were pooled and loaded onto a 6 mL Resource Q ion exchange column (GE Health Care) and eluted with a gradient of NaCl in 20 mM HEPES, pH 7.5. The fractions that contained Pa5106 were pooled and precipitated by the addition of ammonium sulfate to 65% saturation. The sample was centrifuged at 12000*g* for 15 min at 4 °C and then resuspended in a minimal amount of HEPES, pH 7.5. The final step in the purification was accomplished with a High Load 26/60 Superdex 200 prep grade gel filtration column where the protein was eluted with 50 mM HEPES, pH 7.5. The purity of the protein during the isolation procedure was monitored by SDS-PAGE.

Metal Analysis. Metal-free Pa5106 was prepared and reconstituted with Zn^{2+}, as previously described.[16] Purified Pa5106 was treated with 3 mM dipicolinate at 4 °C at pH 5.6 for 48 h. The chelator was removed by loading the mixture onto a PD10 column (GE Health Care) and eluting with metal-free HEPES, pH 7.5. The apo-Pa5106 was reconstituted with 1.0 equiv of Zn^{2+} in 50 mM HEPES, pH 7.5. The metal content of the apo-Pa5106 and the metal-reconstituted enzymes was quantified using inductively coupled plasma emission-mass spectrometry (ICP-MS).

Crystallization and Data Collection. Diffraction quality crystals of SeMet-labeled Pa5106 could not be obtained; therefore, all subsequent crystallization and structure determination efforts utilized native Pa5106. Three different crystal forms of Pa5106 were grown by the hanging drop method at room temperature (Table 1). The first and second crystal forms contained the inhibitor *N*-formimino-L-aspartate (**2**) and the third crystal form contained the inhibitor 2-guanidino-L-glutarate (**3**). For the first *N*-formimino-L-aspartate complex, the solution contained Pa5106 (20 mg/mL) in 20 mM HEPES, pH 8.1, 130 mM NaCl, 1.0 mM $ZnCl_2$, and 50 mM *N*-formimino-L-aspartate (**2**). The precipitant solution contained 1.4 M sodium sulfate, pH 6.9, and 1.0 mM $ZnCl_2$. Crystals appeared in 6−7 days and exhibited diffraction consistent with the space group *C*2, with two molecules of the protein per asymmetric unit (Table 1). A mercury derivative for this crystal form was prepared by soaking the crystals for 5 min in mother liquor containing 1.0 mM ethyl mercury phosphate (EMP).

For the second *N*-formimino-L-aspartate complex, the protein solution contained Pa5106 (10 mg/mL) in 20 mM HEPES, pH 8.1, 130 mM NaCl, 1.0 mM $ZnCl_2$, and 50 mM *N*-formimino-L-aspartate. The precipitant solution contained 1.4 M sodium phosphate, pH 6.9, and 1.0 mM $ZnCl_2$. Crystals appeared in 2−3 days and exhibited diffraction consistent with space group C2, with four molecules of the protein per asymmetric unit.

For the 2-guanidino-L-glutarate (**3**) complex, the protein solution contained Pa5106 (10 mg/mL) in 20 mM HEPES, pH 8.1, 130 mM NaCl, 1.0 mM $ZnCl_2$, and 50 mM *N*-guanidino-L-glutarate. The precipitant contained 1.1 M sodium malonate,

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897

Table 1. Data Collection and Refinement Statistics for Crystals of Pa5106

	Pa5106 Zn^{2+}·*N*-formimino-L-aspartate·Hg^{2+}	Pa5106 Zn^{2+}·*N*-formimino-L-aspartate	Pa5106 Zn^{2+}·*N*-guanidino-L-glutarate
	Data Collection		
space group	*C*2	*C*2	$I4_122$
number of molecules in asymmetric unit	2	4	1
	Cell Dimensions		
a, *b*, *c* (Å)	101.56, 141.83, 86.41	304.05, 67.25, 98.23	133.21, 133.21, 124.84
β (deg)	107.18	91.50	
resolution (Å)	1.86	1.90	1.4
number of unique reflections	97 124	148 848	102 752
R_{merge}	0.083	0.073	0.089
$I/\sigma I$	17.6	16.8	43.2
completeness (%)	99.8	94.8	94.0
	Refinement		
resolution (Å)	25.0−1.86	25.0−1.9	25.0−1.4
R_{cryst}	0.159	0.169	0.179
R_{free}	0.176	0.203	0.189
	RMS Deviations		
bond lengths (Å)	0.007	0.007	0.006
bond angles (deg)	1.06	1.04	1.05
	Number of Atoms		
protein	7009	13 800	3454
waters	906	1054	372
inhibitors	22	44	13
metal ions	2 Zn^{2+}, 6 Hg^{2+}	4 Zn^{2+}	1 Zn^{2+}
PDB entry	4RZB	3MDW	3MDU

0.1 M HEPES, pH 7.0, 0.05% Jeffamine, and 1.0 mM $ZnCl_2$. Crystals appeared in 1 week and exhibited diffraction consistent with space group $I4_122$, with one molecule of the protein per asymmetric unit (Table 1).

Prior to data collection, crystals were transferred to cryoprotectant solutions composed of their mother liquids and 20% glycerol. After incubation for ∼10 s, the crystals were flash-cooled in a nitrogen stream. Three diffraction data sets were collected at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector. Diffraction intensities were integrated and scaled with programs DENZO and SCALEPACK.[17] The data collection statistics are given in Table 1.

Structure Determination and Model Refinement. The structure of the mercury derivative (Pa5106·Zn^{2+}·*N*-formimino-L-aspartate·Hg^{2+}) was determined by single anomalous dispersion (SAD) methods with PHENIX.[18,20] Almost 90% of all residues in the two polypeptide chains were initially modeled, followed by iterative cycles of model rebuilding with COOT,[19] refinement with PHENIX,[20] and automatic model rebuilding with ARP.[21] The final model converged with an R_{cryst} of 0.159 and an R_{free} of 0.176 at 1.86 Å. The inhibitor *N*-formimino-L-aspartate and a single Zn^{2+} are bound to each protein molecule and are well-defined in the electron density maps. Each protein molecule binds three Hg^{2+} ions near Cys-102 and Cys-243.

The structure of the second complex Pa5106·Zn^{2+}·*N*-formimino-L-aspartate was determined by molecular replacement with PHENIX[20] using coordinates of the mercury derivative as the search model. This structure was refined with an R_{cryst} of 0.169 and an R_{free} of 0.203 at 1.9 Å. Each protein molecule binds one inhibitor molecule and one Zn^{2+} ion.

The same search model was used to determine the structure of the Pa5106·Zn^{2+}·*N*-guanidino-L-glutarate complex with PHENIX.[20] A single inhibitor molecule and single Zn^{2+} ion are bound to the protein molecule in this complex. This structure was refined with an R_{cryst} of 0.179 and an R_{free} of 0.189 at 1.4 Å. Final crystallographic refinement statistics for the three inhibitor complexes of Pa5106 are provided in Table 1.

Molecular Docking. Computational docking of substrates and reaction intermediates to Pa5106 was performed using Autodock Vina[22] with the MGL tools 1.5.4 plugin. The 1.4 Å crystal structure of Pa5106 (PDB id: 3MDU) was used as a rigid receptor in all docking calculations. The zinc ion was retained in the active site for all docking calculations, following the removal of all water molecules. A pdbqt file format of the protein was generated by adding polar hydrogens and Kollmann charges using MGL tools 1.5.4. A grid box was centered at the active site of Pa5106 with dimensions 26 × 26 × 26 $Å^3$ with grid points spaced every 1 Å. Initial three-dimensional structures of the ligands as high-energy tetrahedral intermediates were generated using Marvin View (ChemAxon) and Gasteiger charges were added using MGL tools 1.5.4. The nonpolar hydrogens in the ligands were merged and the rotatable bonds were set to generate flexible coordinates to impart conformational freedom. Ligands **1**, **2**, and **3** were docked with the positive charges delocalized among the nitrogen atoms. High-energy tetrahedral intermediates (both stereoisomers) of **1** and **2** were generated with the positive charges residing in the terminal nitrogen. The flexible ligand coordinates were docked in the rigid protein coordinates of Pa5106 with an exhaustiveness of 15. Standard flexible protocols of Autodock Vina using the iterated local search global algorithm were employed to evaluate the binding affinities of the ligand receptor interactions.[22] The resulting docking poses were visualized and the output coordinates were

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897

overlaid using Discovery Studio Client 3.5 (Accelrys). The output structures with the lowest binding energies were considered and the docked poses were filtered for catalytically productive conformations.

RESULTS AND DISCUSSION

Three-Dimensional Structure of Pa5106. Crystal structures of Pa5106 were determined with *N*-formimino-L-aspartate (**2**) and 2-guanidino-L-glutarate (**3**) in the active site. The structure of Pa5106 complexed with *N*-formimino-L-aspartate was refined at 1.9 Å with an R_{cryst} of 0.169 and an R_{free} of 0.203. The final structure contains protein residues 3−451 and one well-defined zinc ion in each monomer of the asymmetric unit. The structure of Pa5106 complexed with 2-guanidino-L-glutarate was refined at 1.4 Å with an R_{cryst} of 0.179 and an R_{free} of 0.189. The final structure contains protein residues 2−451 and one well-defined zinc ion in the single monomer of the asymmetric unit. The four subunits contained within the asymmetric unit of the *N*-formimino-L-aspartate complex have structures virtually identical to each other with a root-mean-square deviation of 0.25 Å. The structure of an individual subunit is presented in Figure 1 and is composed of

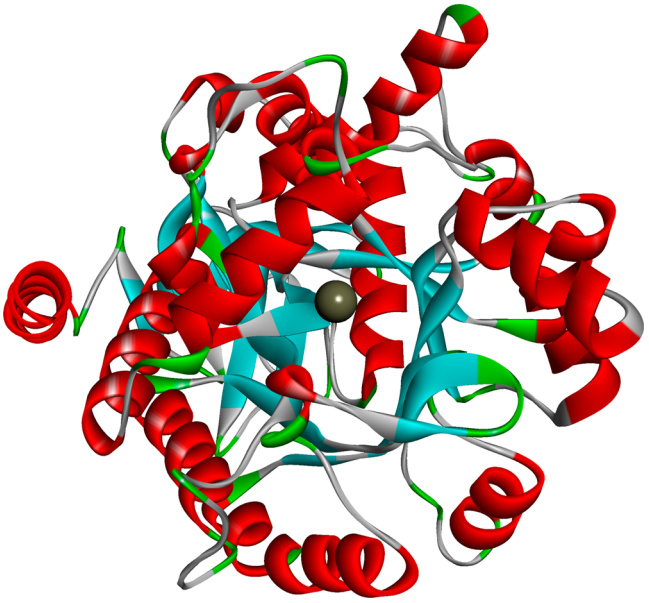


Figure 1. Ribbon representation of Pa5106. Individual monomer of Pa5106 in the presence of *N*-formimino-L-aspartate (PDB id: 3MDW).

two domains. The larger domain extends from Gly-51 to Pro-375 and folds as a distorted $(\beta/\alpha)_8$-barrel. The smaller of the two domains is formed from residues that start at the *N*-terminus (Met-1 to Pro-50) and ends at the C-terminus (Ile-376 to Leu-451).

The subunits of Pa5106 are roughly globular in shape with overall dimensions of approximately 70 Å × 57 Å × 52 Å. The smaller domain of an individual subunit is composed of eight β-strands formed from Ala-3 to Ala-6, Arg-8 to Leu-11, Gly-14 to Arg-17, Val-28 to Ser-24, Val-28 to Ala-30, Asp-386 to Leu-390, Asp-420 to Val-423, and Val-429 to Asp-431, and two α-helices formed by Ala-404 to Phe-411 and Glu-438 to Glu-450. The $(\beta/\alpha)_8$-barrel domain is composed of eight β-strands linked by 11 α-helices and an additional β-strand (between the third β-strand and α-helix). The eight β-strands forming the inside walls of the central barrel structure are formed from Gly-51 to Leu-55, Ala-115 to Tyr-121, Leu-152 to Tyr-159, Gly-202 to Ser-207, Pro-228 to Ala-234, Trp-265 to Ala-270, Val-287 to Cys-291, and Arg-314 to Gly-318.

Quaternary Structure of Pa5106. Previous reports have indicated that Pa5106 from *P. aeruginosa* has a molecular weight of approximately 100 kDa.[4] Because the molecular weight of an individual subunit is 49 215, this observation suggests that two subunits associate to form a dimer. The dimers in the structure of Pa5106 with *N*-guanidino-L-glutarate are formed by two protein molecules connected by a crystallographic 2-fold axis in the space group $I4_122$. Similar dimers are present in the structure of Pa5106 with *N*-formimino-L-aspartate. Here two protein monomers inside the dimers are connected by noncrystallographic 2-fold axis in the space group *C*2. There are two analogous dimers in the asymmetric unit of the monoclinic unit cell. All dimers in both crystal forms are very similar, with an rmsd of 0.7 Å for 902 Cα pairs.

The Active Site. A close-up view of the active site of Pa5106 is illustrated in Figure 2A. The lone zinc ion is

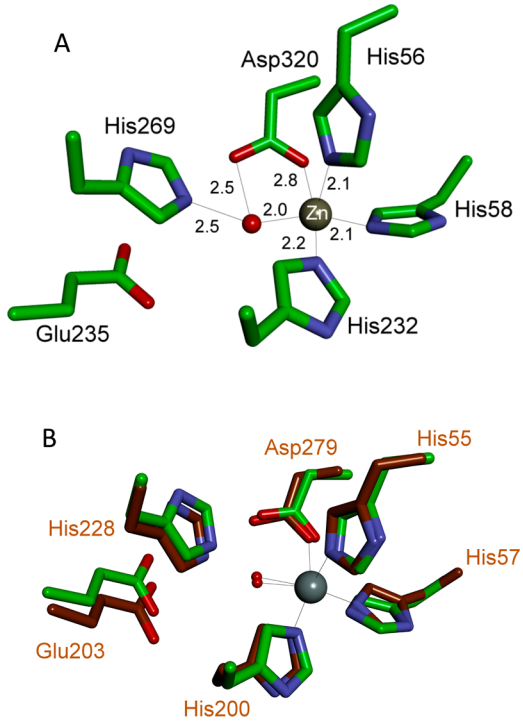


Figure 2. (A) Coordination scheme for the binding of zinc to the active site of Pa5106. The figure was generated from PDB entry 3MDW. The inhibitor, *N*-formimino-L-aspartate, was omitted for clarity. (B) Superimposition of the metal binding residues of Pa5106 (PDB id: 3MDW) with those of Tm0936 from *Thermotoga maritima* (PDB id: 2PLM).

coordinated by His-56 (2.1 Å), His-58 (2.1 Å), His-232 (2.2 Å), and Asp-320 (2.8 Å). A water molecule from solvent is positioned at a distance of 2.0 Å from the zinc to create a distorted trigonal bipyramidal coordination scheme with Asp-320 and His-232 serving as the axial ligands. The bond angles between the axial ligands, zinc, and the equatorial ligands range from 77° to 101°. The nucleophilic water molecule is hydrogen bonded to Asp-320 at a distance of 2.5 Å and to His-269 at a distance of 2.5 Å. This arrangement of coordinating residues to a single divalent cation in the active site of HutF is nearly identical to that of the other structurally characterized enzymes

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897

from cog0402 from the AHS.[3] A comparison of the coordinating environment of Pa5106 with that of Tm0936 from *Thermotoga maritima* is presented in Figure 2B (PDB id: 2PLM). Residues involved in the binding of the divalent cation in the active site of Tm0936 are also conserved in Pa5106, a finding that suggests a similar metal coordination environment within the AHS.

Protein-Inhibitor Interactions. The electrostatic interactions between *N*-formimino-L-aspartate and the protein are depicted in Figure 3A. *N*-Formimino-L-aspartate is a competitive inhibitor of Pa5106 with a K_i value of 30 μM.[4] The α-carboxylate moiety of the inhibitor interacts with His-206 (2.7 Å), the phenolic group of Tyr-121 (2.7 Å), and the side chain amide of Gln-61 (3.1 Å). The side chain carboxylate of the inhibitor is ion-paired with the guanidino groups of Arg-209 (2.9 Å) and Arg-82 (3.1 Å). The terminal nitrogen of the formimino moiety (arbitrarily labeled as β in Scheme 2) of the inhibitor is 3.0 Å from the side chain carboxylate of Glu-235. A water molecule that is not coordinated to the zinc in the active site is hydrogen bonded to the α-amino nitrogen of the inhibitor and to Asp-320 (3.0 Å) and Gln-61 (2.8 Å). The electrophilic carbon of the formimino group of *N*-formimino-L-aspartate is 3.0 Å from the nucleophilic water molecule. The molecular interactions in the active site in the complex with 2-guanidino-L-glutarate are nearly the same as observed for *N*-formimino-L-aspartate. This compounds inhibits the reaction catalyzed by Pa5106 with a K_i value of 890 μM.[4] The image of this inhibitor in the active site of Pa5106 is presented in Figure 3B.

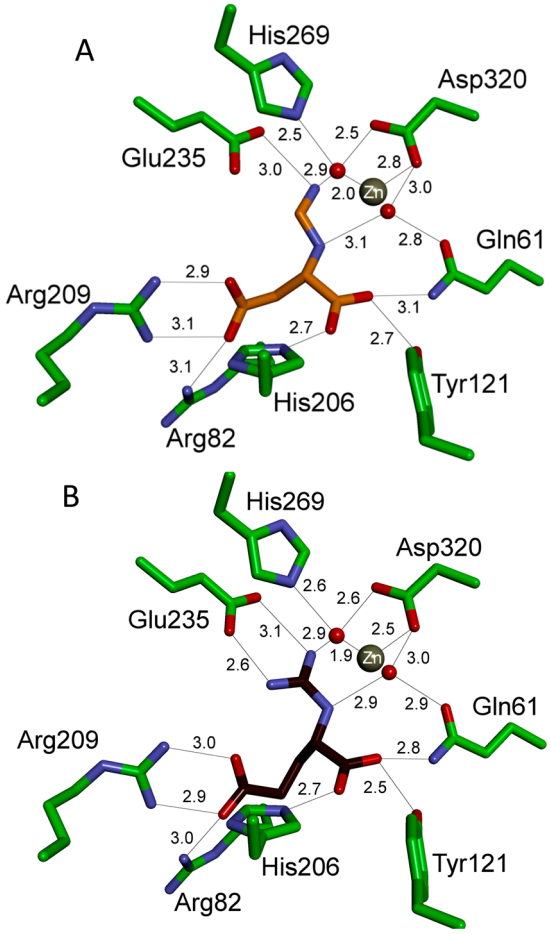


Figure 3. (A) Structure of the active site of Pa5106 complexed with *N*-formimino-L-aspartate (PDB id: 3MDW). (B) Structure of the active site of Pa5106 complexed with *N*-guanidino-L-glutaric acid (PDB id: 3MDU).

Computational Docking of Ligands to the Active Site of Pa5106. Computational docking of the substrate and potential high-energy reaction intermediates to the active site was conducted in an attempt to provide a more detailed view of how the zinc-bound hydroxide attacks the formimino group of the substrate and is hydrolyzed to ammonia and *N*-formyl-L-glutamate.[23] The computationally docked poses, modeled using the program Autodock Vina, of *N*-formimino-L-aspartate (**2**) and *N*-guanidino-L-glutaric acid (**3**) to the active site of Pa5106 are presented in Figure 4A,B, respectively. The active site

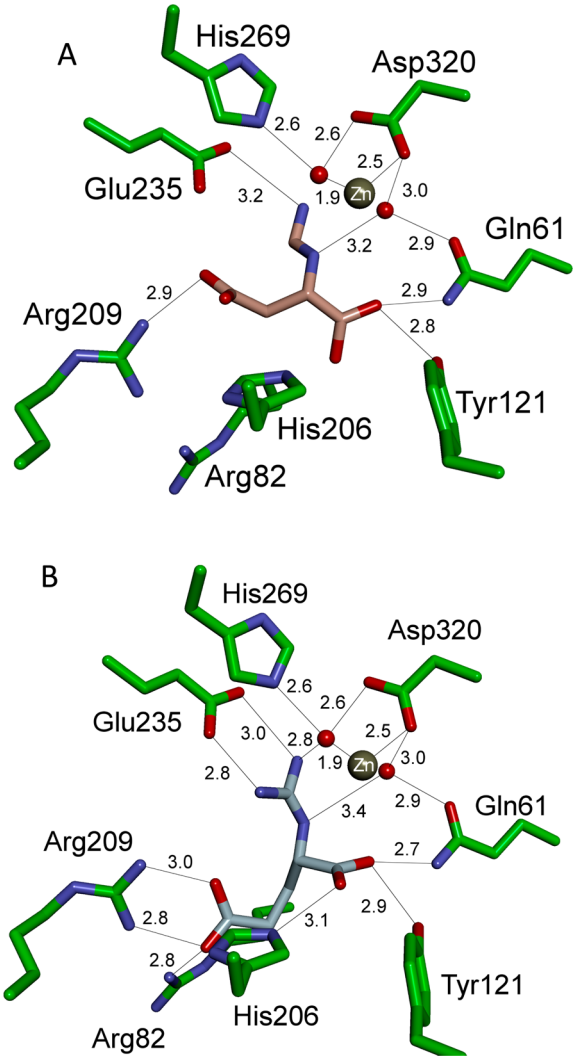


Figure 4. (A) Crystal structure of Pa5106 docked with *N*-formimino-L-aspartate. (B) Crystal structure of Pa5106 docked with *N*-guanidino-L-glutaric acid.

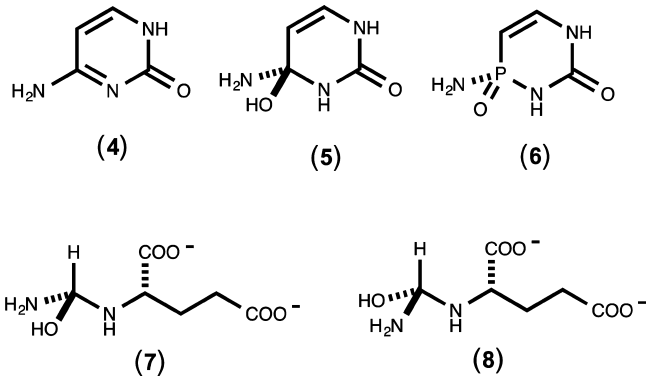
interactions of the computationally docked ligands closely resemble those of the two inhibitors in the crystal structures of Pa5106. The α-amino groups of both ligands interact with a water molecule at a hydrogen bonding distance of 3.2−3.4 Å, an interaction that is also observed in the two crystal structures (Figure 3A,B). This water molecule, which is 4.0 Å from the Zn

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897

atom, is held to the protein by hydrogen bonding interactions with Asp-320 and Gln-61.

The two computationally docked poses of the substrate, *N*-formimino-L-glutamate, as a high-energy tetrahedral intermediate (Scheme 3) in the active site of Pa5106 are shown in Figure

Scheme 3



5A,B.[23] The tetrahedral intermediate **7** was formed by the *re*-face attack of hydroxide on the electrophilic carbon atom of the formimino group, whereas tetrahedral intermediate **8** was formed by nucleophilic attack on the *si*-face. The relative binding orientation of tetrahedral intermediate **7** docked in the active site of Pa5106 is different from that of the two inhibitors in the active site. The side chain carboxylate of intermediate **7** is "flipped" with respect to the α-carboxylate. In the preferred docking pose of intermediate **7**, the α-carboxylate group is ion-paired with the guanidino group of Arg-209, whereas in the X-ray structures of compounds **2** and **3**, Arg-209 interacts with the side chain carboxylate. In the docked pose of the proposed intermediate **7**, the side-chain carboxylate interacts with His-206 and Tyr-121, and the α-amino group interacts with Glu-235. In this docking pose the hydroxyl group of intermediate **7** occupies the original position of the metal-bound hydroxide ion displayed in the crystal structure and the terminal amino group (β in Scheme 2) of the formimino substituent interacts with Asp-320.

The binding orientation of the formimino group in the proposed tetrahedral intermediate **7** is similar to that of a tetrahedral phosphonate mimic (compound **6** of Scheme 3) in the crystal structure of cytosine deaminase.[6] In that structure the terminal amino group of the phosphonocytosine inhibitor (**6**) interacts with Glu-217 (analogous to Glu-235 in Pa5106) and the metal-bound Asp-313 (analogous to Asp-320 in Pa5106) in the active-site. The oxygen atom of the inhibitor mimics the metal-bound hydroxide ion that attacks from the *re*-face of cytosine (compound **5** of Scheme 3) to establish the stereochemistry of nucleophilic attack within the cytosine deaminase reaction.[6] An identical stereochemistry is also observed in a prolidase (Sgx9260b) from the AHS that catalyzes the hydrolysis of an *N*-acyl amide bond to proline containing substrates.[24] The crystal structure of Sgx9260b complexed with *N*-methylphosphonate-L-proline established the relative orientation of the proline moiety in the active site of this enzyme. In that structure the pro-*R* oxygen atom of the phosphonate moiety of the inhibitor mimics the metal bound hydroxide group that undergoes nucleophilic attack on the carbonyl carbon of the substrate. It is thus highly likely that Pa5106, similar to other amidohydrolases described above,

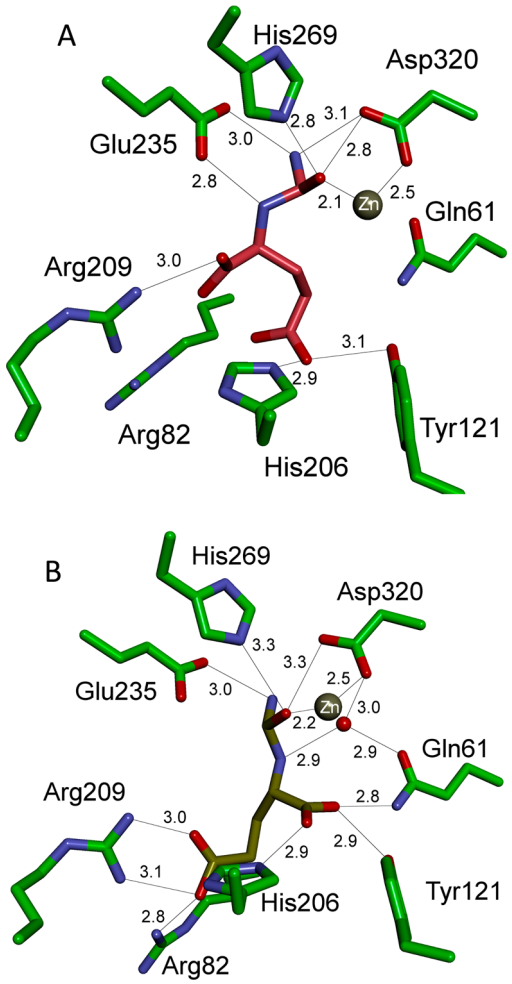


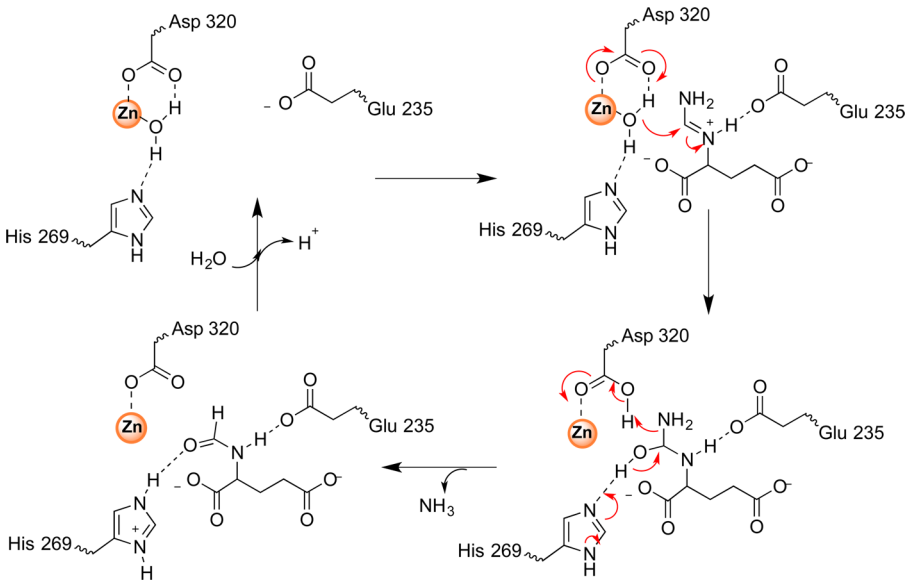
Figure 5. (A) Active site of Pa5106 docked with *N*-formimino-L-glutamate as a high-energy tetrahedral intermediate **4** with an *R* stereocenter formed the attack of water on the *re*-face of the formimino group of the substrate. (B) Active site of Pa5106 docked with *N*-formimino-L-glutamate as a high-energy tetrahedral intermediate **5** with an *S* stereocenter formed by the attack of water on the *si*-face of the formimino group of the substrate.

prefers a *re*-faced attack of hydroxide on the electrophilic carbon of its substrate. All of the structurally characterized enzymes from the amidohydrolyase superfamily show the same relative stereochemical attack of hydroxide on an sp^2-hybridized reaction center.[3]

The docked pose of the proposed intermediate **8** more closely resembles the crystal structures of the bound inhibitors. It is thus possible that although intermediate **8** could lead to a nonsubstrate-like orientation, intermediate **7** would lead to a productive substrate-like orientation and hence catalysis. The two possible tetrahedral reaction intermediates formed from the inhibitor *N*-formimino-L-aspartate (**2**) exhibited docking poses with the side chains and forminino substituents bound in a manner similar to that of the inhibitor structure shown in Figure 3A (data not shown). The reasons for the non-productive poses of the tetrahedral intermediate formed from *N*-formimino-L-aspartate are not immediately clear; this compound is not a substrate for Pa5106 and the unproductive binding poses of the substrate and tetrahedral intermediates may explain why this compound is not a substrate for Pa5106.

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890–897

Scheme 4



Mechanistic Implications. A mechanism for the hydrolysis of *N*-formimino-L-glutamate in the active site of HutF is presented in Scheme 4. Examination of the X-ray crystal structure of *N*-formimino-L-aspartate complexed to Pa5106 corroborates that the most likely residues involved as general acid/base groups are Asp-320, His-269, and Glu-235. Upon binding of the substrate to the active site, catalysis is initiated by the abstraction of a proton from the water molecule bound to zinc by Asp-320. Nucleophile attack on the carbon center of the formimino substituent creates a tetrahedral intermediate with *R*-stereochemistry, which is stabilized through interactions with Glu-235, His-269, and Asp-320. The collapse of the tetrahedral intermediate is facilitated by a proton transfer to His-269 and protonation of the leaving group ammonia by Asp-320. The simple computational docking results with intermediate 7 support the proposed mechanism.

In the crystal structures of the two inhibitors, the α-carboxylate group interacts electrostatically with Gln-61 and Tyr-121, whereas the side chain carboxylate interacts with the guanidino groups of Arg-209 and Arg-82. In the computationally docked pose of the putative tetrahedral intermediate of the actual substrate these interactions are reversed, with the α-carboxylate group ion-paired with the guanidino group of Arg-209 and the side chain carboxylate interacting with Tyr-121. An amino acid sequence alignment of 39 homologues to Pa5106 from various bacterial species has been constructed.[4] This alignment indicates that residues corresponding to Arg-82, Arg-209, His-206, and Tyr-121 are fully conserved but that Gln-61 is not conserved in these protein sequences.

SUMMARY

The three-dimensional structure of *N*-formimino-L-glutamate iminohydrolase (HutF) from *P. aeruginosa* (Pa5106) was determined in the presence of two different inhibitors bound in the active site. Computational docking of the proposed tetrahedral intermediate for the deaminase reaction catalyzed by this enzyme to the active site of Pa5106 has suggested a chemical mechanism for the conversion of substrate to products. In this mechanism a zinc-bound water molecule is activated by interactions with Asp-320 and His-269, whereas the formimino group of the substrate is activated by interaction with Glu-235. The proposed reaction mechanism is similar to those previously described for the deamination of nucleic acid bases by other members of cog0402 from the AHS.

AUTHOR INFORMATION

Corresponding Authors

*Telephone: (718) 430-2746. E-mail: steve.almo@einstein.yu.edu.

*Telephone: (979)-845-3373. E-mail: raushel@tamu.edu.

Funding

This work was supported in part by the National Institutes of Health (GM 71790) and the Robert A. Welch Foundation (A-840). The X-ray coordinates and structure factors for Pa5106 from *Pseudomonas aeruginosa* have been deposited in the Protein Data Bank (PDB accession codes 4RZB, 3MDW, and 3MDU).

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Holm, L., and Sander, C. (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins 28*, 72−82.

(2) Roodveldt, C., and Tawfik, D. S. (2005) Shared promiscuous activities and evolutionary features in various members of the amidohydrolase superfamily. *Biochemistry 44*, 12728−12736.

(3) Seibert, C. M., and Raushel, F. M. (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry 44*, 6383−6391.

(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) Hall, S. R., Fedorov, A. A., Marti-Arabona, R., Fedorov, E. V., Kold, P., Sauder, M. J., Burley, S. K., Shoicket, B. K., Almo, S. C., and Raushel, F. M. (2010) The hunt for 8-oxoguanine deaminase. *J. Am. Chem. Soc. 132*, 1762−1763.

(6) Hall, S. R., Fedorov, A. A., Xu, C., Fedorov, E. V., Almo, S. C., and Raushel, F. M. (2011) Three-dimensional structure and catalytic mechanism of cytosine deaminase. *Biochemistry 50*, 5077−5085.

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897

(7) Hall, S. R., Agarwal, R., Hitchcock, D., Sauder, J. M., Burley, S. K., Swaminathan, S., and Raushel, F. M. (2010) Discovery and structure determination of the orphan enzyme isoxanthopterin deaminase. *Biochemistry 49*, 4374−4382.
(8) Borek, B. A., and Waelsch, H. (1953) The enzymatic degradation of histidine. *J. Biol. Chem. 205*, 459−474.
(9) Tyagi, R., Eswaramoorthy, S., Burley, S. K., Raushel, F. M., and Swaminathan, S. (2008) A common catalytic mechanism for proteins of the HutI family. *Biochemistry 47*, 5608−5615.
(10) Tabor, H., and Wyngarden, L. (1959) The enzymatic formation of formiminotetrahydrofolic acid, 5,10-methenyltetrahydrofolic acid, and 10-formyltetrahydrofolic acid in the metabolism of formiminoglutamic acid. *J. Biol. Chem. 234*, 1830−1846.
(11) Magasanik, B., and Bowser, H. R. (1955) The degradation of histidine by Aerobacter aerogenes. *J. Biol. Chem. 213*, 571−580.
(12) Tabor, H., and Mehler, A. H. (1954) Isolation of *N*-formyl-L-glutamic acid as an intermediate in the enzymatic degradation of L-histidine. *J. Biol. Chem. 210*, 559−568.
(13) Hu, L., and Phillips, A. T. (1988) Organization and multiple regulation of histidine utilization genes in *Pseudomonas putida*. *J. Bacteriol. 170*, 4272−4279.
(14) Hu, L., Mulfinger, L. M., and Phillips, A. T. (1987) Purification and properties of formylglutamate amidohydrolase from *Pseudomonas putida*. *J. Bacteriol. 169*, 4696−4702.
(15) Wickner, R. B., and Tabor, H. (1972) *N*-Formimino-L-glutamate iminohydrolase from histidine-adapted *Pseudomonas*. Purification and properties. *J. Biol. Chem. 247*, 1605−1609.
(16) Marti-Arbona, R., and Raushel, F. M. (2006) Mechanistic characterization of N-formimino-L-glutamate iminohydrolase from *Pseudomonas aeruginosa*. *Biochemistry 45*, 14256−14262.
(17) Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Methods in Enzymology* (Carter, C. W. J., Sweet, R. M., Abelson, J. N., and Simon, M. I., Eds.) pp 307−326, Academic Press, New York.
(18) Long, F., Vagin, A., Young, P., and Murshudov, G. N. (2008) BALBES: A Molecular Replacement Pipeline. *Acta Crystallogr., Sect. D D64*, 125−132.
(19) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr., Sect. D D60*, 2126−2132.
(20) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Jung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr., Sect. D D66*, 213−221.
(21) Lamzin, V. S., and Wilson, K. S. (1993) Automated refinement of protein models. *Acta Crystallogr., Sect. D D49*, 129−147.
(22) Trott, O., and Olson, A. J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J. Comput. Chem. 31* (2), 455−461.
(23) 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.
(24) Xiang, D. F., Patskovsky, Y., Xu, C., Fedorov, A. A., Federov, E. V., Sisco, A. A., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2010) Functional identification and structure determination of two novel prolidases from cog1228 in the amidohydrolase superfamily. *Biochemistry 49*, 6791−6803.

DOI: 10.1021/bi501299y
Biochemistry 2015, 54, 890−897