A Thesis for the Degree of Doctor of Philosophy

Structural and functional studies of two enzymes from *Arabidopsis thaliana* in the ureide pathway: ureidoglycine aminohydrolase and ureidoglycolate amidohydrolase

애기장대 **ureidoglycine aminohydrolase**와 **ureidoglycolate amidohydrolase**의 구조와 기능에 대한 연구

August 2014

Department of Agricultural Biotechnology

Seoul National University

Inchul Shin

Structural and functional studies of two enzymes from *Arabidopsis thaliana* in the ureide pathway: ureidoglycine aminohydrolase and ureidoglycolate amidohydrolase

애기장대 ureidoglycine aminohydrolase와 ureidoglycolate amidohydrolase의 구조와 기능에 대한 연구

지도교수: 이 상 기

이 논문을 농학박사학위논문으로 제출함

2014 년 7 월

서울대학교 대학원

농생명공학부 응용생명화학전공

신인철

신인철의 박사학위논문을 인준함

2014 년 7월

위 원 장 김 정 한 (인)

부위원장 이 상 기 (인)

위 원 오 기 봉 (인)

위 원 배 의 영 (인)

위 원 양 진 국 (인)

Structural and functional studies of two enzymes from *Arabidopsis thaliana* in the ureide pathway: ureidoglycine aminohydrolase and ureidoglycolate amidohydrolase

Advisor : Sangkee Rhee

A Dissertation Submitted in Partial Fulfillment of the Requirement for the Degree of
DOCTOR OF PHILOSOPHY
to the Faculty of
Department of Agricultural Biotechnology
At
SEOUL NATIONAL UNIVERSITY
by
Inchul Shin

Date Approved

_____

_____

_____

_____

_____

_____

ABSTRACT

In plants, the ureide pathway is a metabolic route that converts the ring nitrogen atoms of purine to ammonia via sequential enzymatic reactions, playing an important role in nitrogen recovery. By using x-ray crystallography, crystal structures of two enzymes in this pathway were solved in this study: (*S*)-ureidoglycine aminohydrolase (UGlyAH) and the (*S*)-ureidoglycolate amidohydroalse (UAH), both from *Arabidopsis thaliana*. UGlyAH enzyme converts (*S*)-ureidoglycine into (*S*)-ureidoglycolate and ammonia, providing the final substrate to the pathway. Here, a structural and functional analysis of this enzyme from *Arabidopsis thaliana* (AtUGlyAH) is reported. The crystal structure of AtUGlyAH in the ligand-free form shows a monomer structure in the bi-cupin fold of the β-barrel and an octameric functional unit, as well as an Mn^{2+} ion binding site. The structure of AtUGlyAH in complex with (*S*)-ureidoglycine revealed that the Mn^{2+} ion acts as a molecular anchor to bind (*S*)-ureidoglycine and its binding mode dictates the enantioselectivity of the reaction. Further kinetic analysis characterized the functional roles of the active site residues, including the Mn^{2+} ion binding site and residues in the vicinity of (*S*)-ureidoglycine. These analyses provide molecular insights into the structure of the enzyme and its possible catalytic mechanism. In the final step of the pathway, UAH catalyzes the conversion of (*S*)-ureidoglycolate into glyoxylate and releases two molecules of ammonia as by-products. UAH is homologous in structure and sequence with allantoate amidohydrolase (AAH), an upstream enzyme in the pathway with a similar function as that of an amidase but with a different substrate. Both enzymes

exhibit strict substrate specificity and catalyze reactions in a concerted manner, resulting in purine degradation. Here, three crystal structures of *Arabidopsis thaliana* UAH: bound with substrate, reaction intermediate, and product, and a structure of *Escherichia coli* AAH complexed with allantoate are reported. Structural analyses of UAH revealed a distinct binding mode for each ligand in a bimetal reaction center with the active site in a closed conformation. The ligand directly participates in the coordination shell of two metal ions and is stabilized by the surrounding residues. In contrast, AAH, which exhibits a substrate-binding site similar to that of UAH, requires a larger active site due to the additional ureido group in allantoate. Structural analyses and mutagenesis revealed that both enzymes undergo an open-to-close conformational transition in response to ligand binding, and that the active site size as well as the interaction environments in each UAH and AAH are determinants for substrate specificity between two structurally homologous enzymes.

Keywords: Enzyme kinetics, Metalloenzymes, Protein structure, Nitrogen metabolism, Purine catabolism, Ureide pathway, Substrate specificity

Student Number: 2008-30927

CONTENTS

LIST OF FIGURES

CHAPTER III

LIST OF TABLES

CHAPTER II

CHAPTER III

LIST OF ABBREVIATION

AAH	allantoate amidohydrolase
AtAAH	AAH from *Arabidopsis thaliana*
EcAAH	AAH from *Escherichia coli*
UGlyAH	(*S*)-ureidoglycine aminohydrolase
AtUGlyAH	UGlyAH from *Arabidopsis thaliana*
UAH	(*S*)-ureidoglycolate amidohydrolase
AtUAH	UAH from *Arabidopsis thaliana*
EDTA	ethylenediaminetetraacetic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
MES	2-(*N*-morpholino)ethanesulfonic acid
NAD^+	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate hydrogen
PCR	polymerase chain reaction
rmsd	root mean-square deviation
SAD	single wavelength anomalous dispersion
SeMet	seleno-L-methionine
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
WT	wild-type

CHAPTER I. INTRODUCTION

Nitrogen metabolism is a biological process present in a wide variety of living organisms, and it is crucial in cell physiology since it supplies an essential nutrient. Nitrogen metabolism includes metabolic routes for nitrogen uptake and assimilation, and catabolism of nitrogen-rich metabolites (1). Of these, nitrogen assimilation is a universal, indispensable process common to all organisms, as identified in *de novo* biosynthesis of purines and pyrimidines (2, 3). In contrast, both nitrogen uptake and degradation of nitrogenous compounds manifest in distinct metabolic pathways in different organisms (4).

In contrast to animals, plants depend on inorganic nitrogen such as nitrate and ammonium as external nitrogen sources and uptake them from soil using nitrate transport systems (5) and ammonium transporter (6). Either directly absorbed or reduced from nitrate, ammonium is incorporated into glutamine and glutamate by glutamine synthetase/glutamate synthase (7). Ammonium can also be assimilated into glutamate using α-ketoglutarate by glutamate dehydrogenase (8). Glutamine and glutamate serve as the amino group donor. The assimilated nitrogen in amino acids is further used to build nucleotides as well as other organic compounds. All of the ring nitrogen atoms of purine and pyrimidine base originate from amino acids. The ring nitrogen atoms of purine come from glutamine, glycine, and aspartate and those of pyrimidine are resulted from glutamine and aspartate (9).

Catabolism of the nitrogenous purine and pyrimidine provides a chance to recycle the ring nitrogen. The unique nitrogen utilization features of the pyrimidine and purine catabolic pathways have been revealed. For example, two oxidative and

reductive pyrimidine degradation pathways have been characterized in *Escherichia coli* (10). An additional Rut pathway was also identified in this bacterium (11, 12). In contrast, an oxidative purine degradation pathway is conserved in most organisms, the primary product of which is uric acid (1). Although uric acid has a high content of nitrogen, with an N to C ratio of 0.8 (Fig. 1), humans and most mammals are unable to utilize the ring nitrogen in uric acid due to a lack of the necessary catabolic pathways. However, there exists an elucidated uric acid degradation pathway that could remobilize the ring nitrogen atoms for further assimilation. This catabolic pathway in purine metabolism, called the ureide pathway, has been characterized in *Arabidopsis thaliana* and the presence of genes for the pathway were also identified in all plants that have been sequenced to-date, as well as in some bacteria and fungi (13–15). These two independent catabolic pathways - the Rut pathway for pyrimidines and the ureide pathway for purines - do not show any genetic or biochemical resemblance, but exhibit common features for producing ammonia as a reaction byproduct. Therefore, the possible physiological role of these degradation pathways has been suggested for nitrogen mobilization from nitrogen-rich metabolites in their respective organisms (4, 12, 14, 15).

The ureide pathway in *A. thaliana* is composed of seven enzymes and has been studied extensively via biochemical and structural investigations (14–22). In plants, the ureide pathway can be divided into two phases (Fig. 1). The first phase, in which the pyrimidine ring of uric acid is cleaved via three enzymatic reactions to produce (*S*)-allantoin, occurs in peroxisomes (13, 23, 24). In the second phase, which takes place in the endoplasmic reticulum, (*S*)-allantoin is converted to glyoxylate in a stepwise manner by four different enzymes, resulting in the release

of four molecules of ammonia (Fig. 1) (14, 15). Specifically, allantoinase opens up the hydantoin ring of (*S*)-allantoin to form allantoate with two ureido groups. Subsequent stepwise enzymatic reactions by allantoate amidohydrolase (AAH), (*S*)-ureidoglycine aminohydrolase (UGlyAH), and (*S*)-ureidoglycolate amidohydrolase (UAH) degrade allantoate, liberating four molecules of ammonia and a glyoxylate. In this way, all four of the ring nitrogen atoms in uric acid are released as ammonia. In contrast to this ureide pathway, which is found in plants, some bacteria and fungi oxidize (*S*)-ureidoglycolate, a substrate for UAH, to oxalurate (25) or degrade it to glyoxylate and urea (26). As non-plant reactions, (*S*)-ureidoglycoate can be directly converted from allantoate (4) and oxalurate can be formed from (*S*)-ureidoglycine by ureidoglycine transaminase (27).

The enzyme UGlyAH (EC 3.5.3.-) catalyzes the conversion of (*S*)-ureidoglycine into (*S*)-ureidoglycolate in an enantioselective manner using an Mn^{2+} ion as a co-factor and releases ammonia (Fig. 1) (14). In this study, structural and functional analyses of UGlyAH from *A. thaliana* (AtUGlyAH) were carried out. In particular, the structure of AtUGlyAH in the ligand-free form and as a complex with (*S*)-ureidoglycine substrate was determined, thus providing molecular insights into the structure of this enzyme and its mechanism.

A UAH enzyme that is unique to plants is responsible for the final reaction in the ureide pathway, which converts (*S*)-ureidoglycolate to glyoxylate with the release of two molecules of ammonia via (*S*)-hydroxyglycine as a reaction intermediate (Figs. 1 and 2) (4). Ureidoglycolate-degrading activity had been described in plants (28, 29) but the gene encoding UAH was identified later in *Arabidopsis thaliana* with the aid of bioinformatics and comparative genomics (15). This UAH exhibited a high sequence similarity (30% identity) with AAH. These

similarities originate from catalytic features common to both enzymes. Both UAH and AAH cleave an amide bond and the substrates for these enzymes are similar in chemical structure, differing only by the presence of an additional ureido group in allantoate (Fig. 1).

UAH and AAH are both present in the endoplasmic reticulum of plants (15, 30). Thus, plants have two catalytically homologous enzymes in the identical catabolic pathway. Stringent substrate specificity between these two enzymes is therefore a key factor that enables the purine catabolic pathway to operate in a concerted way. This study investigated the substrate specificity of these two enzymes and provides three crystal structures of UAH from *Arabidopsis thaliana* (AtUAH). Each of these structures contains a different ligand: (*S*)-ureidoglycolate substrate, (*S*)-hydroxyglycine reaction intermediate, and the glyoxylate product. The crystal structure of AAH from *Escherichia coli* (EcAAH) was also determined as a complex with the allantoate substrate. Structural comparisons and enzyme kinetic analyses provide a molecular basis for the observed substrate specificity between two homologous enzymes.

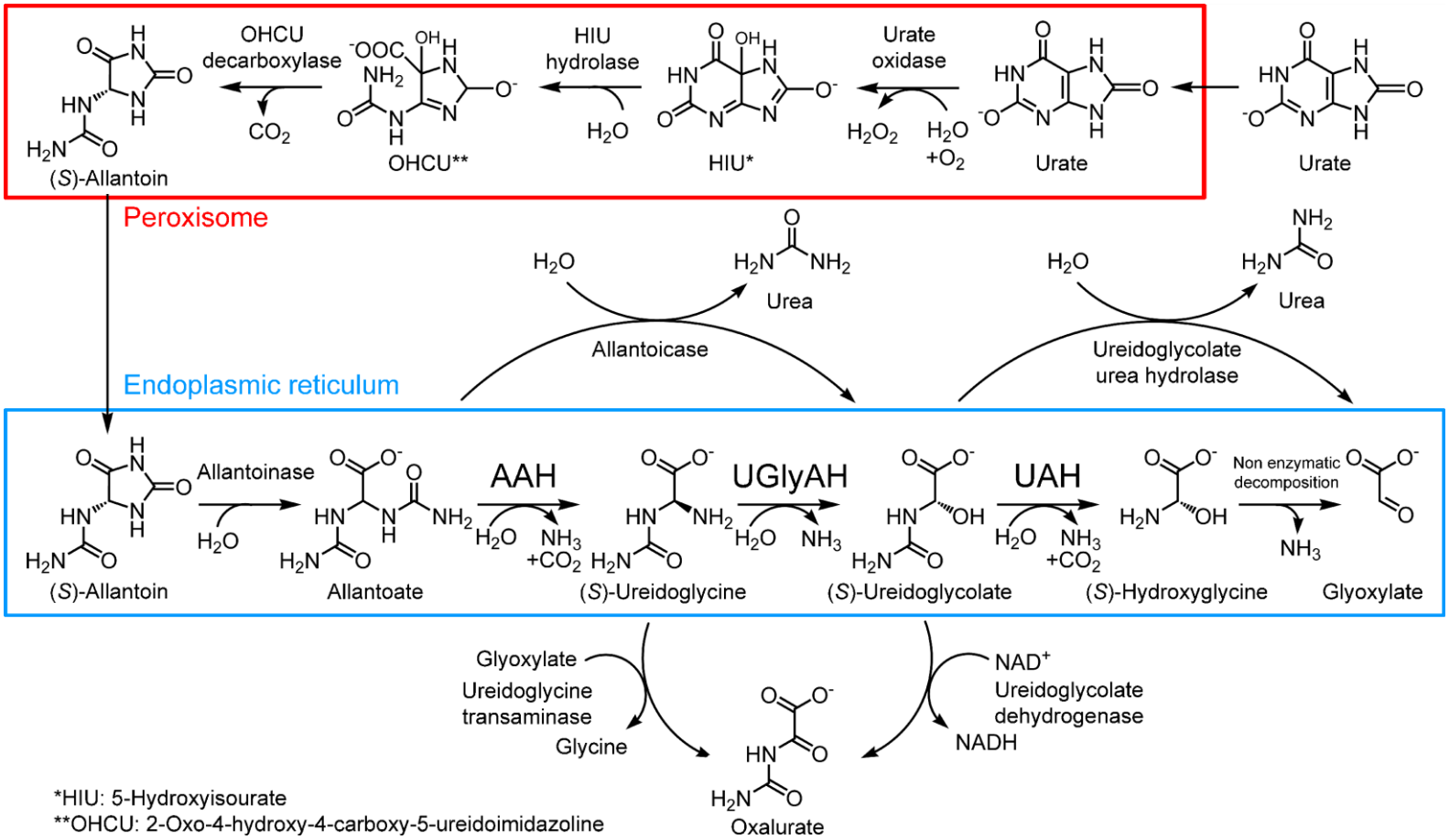

OHCU decarboxylase
HIU hydrolase
Urate oxidase
(S)-Allantoin
OHCU**
HIU*
Urate
Urate
Peroxisome
Endoplasmic reticulum
Urea
Allantoicase
Urea
Ureidoglycolate urea hydrolase
Allantoinase
AAH
UGlyAH
UAH
Non enzymatic decomposition
(S)-Allantoin
Allantoate
(S)-Ureidoglycine
(S)-Ureidoglycolate
(S)-Hydroxyglycine
Glyoxylate
Glyoxylate
Ureidoglycine transaminase
Glycine
NAD+
Ureidoglycolate dehydrogenase
NADH
Oxalurate
*HIU: 5-Hydroxyisourate
**OHCU: 2-Oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline

FIGURE 1. **The ureide pathway.** The full ureide pathway is depicted. In plants, the first half of the ureide pathway occurs in the peroxisome (red box) and the second half takes place in the endoplasmic reticulum (blue box). In this pathway, uric acid is degraded to glyoxylate with releasing four purine ring nitrogen atoms of urate as a form of ammonia.

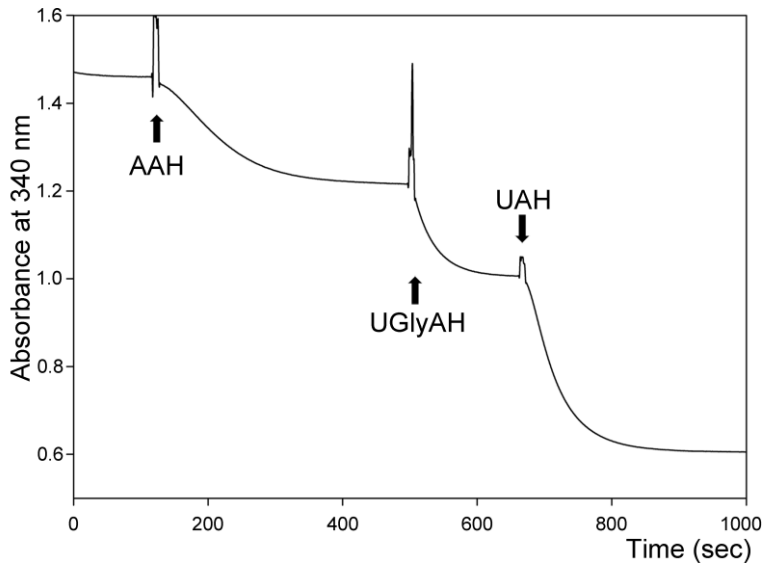


FIGURE 2. **Ammonia production by AtAAH, AtUglyAH and AtUAH.** A time-dependent reaction profile for ammonia release is shown. The enzyme reaction using 30 μM allantoate as a substrate was initially catalyzed by AtAAH followed by successive addition of AtUGlyAH and AtUAH. The ammonia thus produced was detected in a coupled-reaction described in the MATERIALS AND MATHODS. Note that an almost twofold reduction in absorbance was measured after addition of UAH, which is consistent with the production of two molecules of ammonia, in contrast to the one molecule of ammonia released in the preceding two reactions.

CHAPTER II.

Structural and Functional Insights into (*S*)-Ureidoglycine Aminohydrolase, Key Enzyme of Purine Catabolism in *Arabidopsis thaliana*

MATERIALS AND METHODS

1. Construct

The UGlyAH gene from *A. thaliana* (gene access No. NM_117809) was modified and used as a template for PCR. The modifications were performed using the QuickChange method (Agilent) to introduce a silent mutation at the sequences for NdeI and XhoI sites residing in the internal region of the AtUGlyAH gene. The N-terminal truncated variants of AtUGlyAH, from which the signal peptide (residue no. 1–20) and an additional N-terminal region had been removed, were amplified using a pair of sequence-specific primers (Table 1). The resulting PCR products were introduced into the NdeI and XhoI sites of a modified pET-28a vector (Merck) containing a *Tobacco etch virus* protease cleavage site at the junction between a His_5-tag and a multiple cloning site. For the functional assay, all AtUGlyAH mutants were also prepared using the QuickChange method with mutagenic primers (Table 1).

(*S*)-Ureidoglycine was produced by enzyme reactions, and the enzyme activity of AtUGlyAH was assayed using *E. coli* (*S*)-ureidoglycolate dehydrogenase as a coupled enzyme (Fig. 1) (25). Therefore, expression vectors were constructed as described above for genes encoding the enzymes necessary for producing (*S*)-ureidoglycine from a racemic allantoin mixture and also for the gene encoding (*S*)-ureidoglycolate dehydrogenase. Those were as follows: allantoinase from *E. coli* (19); allantoate amidohydrolase from *Arabidopsis thaliana* (AtAAH; gene access No. NM_11826) and *E. coli str. K-12 substr. DH10b* (EcAAH; gene access No. NC_010473:482612–483847); and (*S*)-ureidoglycolate dehydrogenase (gene access No. NC_010473:483869–484918) from *E. coli str. K-12 substr. DH10b*.

TABLE 1. Primers for AtUGlyAH and its mutants

AtUGlyAH(Δ35)	
Forward	GGAATTCCATATGAAGACGAACCCCATTTATTG
Reverse	AACCGCTCGAGTCACAATGGATTTCGATTCAC
H221A	GCTTATGACTTCAACATC**GCT**ACTATGGATTTTCAGCC
E235Q	GGAGAGTTTCTGAATGTTAAG**CAA**GTTCATTATAACCAG
E235A	GAGTTTCTGAATGTTAAG**GCA**GTTCATTATAACCAGC
H237A	CTGAATGTTAAGGAAGTT**GCT**TATAACCAGCATGGTTTG
H241A	GTTCATTATAACCAG**GCT**GGTTTGTTGCTTCTGG
Y252F	GGAGGGGCAAGGCATT**TTT**CGCTTGGGTGATAAC
Q275A	GGCTCCTTTTGTCCCT**GCA**TGGTATGCTGCACTC
Y287F	GAAAGACTAGGTCTCGG**TTT**TTGTTGTACAAAG
Y287A	GAAAGACTAGGTCTCGG**GCT**TTGTTGTACAAAG
K291A	CGGTATTTGTTGTAC**GCA**GATGTGAATCGAAATC
K291R	CGGTATTTGTTGTAC**CGA**GATGTGAATCGAAATC

Sequences are written from 5' to 3'. For mutagenic primers, forward versions are listed. Restriction sites used for cloning are underlined, and the boldfaced-underlines show the mutated sequences.

2. Protein Expression, Purification, and Crystallization

For structural studies, seleno-L-methionine (SeMet)-substituted, N-terminal His-tagged AtUGlyAH was expressed in *E. coli* B834 (DE3) methionine auxotroph cells (Merck) which were grown at 37℃ in minimal medium supplemented with SeMet. When the optical density of the culture medium reached 0.7 at 600 nm, the recombinant protein was expressed for 18 h at 20℃ with the addition of 0.5 mM isopropyl-1-thio-β-d-galactopyranoside. The cells were harvested and sonicated in buffer A (50 mM NaH_2PO_4, pH 7.4, and 200 mM NaCl), and the supernatant was obtained by centrifugation at 30,000 × *g* for one hour at 4℃. The enzyme was purified using an immobilized metal affinity column (GE Healthcare) that had been equilibrated with buffer A, and was eluted with buffer B (buffer A containing 500 mM imidazole). After the pooled elution fractions were dialyzed against buffer C (50 mM NaH_2PO_4, pH 7.4, and 100 mM NaCl), the N-terminal His-tag was removed by treatment of *Tobacco etch virus* protease for 12 h at 4℃. AtUGlyAH in the absence of a His-tag was further purified by immobilized metal affinity chromatography, followed by gel filtration chromatography using Superdex-200 (GE Healthcare) with buffer D (50 mM HEPES, pH7.4, and 100 mM NaCl). The protein was concentrated to ~10 mg/mL, and 1 mM $MnCl_2$ and 5 mM dithiothreitol were added before crystallization.

Crystallization using SeMet-AtUGlyAH was initially carried out by the sitting-drop vapor-diffusion method at 295 K. Of the various constructs, AtUGlyAH (Δ35), from which the N-terminal 35 residues had been removed, produced crystals suitable for further structural determination under two different conditions: crystal-A grown under 0.1 M phosphate citrate, pH 4.2, 10% (w/v) polyethylene glycol 3000, 0.2 M NaCl (Fig. 3*A*), and crystal-B grown under 0.1 M HEPES, pH 7.5, 7%

(w/v) polyethylene glycol 8000, 8% (v/v) ethylene glycol (Fig. 3*B*). These conditions were further optimized using the hanging-drop vapor-diffusion method at 295 K.

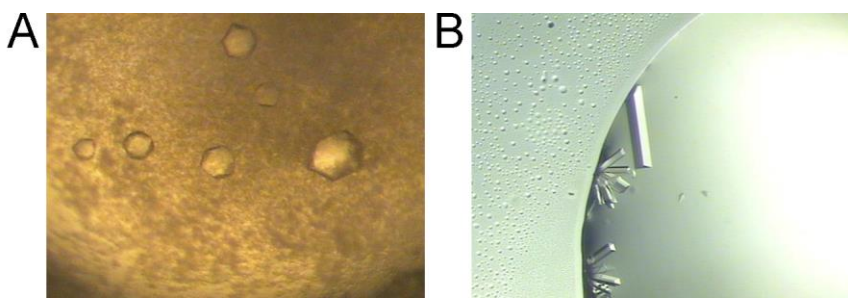


FIGURE 3. **Crystals of AtUGlyAH.** *A*, crystal-A was grown under 0.1 M phosphate citrate, pH 4.2, 10% (w/v) polyethylene glycol 3000, 0.2 M NaCl. *B*, crystal-B was grown under 0.1 M HEPES, pH 7.5, 7% (w/v) polyethylene glycol 8000, 8% (v/v) ethylene glycol.

3. Data Collection and Structure Determination

X-ray diffraction data were collected on beamline 4A and 6C at the Pohang Accelerator Laboratory, Korea. Both crystal-A and -B were cryoprotected by soaking each in crystallization solution containing additional 20-25% (v/v) glycerol and then flash-frozen in a liquid nitrogen stream at 100K. Initially, single-wavelength anomalous diffraction data were collected to 3.0 Å resolution on crystal-A at a wavelength of 0.97969 Å, corresponding to an absorption peak of selenium, and 2.5 Å resolution on crystal-B. Subsequently, data for the enzyme-substrate binary complex using crystal-B were collected to 2.6 Å resolution. Formation of the binary complex was achieved by a soaking experiment, in which crystal-B was soaked for 60 sec in a solution containing crystallization solution, 20% glycerol, and 3 mM (*S*)-ureidoglycine. To prepare (*S*)-ureidoglycine, 15 mM allantoate (Sigma) was incubated with AtAAH (2.5 μg/ul; 49.2 μM) and 1 mM $MnCl_2$ for 2 min (Fig. 4*A*). The reaction product was then diluted with crystallization solution containing 20% glycerol. The collected data were processed using HKL2000 (31). Crystal-A belongs to a $P3_221$ space group with a tetramer in an asymmetric unit, while crystal-B to a $P2_1$ space group with sixteen monomers in an asymmetric unit (Table 2).

Initially, the structure of crystal-A was determined using the SOLVE (32) and RESOLVE (33, 34) programs for phasing and density modification, respectively. Even with 3.0 Å resolution data, the initial electron density map was sufficient to trace most of the residues in each monomer. Molecular model building and refinement was performed using COOT (35) and CNS (36), respectively, and the PHENIX program (37) was used for molecular replacement. Of the four monomers in the crystal-A structure, one was manually modeled and then used as a search

model of molecular replacement for locating three other subunits. Further refinement of the tetrameric structure in crystal-A was carried out, and R_{work} and R_{free} values were 23.5 and 30.6%, respectively. Subsequently, higher resolution data were available using a crystal-B, and molecular replacement was conducted using the refined monomer in crystal-A as a search structure. In the 2.5 Å-resolution structure of crystal-B, a total of 12 monomers of the 16 in the asymmetric unit were initially found. At this stage, one particular monomer was manually modeled and again used as a search model for further molecular replacement. Eventually, all 16 monomers were found and automatically built in PHENIX. These 16 monomers were then manually inspected and refined to R_{work} and R_{free} values of 22.4 and 28.6%, respectively. For the AtUGlyAH in complex with (*S*)-ureidoglycine, the refined structure of the ligand-free form was used as a starting model for refinement. During refinement, an electron density corresponding to (*S*)-ureidoglycine was noticed in the metal-binding site, and its model was built and included in the structure for further refinement. The stereochemistry of the refined models was evaluated using a program MolProbity (38). Details of data collection and refinement are in Table 2. Structural analyses were carried out using programs in the CCP4 suite (39). The figures presented in this study were prepared using PyMOL (The PyMOL Molecular Graphics System).

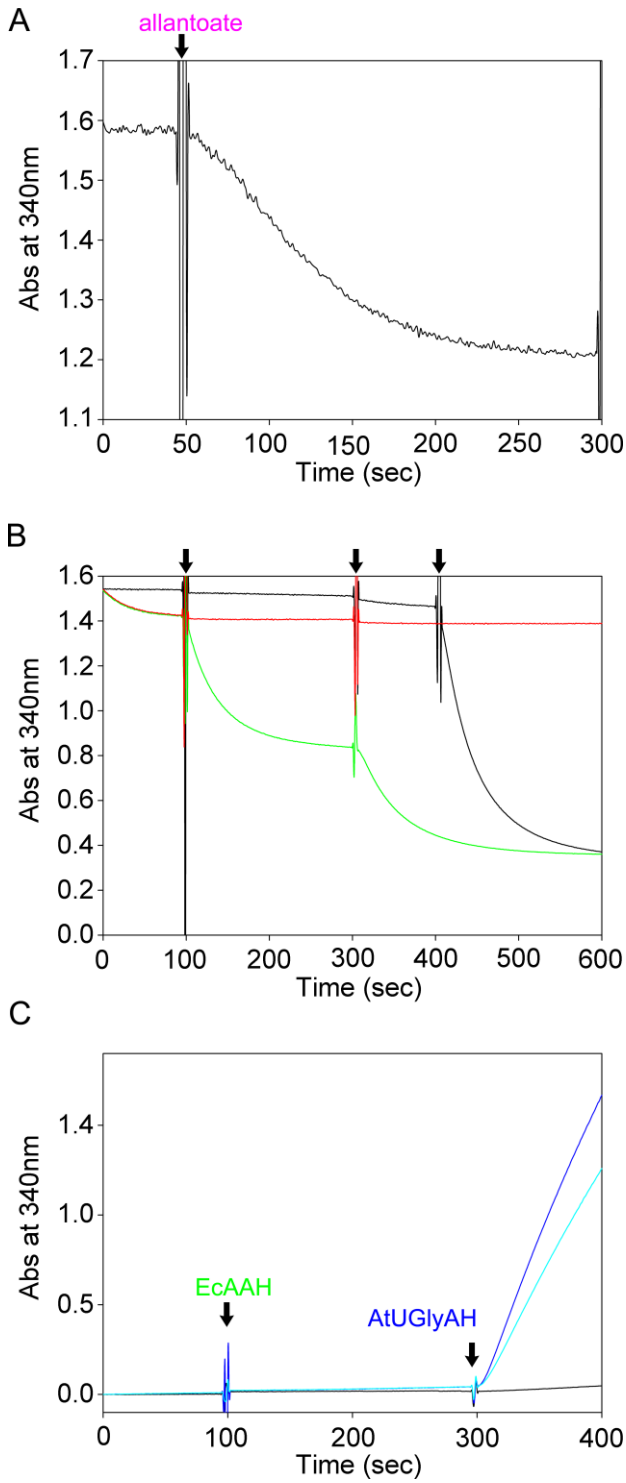

A
allantoate
Abs at 340nm
Time (sec)
B
Abs at 340nm
Time (sec)
C
EcAAH
AtUGlyAH
Abs at 340nm
Time (sec)

FIGURE 4. **Functional analysis of the enzymes in the ureide pathway.** *A,* (*S*)-ureidoglycine was produced by enzymatic reaction of AtAAH (2.1 μM) with 75 μM allantoate (Sigma). Ammonia (64 μM) was released during this reaction. *B,* (*S*)-ureidoglycine and (*S*)-ureidoglycolate were produced in sequential reactions of EcAAH (8.3 μM) and AtUGlyAH (0.23 μM) with allantoate prepared from allantoin. Reaction progress curves were recorded as follows: green (EcAAH, component present at 0 sec; 75 μM allantoate, component injected at 100 sec; AtUGlyAH, component injected at 300 sec; none, component injected at 400 sec), red (EcAAH; 75 μM (*S*)-allantoin; AtUGlyAH; none), black (without EcAAH; 75 μM allantoate; AtUGlyAH; EcAAH). *C,* the initial velocity of AtUGlyAH was measured by a coupled reaction with (*S*)-ureidoglycolate dehydrogenase. 6 mM allantoate generated from allantoin was converted into (*S*)-ureidoglycine by injecting EcAAH (0.4 mg/mL; 8.3 μM) at 100 sec. After incubation for 200 sec, AtUGlyAH (1.9 μg/mL; 58 nM) was added, resulting in an increase in absorbance at 340 nm. The progress curve was recorded in blue, while the cyan line shows a curve representing half of the AtUGlyAH concentration. The black curve was recorded in the absence of (*S*)-ureidoglycolate dehydrogenase from the pre-incubated reaction mixture.

TABLE 2. Data collection and refinement statistics for AtUGlyAH

	SeMet AtUGlyAH Crystal-A (Fig. 3*A*)	Ligand-free form Crystal-B (Fig. 3*B*)	Binary complex Crystal-B
PDB ID		4E2Q	4E2S
Data collection	SAD	*Single wavelength*	*Single wavelength*
Wavelength (Å)	0.97969	0.97973	1.23985
Space group	$P3_221$	$P2_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	147.4, 147.4, 112.4	93.9, 175.8, 155.2	93.1, 174.9, 154.3
α, β, γ (°)	90, 90, 120	90, 99.4, 90	90, 99.3, 90
Resolution (Å)	50-2.95(3.00-2.95) *	50-2.5(2.59-2.50)	50-2.6(2.69-2.60)
R_{merge} †	0.136 (0.821)	0.099 (0.564)	0.162 (0.513)
$I / \sigma I$	29.5 (4.9)	18.1 (2.6)	14.1 (2.3)
Completeness (%)	99.9 (100)	98.1 (94.6)	99.6 (99.5)
Redundancy	22.5 (21.6)	6.9 (6.3)	6.6 (6.1)
FOM	0.71		
Refinement			
Resolution (Å)		50-2.5	50-2.6
No. reflections		152,515	140,969
R_{work} ‡ / R_{free} ‡		22.4/28.6	20.9/27.1
Number of atoms			
Protein§		33,216	33,216
Mn^{2+} ion		16	16
(*S*)-Ureidoglycine		-	144
Water		719	598
B-factors			
Protein		41.4	40.6
Mn^{2+} ion		38.6	35.5
(*S*)-Ureidoglycine		-	48.4
Water		35.2	33.6
rmsd			
Bond lengths (Å)		0.00911	0.00850
Bond angles (°)		1.49	1.46
Ramachandran plot			
Favored (%)		92.1	93.3
Allowed (%)		7.0	5.9
Outlier (%)¶		0.9	0.9

*Values in parentheses are for the highest-resolution shell.

†$R_{merge} = \sum_{hkl}\sum_i|I_i-\langle I\rangle|/\sum_{hkl}\sum_i I_i$.

‡$R_{work} = \sum_{hkl}||F_{obs}|-k|F_{calc}||/\sum_{hkl}|F_{obs}|$. R_{free} was calculated using 10% of the data excluded from refinement.

§Ordered residues; Both ligand-free form and the binary complex contain Pro-39 to Leu-298, except for the disordered residues Thr-138 and Ser-139.

¶Outliers identified:

A total of 36 outliers in the ligand-free form, Ser-141, Ala-167 for subunit A, Ser-141 for subunit B, Pro-119 for subunit C, Ser-142 for subunit D, Pro-119 for subunit G, Pro-119 for subunit H, Ser-141, Gly-242 for subunit I, Pro-119 for subunit J, Pro-58, Ser-115, Ser-142, His-159, Glu-165, Leu-186 for subunit K, Ser-142, Asp-148, Ala-151, Gln-194 for subunit L, Pro-119, Ser-142 for subunit M, Pro-81, Asn-294 for subunit N, Pro-58, Val-147, Asp-148, Glu-165 for subunit O, Ile-40, Asn-87, Met-112, Ser-115, Ser-141, Thr-146, Pro-154, Pro-211 for subunit P.

A total of 35 outliers in the binary complex, Ser-141, Gly-242 for subunit A, Pro-96, Ser-141 for subunit B, Glu-75, Pro-96, Ser-141 for subunit C, Pro-119, Gly-180, Asn-294 for subunit D, Glu-75, Pro-96 for subunit E, Glu-235 for subunit F, Ser-141, Glu-165, Ala-167, Gly-180 for subunit G, Ser-141 for subunit H, Ser-142, Phe-157 for subunit J, Ser-141, Glu-165, His-182, Thr-191, Val-214 for subunit K, Ser-141 for subunit L, Ser-141, Glu-235, Gly-242 for subunit M, Ser-141 for subunit N, Pro-74, Asp-148, Gly-242 for subunit O, Pro-154, Phe-157 for subunit P.

4. Enzyme Assay

For functional analysis, all enzymes used in this study were prepared as described above, with the exception of allantoinase. For expression, *E. coli* BL21 (DE3) (Merck) harboring the plasmid of interest was cultured in Luria-Bertani medium. Subsequently, the N-terminal His-tagged protein was purified by immobilized metal affinity chromatography and dialyzed against buffer (50mM Tris-HCl, pH 7.4, and 200 mM NaCl). For allantoinase, functionally active enzyme was expressed in Terrific-Broth medium with 10% 4-morpholinepropanesulfonic acid, 1 mM $CoCl_2$ and 50 μM of isopropyl β-D-1-thiogalactopyranoside. The His-tagged protein was purified by increasing imidazole concentration in buffer (50 mM HEPES, pH 7.4, 50 mM KCl). The purified enzyme was dialyzed against buffer (50 mM HEPES, pH 7.4) (19).

Enzyme assays were performed at 30°C using a UV-visible spectrophotometer (V-560; Jasco) equipped with a cuvette holder and connected to a temperature-controlling water circulator. Two different assay protocols were employed: one for monitoring the stepwise release of ammonia during the formation of (*S*)-ureidoglycolate, and the other for measuring the initial velocity of the formation of (*S*)-ureidoglycolate in a coupled reaction with an NAD-dependent (*S*)-ureidoglycolate dehydrogenase reaction.

In the first method, the release of ammonia was monitored in a coupled assay with an NADPH-dependent glutamate dehydrogenase in the presence of α-ketoglutarate by measuring the decrease in absorbance at 340 nm due to the oxidation of NADPH (14, 40). This method was performed for various control experiments and used to verify the substrate specificity of each purified enzyme (Fig. 4, *A* and *B*).

The reaction mixture (2 mL) contained 0.2 M Tris-HCl buffer (pH 8.5), 0.3 mM NADPH, 2.5 mM α-ketoglutarate, 10 units of glutamate dehydrogenase from *Proteus* sp. (Sigma), 100 μM $MnCl_2$, and AtAAH (2.1 μM) or EcAAH (8.3 μM). The reaction was then initiated by addition of 75 μM allantoate purchased from Sigma or generated from allantoin, followed by addition of AtUGlyAH (0.0581–0.232 μM). The reaction progress curve proved that the enzymes used in this assay were functionally active and exhibited high substrate specificity: allantoin is not a substrate of either AAH or AtUGlyAH (red curve in Fig. 4*B*), and allantoate acts as a substrate for AAH, but not AtUGlyAH (black curve in Fig. 4*B*). In particular, the concentration of ammonia released during production of (*S*)-ureidoglycine was calculated to be about 93 μM, while the release of an additional 73 μM of ammonia by AtUGlyAH was detected (green curve in Fig. 4*B*). Therefore, under these experimental conditions we found that allantoate was completely converted into (*S*)-ureidoglycolate.

Measurements of the steady-state kinetic parameters of AtUGlyAH wild-type (WT) and mutant enzymes were performed using the second method, mainly because the first protocol was impractical at high substrate concentrations. Specifically, in the protocol, formation of (*S*)-ureidoglycolate by AtUGlyAH was monitored in a coupled assay with (*S*)-ureidoglycolate dehydrogenase, in which the (*S*)-ureidoglycolate produced is subsequently converted into oxalurate, with concurrent reduction of NAD^+ into NADH accompanying an increase in absorbance at 340 nm (25). The following requirements were verified for reliable kinetic measurements: under the experimental conditions outlined below, the increase in absorbance at 340 nm was dependent only on the concentrations of (*S*)-ureidoglycine and AtUGlyAH, but not of other enzymes (Fig. 4*C*); and allantoin

was completely converted into (*S*)-ureidoglycine (Fig. 4*B*). Initially, a 50 mM allantoin racemic mixture (i.e., 25 mM (*S*)-allantoin) was dissolved in water at 37℃ for one hour and cooled to 30℃, and then diluted with enzyme reaction buffer (50 mM HEPES, pH 7.5) to a final concentration of 20 mM (*S*)-allantoin. The reaction was initiated by addition of allantoinase (0.525 mg/mL; 10 μM) at 30℃ for one hour to produce 20 mM allantoate. This allantoate solution was stored at -70℃ and used without further purification. The assay mixture solution (2 mL) contained 50 mM HEPES buffer (pH 7.5), 100 μM $MnCl_2$, (*S*)-ureidoglycolate dehydrogenase (0.2 mg/mL; 5.0 μM), 3 mM NAD^+, and allantoate at the concentrations indicated (0.2–16 mM). After incubation for 60 sec, further conversion of allantoate into (*S*)-ureidoglycine was catalyzed by addition of EcAAH (0.4 mg/mL; 8.3 μM), and (*S*)-ureidoglycolate formation was subsequently initiated by the addition of AtUGlyAH (3.9 nM–0.15 μM) (Fig. 4*C*). Since (*S*)-ureidoglycine is chemically labile, the time period required for complete conversion of allantoate to (*S*)-ureidoglycine was determined experimentally by measuring the initial velocity of AtUGlyAH as a function of incubation time. Typically, 60 sec of incubation was the optimal time for complete formation of (*S*)-ureidoglycine under the experimental conditions. Shorter or longer incubations caused a lower velocity, which indicates that either the conversion was incomplete or chemically labile (*S*)-ureidoglycine underwent spontaneous degradation, respectively. The initial velocity was determined by measuring the linear increase in absorbance at 340 nm for the first 20 or 30 sec, and was calculated as the NADH concentration produced per min, with a molar extinction coefficient of 6,220 $M^{-1}cm^{-1}$ at 340 nm. The K_m and V_{max} values were obtained using SigmaPlot (Systat

Software), and k_{cat} values were computed by dividing V_{max} by the AtUGlyAH concentration used.

5. Metal Analysis

For metal analysis, the WT enzyme, and E235Q and E235A mutants were expressed and purified as described in a section of Enzyme Assay. Specifically, the N-terminal His-tagged protein was purified by immobilized metal affinity chromatography and the His-tag was subsequently removed by treatment of *Tobacco etch virus* protease. The His-tag free enzyme was then incubated in the presence of 5 mM EDTA for an hour at 4℃ and was purified by gel filtration chromatography. After addition of 1 mM $MnCl_2$, the resulting enzyme solution was again subjected to gel filtration chromatography with buffer (50 mM HEPES, pH7.4, and 100 mM NaCl). Protein concentrations were about 0.11~0.14 mM and the metal content of the enzyme was determined by inductively coupled plasma-atomic emission spectrophotometer.

RESULTS

1. Molecular Architecture of Octameric AtUGlyAH in a Ligand-Free From

The crystal structure of AtUGlyAH was determined using AtUGlyAH$^{\Delta 35}$, which was missing 35 residues from the N-terminus including a cleavable signal peptide (Fig. 5 and Table 2). Two octamers were characterized with no non-crystallographic symmetry in an asymmetric unit of the $P2_1$ space group. The apparent molecular mass of AtUGlyAH (30.2 kDa monomer) was estimated by size-exclusion chromatography to be more than 200 kDa (Fig. 6), suggesting that the functional unit of AtUGlyAH is likely an octamer.

The crystal structure of monomeric AtUGlyAH, which contains the ordered residues Pro-39 to Leu-298, is composed of 19 β-strands and 4 short 3_{10}-helices (Fig. 5). Its structure is divided into two structurally similar β-barrel domains, called N- and C-domains, with an N-terminal segment protruding from the rest of the structure (Fig. 7*A*). Each domain represents a typical cupin fold of a β-barrel, in which two layers of β-sheet are packed in a parallel manner with a funnel-like structure (41). Specifically, the N-domain includes β3 to β11. Six strands are arranged in the spatial order β3, β4, β10, β5, β8, and β11, with an anti-parallel orientation. They form a central face of a β-barrel fold, whereas the outer face of the fold consists of β9, β6, and β7, with their orientations being parallel to the central β-sheet. The C-domain also exhibits a similar architecture, in which N-terminal anti-parallel β1 and β2 strands form part of the central anti-parallel β-sheet that comprises β1, β2, β17, β14, β19, β13, and β12 in a spatial order. Three strands, β16, β15, and β18, line along the outer surface of the β-barrel fold. Therefore, the AtUGlyAH monomer is a bi-cupin molecule and is stabilized by packing the

central β-sheets from each domain in a parallel manner (Fig. 7*A*). These interactions result in 2830 Å^2 of buried surface area and a root mean-square deviation (rmsd) of 1.9 Å for the structurally equivalent 91 Cα atoms between the two cupin folds.

The AtUGlyAH functional unit is assembled from two layers of tetramers, each layer with a dorsal and a ventral surface (Fig. 7*B*). Within a layer, four monomers are related by a non-crystallographic four-fold symmetry along the vertical axis running through the center of the octamer, as shown by the top view of a dorsal surface. These four monomers are structurally identical within 0.34 - 0.42 Å rmsd for 258 Cα atoms. The two layers are stabilized by interacting through the ventral surface, but are offset from the vertical axis by ~22.5° of rotation (Fig. 7*B*). Monomers in different layers are related by a two-fold symmetry along the two horizontal axes, which are orthogonal to each other. Therefore, a *422*-symmetry along a vertical axis and two horizontal axes, respectively, fully describes the molecular architecture of octameric AtUGlyAH using a monomer as a protomer.

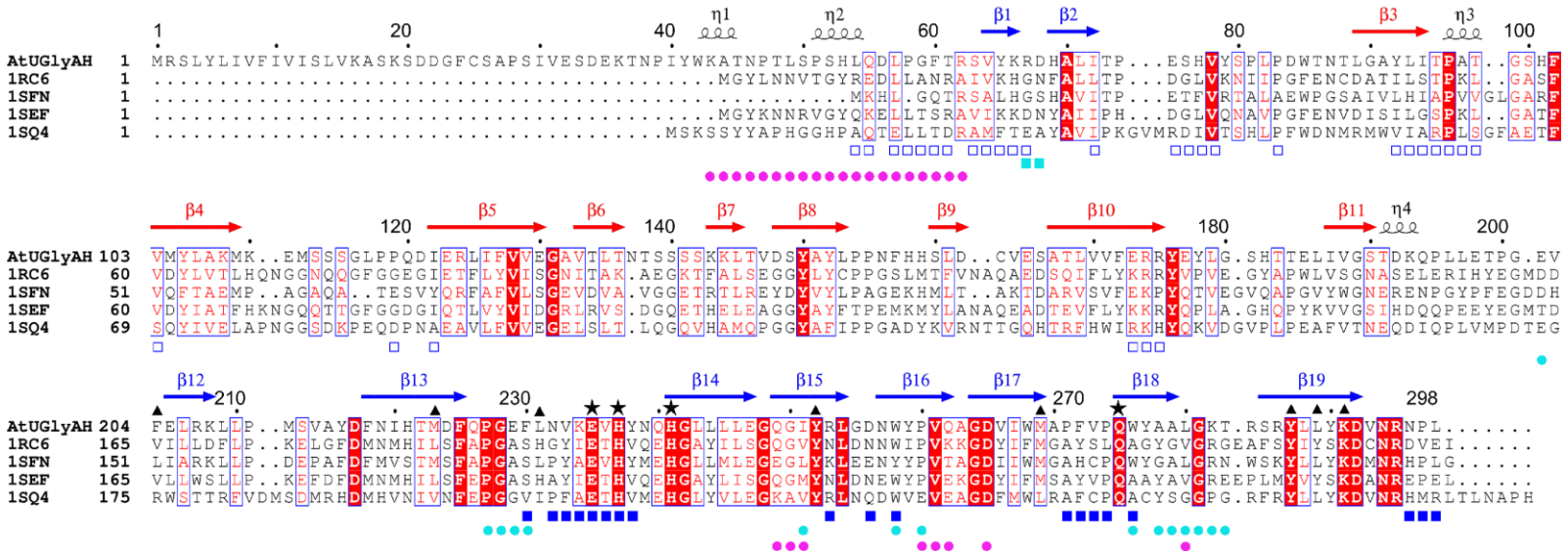

AtUGlyAH
1RC6
1SFN
1SEF
1SQ4

FIGURE 5. **Sequence alignment of AtUGlyAH.** The amino acid sequences of AtUGlyAH, including the signal peptide (residues 1–20) (14), are compared with those of the *E. coli* ortholog and structural homologs from other bacteria, all with known structures in the absence of their functional roles. Structures of these proteins were previously determined by Structural Genomics Centers: an *E. coli* ortholog (PDB ID 1RC6; resolution, 2.6 Å; Z-score, 29.6; rmsd, 2.1 Å; sequence identities, 28%), and structural homologs from *Deinococcus radiodurans* R1 (1SFN;2.46 Å;34.8;1.4 Å;40%), *Enterococcus Faecalis* V583 (1SEF;2.05 Å;32.3;1.7 Å;31%), and *Pseudomonas aeruginosa* (1SQ4;2.7 Å;28.3;2.5 Å;26%). Highly conserved residues are shown in red and boxed in blue. Strictly conserved residues are shown on a red background. The secondary structural elements defined in a ligand-free form are shown for the corresponding AtUGlyAH sequences, with the N- and C-domains in red and blue, respectively. Black asterisks and triangles represent residues involved in the Mn^{2+} coordination shell and the (*S*)-ureidoglycine-binding site, respectively. Residues involved in the interactions between monomers within the layer are indicated by blue rectangles; open and closed ones represent residues from different monomers. Cyan circles indicate residues involved in inter-layer interactions between monomers located vertically and magenta circles indicate those located diagonally across the layers. This figure was prepared using ESPript (42).

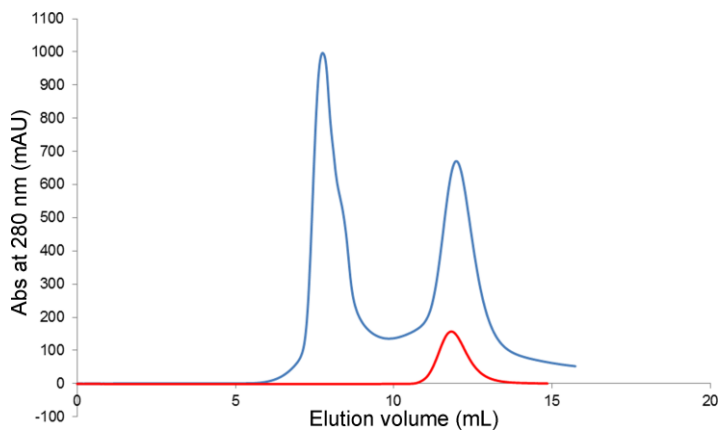


FIGURE 6. **Size-exclusion chromatography of AtUGlyAH.** Elution profile of size-exclusion chromatography is shown for AtUGlyAH (red). Blue dextran (2000 kDa) and β-amylase (200 kDa) were used as size marker (blue).

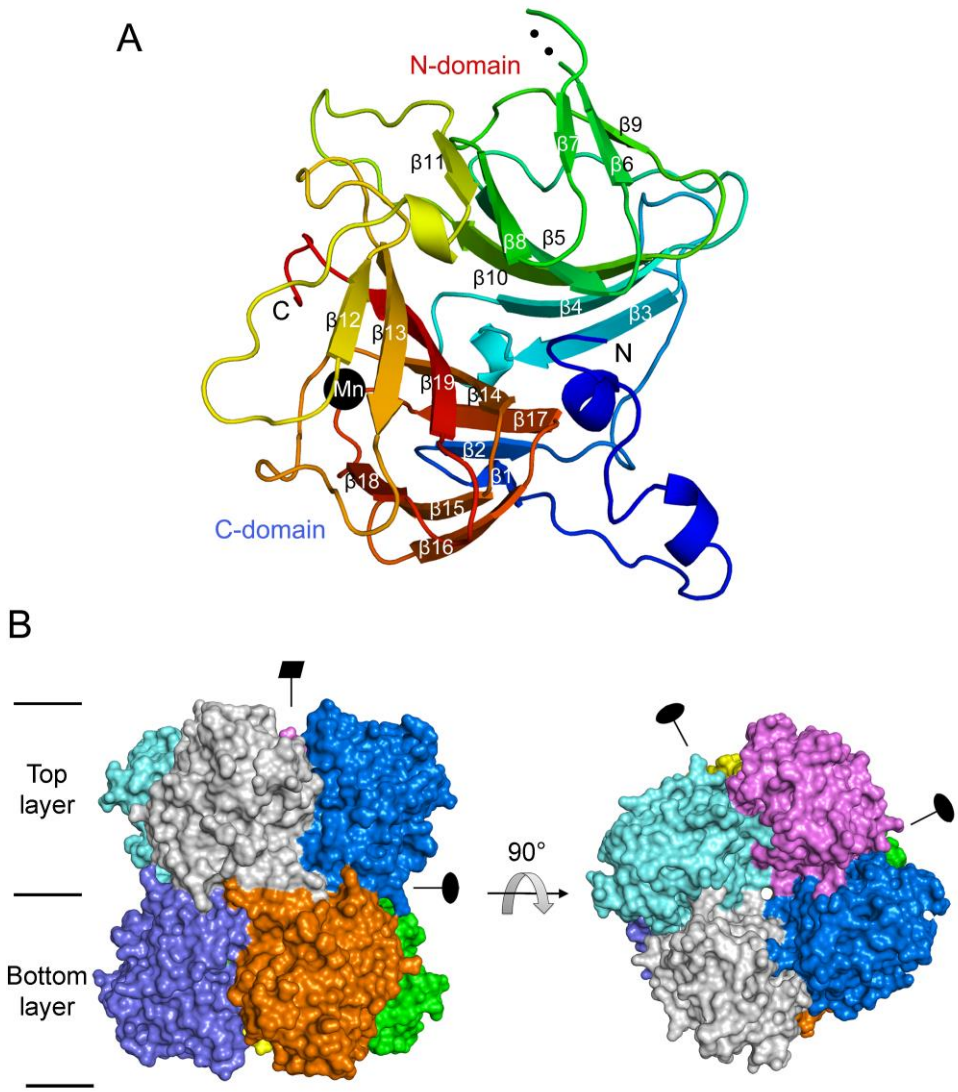


FIGURE 7. **Overall structure of AtUGlyAH.** *A,* monomer structure, with an orientation identical to that of the monomer in grey in the left panel of *B*. Presence of an Mn^{2+} ion is indicated by a black circle. Two residues, Thr-138 and Ser-139, are disordered in the loop between β6 and β7. *B,* surface representation of an octamer in two orientations. A side view, with each monomer in a different color, is presented in the left panel, while the right panel shows a top view of the layer.

2. Interactions between Monomers in the Intra- and Inter-layers

Within each layer, monomeric AtUGlyAH is packed against its neighbors, whose positions differ by 90° of rotation via hydrophobic interactions. These interactions are asymmetrical in that distinct regions from neighboring monomers participate in the packing. Specifically, residues in the loop regions of the protruding N-terminal segment and those connecting β2 and β3, and β3 and β4 mediate largely hydrophobic interactions with other loop residues, mainly in the C-domain β13*–β14* and β17*– β18* of the adjacent monomer (hereafter, an asterisk denotes an element or a residue of a neighboring monomer) (Fig. 8*A*). In addition, the C-terminal residues Asn-296* to Leu-298* are plugged into the concave surface formed by the residues in the vicinity of the N-domains β3 and β4. Several hydrogen bonds were characterized in this packing within a layer (Fig. 8*A*). A four-fold arrangement of monomers in each layer generates four independent, but structurally identical, interfaces; each interface exhibits a buried surface area of 1069 $Å^2$ corresponding to ~8.8 % of the surface area of the monomer.

More extensive, diverse packings were identified in the interactions between the two layers, with a total buried surface area of 4677 $Å^2$. The inter-layer interactions are symmetrical, in which structurally and sequentially identical residues are involved in the packing of monomers in different layers. Two types of interactions are observed (Fig. 7*B*): one between monomers located vertically in different layers and the other between monomers located diagonally across the layers. Specifically, residues in the outer surface of the C-domain cupin fold, including β16, β15, and β18, are major structural elements in stabilizing monomers orientated vertically in different layers (Fig. 8*B*). In the other type of interaction, the protruding N-terminal region containing Pro-39 to Ser-63, as well as residues on the outer surface of the

C-domain cupin fold in the vicinity of β16 and β15, are packed in a two-fold symmetrical manner against the equivalent elements of a monomer in a different layer (Fig. 8*C*). Overall, these interactions serve as key structural features in stabilizing the two layers, in particular by inserting the protruding N-terminal segment into the concave surface of another monomer.

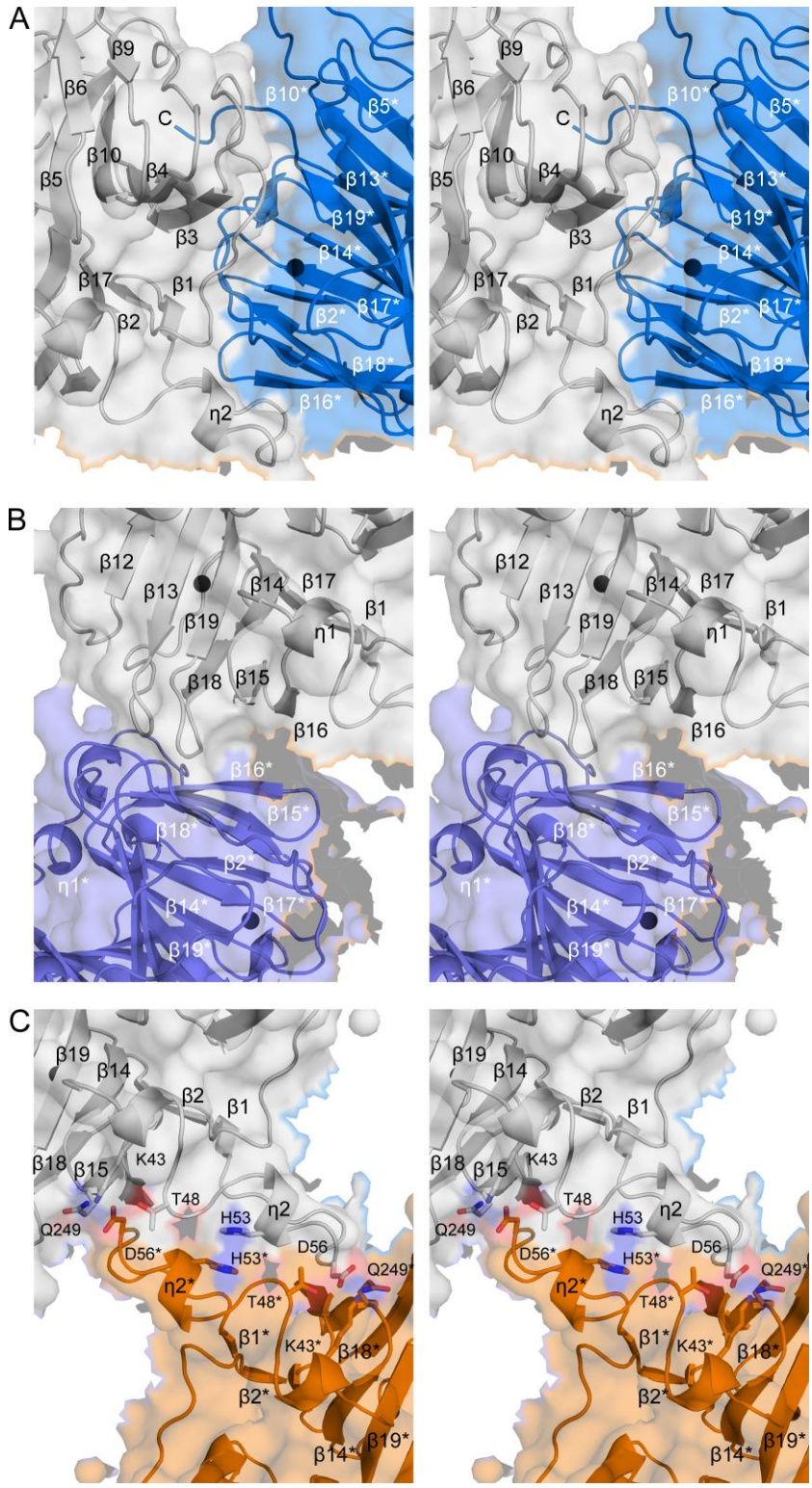

A
β9
β6
C
β10*
β5*
β10
β4
β5
β13*
β19*
β3
β14*
β17
β1
β2*
β17*
β2
β18*
β16*
η2
B
β12
β13
β14
β17
β19
β1
η1
β18
β15
β16
β16*
β15*
β18*
β2*
η1*
β14*
β17*
β19*
C
β19
β14
β2
β1
β18
β15
K43
T48
H53
η2
Q249
D56*
H53*
D56
Q249*
η2*
T48*
β1*
K43*
β18*
β2*
β14*
β19*

FIGURE 8. **Interactions of monomers within and between layers.** Interactions observed within and between layers are represented in stereoscopic view. Color code and orientation for each monomer in these figures are identical to those in the left panel of Fig. 7*B*, and residues involved in these interactions are indicated with symbols as in Fig. 4. *A*, interactions of two neighboring monomers within the layer. The carbonyl oxygens of Gly-59 and Tyr-65, and the side chain of Ser-76 interact with the side chains of Arg-253* and Arg-67*, and the carbonyl oxygen atom of Lys-234*, respectively. *B*, two monomers located vertically in different layers are presented. Residues, such as Phe-230, Ile-251, Trp-258, Trp-276, Ala-278, Ala-279, and Leu-280, and equivalent residues from another monomer are clustered into this region to mediate hydrophobic interactions, as well as hydrogen bonding between the main chain Gly-228 and Gly-281*. *C*, interactions occurring between two monomers located diagonally across the layers are shown. In particular, Lys-43, Thr-48, and Gln-249 participate in hydrogen bonding with Asp-56*, whereas His-53 in the middle of the protruding region mediates a stacking interaction with the equivalent His-53* of a different monomer.

3. Active Site in the Ligand-Free Form of AtUGlyAH

In the monomer, each N- and C-domain with a funnel-like β-barrel structure exhibits two possible openings, one at each end of the long axis of the funnel (Fig. 7*A*). However, various interactions in monomers and octamers effectively seal off these openings, except for one end in the C-domain. Specifically, a smaller opening near β3 and β4 in the N-domain is occluded by the C-terminal segment from the adjacent monomer within the layer (Fig. 8*A*). The entrance to a larger opening in the N-domain is closed off by several layers of hydrophobic residues, whose side chains point into the cavity, from the central and the outer side β-strands in a cupin fold (Fig. 9). In contrast, the inside cavity of the C-domain is readily accessible from the surface of the enzyme through the larger opening, with its location in the direction of the dorsal surface in the layer (Fig. 7*A*). A smaller end of the C-domain funnel points toward the center of the octamer, but is sealed off by the presence of β1 and β2 (Fig. 7*A*).

Consistent with the requirement of Mn^{2+} ion for catalytic activity (14), the location of the active site in AtUGlyAH was indicated by the presence of a metal binding site in the C-domain. Indeed, crystallization of AtUGlyAH was performed with 1 mM $MnCl_2$. The metal binding site is embedded in the middle of a funnel-like cavity in the C-domain, which is located at 14 Å-deep from the surface (Fig. 10*A*). Four residues from the central side and the outer side of a β-barrel are involved in the metal coordination shell with one or two water molecules, depending on the subunit (Fig. 10*B*). Specifically, Glu-235 and His-237 in the long loop between β13 and β14, which transverses and folds back between the central and outer side of the β-barrel, as well as His-241 in the central β14 and Gln-275 in the outer side β18, are part of an octahedral coordination with the bound metal. In

this coordination, the metal ion lies in the center of an equatorial plane formed by Glu-235, His-237, His-241, and a water molecule across from His-237 (Fig. 10*B* and Table 3*A*). Gln-275 and another water molecule *trans* to Gln-275 serve as the axial ligands, which are normal to the equatorial plane; the electron density for the axial water molecule is relatively disordered in different monomers. Under the crystallization condition, the metal ion was identified as Mn^{2+} ion by inductively coupled plasma-atomic emission spectroscopy (Table 4). It is noticeable that two water molecules ligating with the Mn^{2+} ion are near the entrance, and therefore their locations likely serve as the binding site for the incoming (*S*)-ureidoglycine substrate molecule.

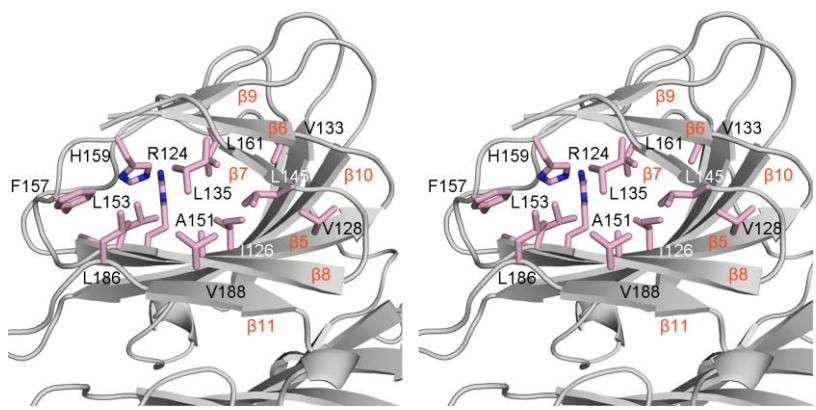


FIGURE 9. **The closed opening in the N-domain of AtUGlyAH.** Hydrophobic residues effectively seal off the possible opening in the N-domain. Residues involved in these features are β11 (Leu-186, Val-188), β8 (Ala-151, Leu-153), and β5 (Ile-122, Arg-124, Ile-126, Val-128) from the central β-strands, and β7 (Leu-145), β6 (Val-133, Leu-135) and β9 (Phe-157, His-159, Leu-161) from the outer β-strands.

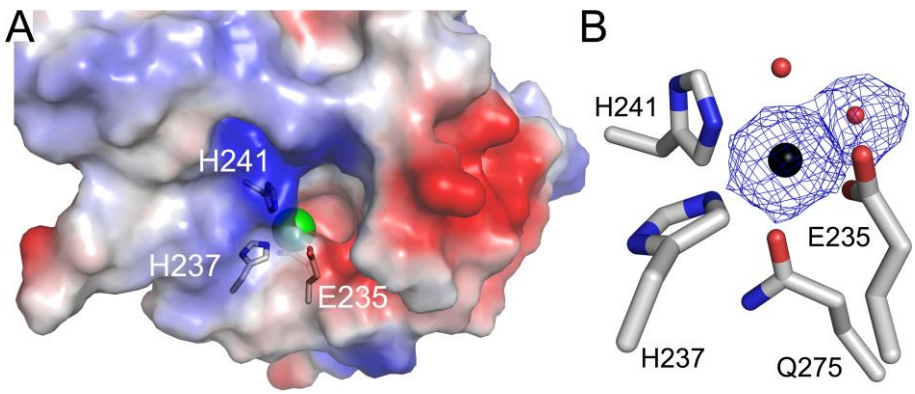
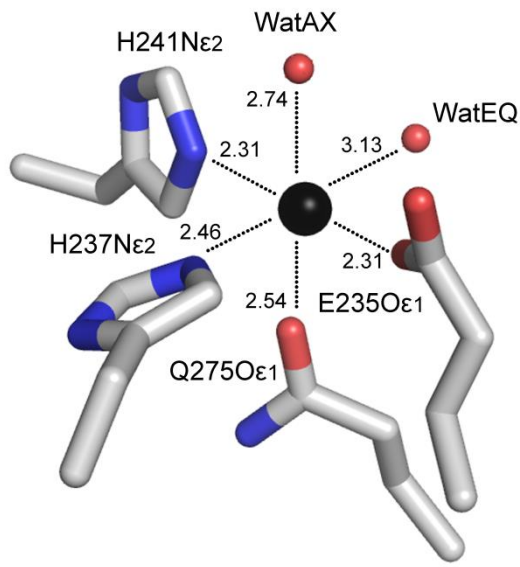


FIGURE 10. **Active site in the C-domain cupin fold of the ligand-free form AtUGlyAH.** *A*, surface representation of the opening in the C-domain of monomer is shown. In this orientation, the N-terminal protruding region is located on the back side. The active site residues are indicated with the bound Mn^{2+} ion in green. *B*, the Mn^{2+} binding residues, with the Mn^{2+} ion in a black sphere and water molecules indicated by red circles. The Mn^{2+} ion and water molecule *trans* to His-237 are overlaid with a $2F_o–F_c$ electron density map contoured at 3.0σ and 1.0σ, respectively. The axial water molecule across from Gln-275 is present in some monomers, and its density is relatively disordered, but obvious, at 0.8 σ of the $2F_o–F_c$ electron density map. Stereochemical details of the Mn^{2+} binding site are shown in Table 3.

TABLE 3. Geometry of the metal coordination in (A) the ligand-free form and (B) the complex

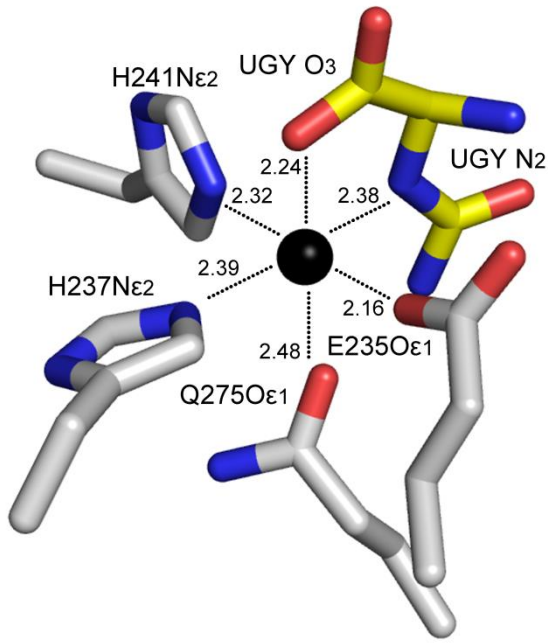
(A) Ligand-free form



Defining Atoms	Observed Angles (degrees)
H241N$^{\varepsilon 2}$-Mn-H237N$^{\varepsilon 2}$	87.5 (6.12) [a]
H237N$^{\varepsilon 2}$-Mn-E235O$^{\varepsilon 1}$	83.1 (13.8)
E235O$^{\varepsilon 1}$-Mn-WatEQ	93.0 (15.4)
WatEQ-Mn-H241N$^{\varepsilon 2}$	96.5 (5.35)
H241N$^{\varepsilon 2}$-Mn-WatAX	88.4 (9.92)
H237N$^{\varepsilon 2}$-Mn-WatAX	97.8 (6.04)
E235O$^{\varepsilon 1}$-Mn-WatAX	87.6 (10.1)
WatEQ-Mn-WatAX[b]	94.2 (10.1)
H241N$^{\varepsilon 2}$-Mn-Q275O$^{\varepsilon 1}$	90.4 (5.24)
H237N$^{\varepsilon 2}$-Mn-Q275O$^{\varepsilon 1}$	90.4 (9.83)
E235O$^{\varepsilon 2}$-Mn-Q275O$^{\varepsilon 1}$	99.5 (14.4)
WatEQ-Mn-Q275O$^{\varepsilon 1}$	81.7 (13.6)
WatAX-Mn-Q275O$^{\varepsilon 1}$	169.3 (3.807)
H241N$^{\varepsilon 2}$-Mn-E235O$^{\varepsilon 1}$	159.5 (11.12)
H237N$^{\varepsilon 2}$-Mn-WatEQ	165.8 (7.339)

[a]Values in parentheses are deviation.
[b]WatAX and WatEQ represent water molecules located in the axial and equatorial position, respectively.

(B) Binary complex



Defining Atoms	Observed Angles (degrees)
H241N$^{\varepsilon 2}$-Mn-H237N$^{\varepsilon 2}$	88.4 (6.26)
H237N$^{\varepsilon 2}$-Mn-E235O$^{\varepsilon 1}$	93.3 (10.4)
E235O$^{\varepsilon 1}$-Mn-UGY N2	81.3 (9.03)
UGY N2-Mn-H241N$^{\varepsilon 2}$	96.0 (11.0)
H241N$^{\varepsilon 2}$-Mn-UGY O3	81.2 (8.25)
H237N$^{\varepsilon 2}$-Mn-UGY O3	94.3 (6.34)
E235O$^{\varepsilon 1}$-Mn-UGY O3	94.1 (8.34)
UGY N2-Mn-UGY O3[a]	74.1 (3.01)
H241N$^{\varepsilon 2}$-Mn-Q275O$^{\varepsilon 1}$	91.0 (7.95)
H237N$^{\varepsilon 2}$-Mn-Q275O$^{\varepsilon 1}$	95.6 (6.29)
E235O$^{\varepsilon 1}$-Mn-Q275O$^{\varepsilon 1}$	92.8 (8.67)
UGY N2-Mn-Q275O$^{\varepsilon 1}$	96.0 (4.76)
UGY O3-Mn-Q275O$^{\varepsilon 1}$	165.1 (6.869)
H241N$^{\varepsilon 2}$-Mn-E235O$^{\varepsilon 1}$	167.6 (6.081)
H237N$^{\varepsilon 2}$-Mn-UGY N2	164.1 (4.856)

[a]UGY O3 and UGY N2 correspond to WatAX and WatEQ in the ligand-free form, respectively.

TABLE 4. Metal analysis of the WT Enzyme[a], and E235A and E235Q mutants

	Mn^{2+} (ppm)	Co^{2+} (ppm)	Mg^{2+} (ppm)	Ni^{2+} (ppm)	No. of Mn^{2+} ion per subunit
WT(–)[b]	–[c]	–	–	–	–
WT(+)	0.918	–	–	–	0.13
E235A	–	–	–	–	–
E235Q	0.994	–	–	–	0.13

[a]In this experiment, we found that concentration of Mn^{2+} ion is about 10 to 200-fold higher than those of other divalent cations (Co^{2+}, Mg^{2+}, Ni^{2+}), and number of Mn^{2+} ion is about 0.1 per subunit of the enzyme. Therefore, the binding affinity of Mn^{2+} ion to the enzyme is relatively low under the experiment conditions.
[b]For the control experiment, additional 1 mM $MnCl_2$ was not included in this WT enzyme (pre-treated with 5 mM EDTA).
[c]Values under 0.200 ppm were determined in these measurements, which is the detection limit of the instrument.

4. Binding Mode of (*S*)-Ureidoglycine in the Binary Complex

The structure of AtUGlyAH in complex with (*S*)-ureidoglycine was determined in a soaking experiment. (*S*)-ureidoglycine, which is commercially unavailable, was produced by an enzyme reaction using AAH with allantoate as the substrate (Figs. 1 and 4*A*). By using the X-ray data from a crystal of AtUGlyAH soaked with (*S*)-ureidoglycine, an extra electron density was identified in the proximity of the bound Mn^{2+} ion (Fig. 11). The density was assigned as that of (*S*)-ureidoglycine, given that AAH produces (*S*)-ureidoglycine in a stereospecific manner (14) and only the (*S*)-configuration of ureidoglycine fits the electron density identified in the binary complex. The product of UGlyAH, ureidoglycoate (Fig. 1) was also modeled into the observed density; however, only the (*R*)-configuration fits the density, which is inconsistent with the previous assignment of absolute stereochemistry of the ureidoglycolate enantiomer produced by UGlyAH (43). Therefore, it is likely that low catalytic activity of crystalline AtUGlyAH relative to that in solution, as well as the cryogenic conditions during X-ray data collection, stabilizes chemically labile (*S*)-ureidoglycine, rather than catalyzing the formation of (*S*)-ureidoglycolate.

There were no significant changes in conformation between the ligand-free form and the binary complex, with an rmsd of less than 0.5 Å and 0.32 Å for 258 and 2064 Cα atoms in the monomer and octamer, respectively. When bound to the Mn^{2+} coordination shell, (*S*)-ureidoglycine is in an almost planar conformation (Figs. 11 and 12); the carboxyl and the ureido groups are in an *anti* conformation along the bond between the chiral carbon atom and the nitrogen atom in the amide bond, with a dihedral angle of 162° (± 16.0) throughout the 16 monomers. The amino group and hydrogen attached to the chiral carbon atom are the only ones out of the

plane. The substrate (*S*)-ureidoglycine is almost parallel to the axial plane of the Mn^{2+} coordination shell (Fig. 12); the ureido group is located at the innermost part of the active site, while the carboxylate group points toward an entrance to the active site. Specifically, the binding of (*S*)-ureidoglycine is achieved by replacing the two water molecules (one each for the axial and equatorial ligands in the Mn^{2+} coordination shell of the ligand-free form). One of the oxygen atoms in the carboxylate group and the nitrogen atom in the amide bond then take the positions of the axial and equatorial ligands, respectively (Fig. 12), without distorting the geometry of the octahedral coordination with the bound metal (Table 3*B*).

The binding of (*S*)-ureidoglycine in the complex is further stabilized by hydrogen bonding along the molecule (Figs. 12 and 13). Specifically, the carboxyl group of (*S*)-ureidoglycine mediates several interactions within 3.0 Å, i.e., a possible charge neutralization with Lys-291 and hydrogen bonding with an equatorial ligand His-241. The side chain of an equatorial ligand Glu-235 also stabilizes the binding of the amino group within a distance of 2.8 Å and the nitrogen atom in the amide bond within a distance of 3.0 Å, allowing the (*S*)-configuration of ureidoglycine. The ureido group is also within a hydrogen bonding distance of 2.8-3.0 Å from Tyr-252, Tyr-287, and an axial ligand Gln-275, facilitating the binding of the ureido group in the innermost part of the active site. In addition, hydrophobic residues comprise a part of the active site within 5.0 Å from the bound substrate, and position the incoming (*S*)-ureidoglycine into the productive binding mode for catalysis. These residues include Phe-204, Met-223, and Leu-231 near the ureido group, and Met-269 and Leu-289 near the amide bond (Figs. 12 and 13). It is noteworthy that the Mn^{2+}-binding residues and those hydrophilic residues near the (*S*)-ureidoglycine binding site are invariant among UGlyAHs and its structural

homologs (Fig. 5), suggesting the functional roles of these residues in catalysis. In particular, His-221, which is conserved in AtUGlyAH and its *E. coli* ortholog (Fig. 5; PDB ID 1RC6), is located on the side of the hydrogen attached to the chiral carbon atom, 5.5 Å from the chiral carbon atom.

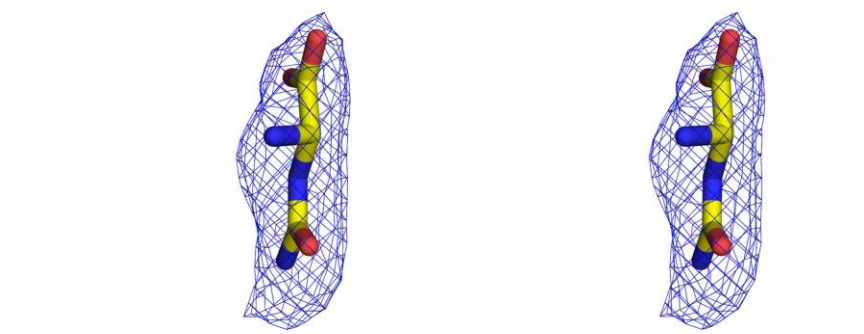


FIGURE 11. **The electron density of (*S*)-ureidoglycine in the binary complex.** Stereo view of the F_o–F_c electron density map contoured at 2.5 σ is shown with the modeled (*S*)-ureidoglycine in the binary complex.

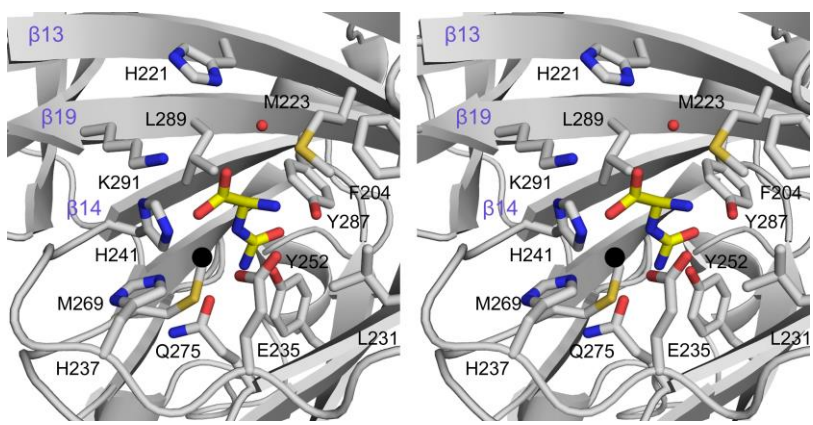


FIGURE 12. **The binding mode of (*S*)-ureidoglycine in the binary complex.** The active site in the binary complex is shown in a stereoscopic view, with neighboring residues less than 5.0 Å from the bound (*S*)-ureidoglycine.

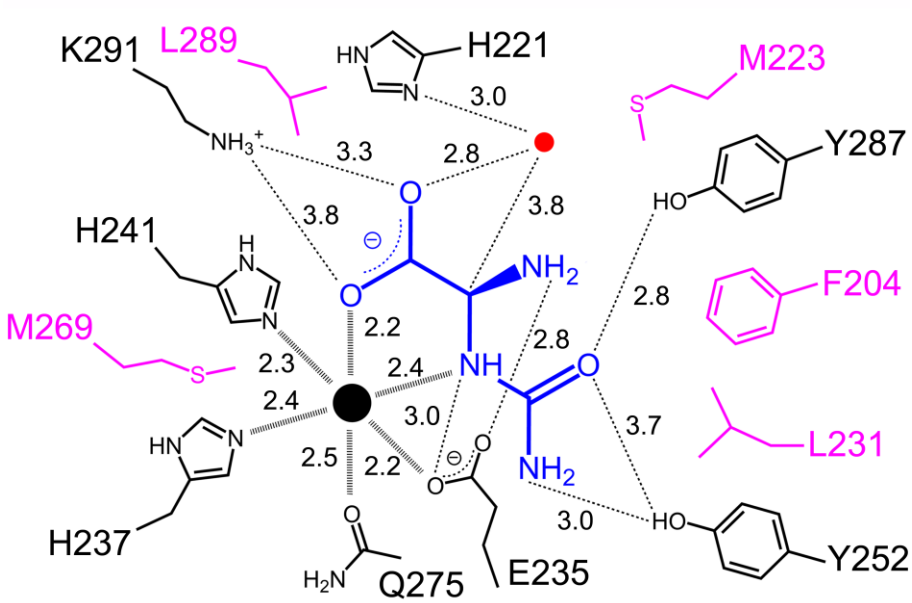


FIGURE 13. **Schematic representation of the interactions around the Mn^{2+} ion and (*S*)-ureidoglycine in the binary complex.** The active site in the binary complex is shown in a schematic view. Residues involved in hydrophobic interactions and hydrogen bonding are indicated in magenta and black, respectively, with the interatomic distance values averaged among 16 monomers. For clarity, the following possible hydrogen bonds were not indicated: the ring nitrogen of His-241 to the carboxyl group of (*S*)-ureidoglycine within 3.2 Å, the side chain carbonyl oxygen of Gln-275 to the amino group of the ureido group within 2.8 Å, and the two hydroxyl groups in Tyr-252 and Tyr-287 within 3.7 Å.

5. Functional Analysis

(*S*)-ureidoglycine was produced *in situ* by employing the two preceding enzymes, allantoinase (19) and AAH (16), with allantoin as the substrate (Figs. 1 and 4*B*). A coupled reaction was also used in which (*S*)-ureidoglycolate produced by AtUGlyAH was oxidized into oxalurate by an NAD-dependent *E. coli* (*S*)-ureidoglycolate dehydrogenase, with concomitant formation of NADH (Figs. 1 and 4*C*). Therefore, the steady-state kinetic parameters for various AtUGlyAH mutant enzymes were measured by monitoring the increase in absorbance at 340 nm caused by the formation of NADH.

A total of eight residues were selected for site-directed mutagenesis (Fig. 13); their kinetic parameters are listed in Table 5 (Fig. 14). The residues involved in the Mn^{2+} coordination shell were sensitive to mutation. Mutant enzymes such as E235A, H237A, H241A, and Q275A were catalytically incompetent, suggesting that these residues are essential for coordinating with the Mn^{2+} ion. It is noticeable that E235Q mutant is an essentially inactive enzyme (Table 5) but is capable of binding the Mn^{2+} ion as much as the WT enzyme does (Table 4). Unexpectedly, the K_m value of Y252F was comparable, but its k_{cat} value was reduced by ~10-fold compared to that of WT enzyme. However, enzyme activity was essentially abolished in the Y287F and Y287A mutants, suggesting that the side chain hydroxyl group in Tyr-287, but not Tyr-252, plays an essential role in catalysis and/or substrate binding. The K_m value of K291R was ~twofold higher, and its k_{cat} value was ~65% of the WT enzyme, while enzyme activity was completely absent in the K291A mutant. This suggests that the positive charge at Lys-291 plays a pivotal role in substrate binding, possibly by neutralizing the negative charge of the carboxyl group of (*S*)-ureidoglycine (Figs. 12 and 13). Unlike other mutant

enzymes, initial velocity measurements of the H221A mutant were unsuccessful, mainly due to non-linear changes in absorbance as a function of reaction time under various assay conditions. However, its activity was estimated to be similar to that of Y252F (Fig. 14*D*). Since all AtUGlyAH mutant enzymes used in this functional analysis were estimated to be in a native conformation based on circular dichroism spectra that was similar to the WT enzyme (Fig. 15), a lack of activity in some of these mutants is related to the functional features of those residues, not due to a possible conformation change by mutation.

TABLE 5. Kinetic parameters of AtUGlyAH and its mutants[a]

	K_m (mM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($mM^{-1}sec^{-1}$)
WT	1.77(0.104) [b]	761(14.1)	429
Y252F	2.30(0.0910)	73(0.73)	32
K291R	3.46(0.196)	493(10.9)	143

[a]Mutants as follows were inactive: E235A, E235Q, H237A, H241A, Q275A, Y287A, Y287F, K291A.
[b]Values in parentheses are standard errors estimated by the program SigmaPlot.
[c]The initial velocity measurements of the H221A mutant were unsuccessful, mainly due to non-linear changes in absorbance as a function of reaction time under various assay conditions.

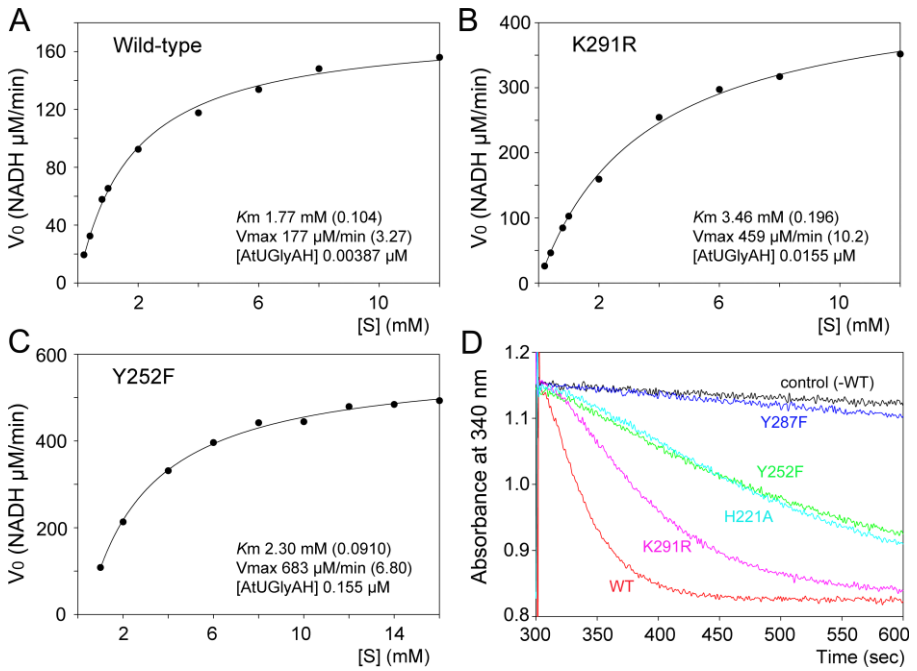


FIGURE 14. **Kinetic data of various AtGlyAH mutant enzymes.** *A-C*, the measured initial velocity (Fig. 4*C*) *versus* substrate concentration was plotted. The kinetic parameters were computed using SigmaPlot, and the enzyme concentration used is indicated. *D*, reaction progress curves for various mutant enzymes are shown. For this experiment, the procedures shown in Fig. 4 were followed. Allantoate (75 µM) was added to the reaction mixture at 50 sec and either the AtUGlyAH mutant or WT enzyme (0.155 µM) was added to the reaction mixture after 250 sec of incubation to produce (*S*)-ureidoglycolate. The reaction progress curve for ammonia release by AtUGlyAH was recorded at 300 sec and the decrease in absorbance at 340 nm was due to oxidation of NADPH to $NADP^+$ in a coupled reaction with an NADPH-dependent glutamate dehydrogenase. Consistent with kinetic parameters (Table 5), the activity of the K291R mutant was closest to that of the WT enzyme. While the Y287F mutant was almost inactive (activity similar to that in the absence of enzyme), the catalytic activity of the H221A mutant was more or less similar to that of the Y252F mutant.

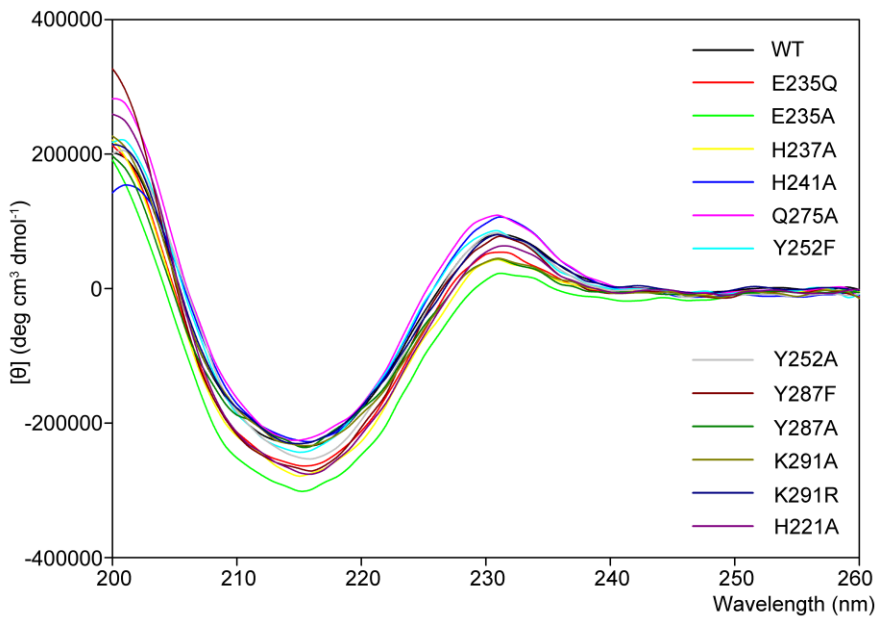


FIGURE 15. **Circular dichroism spectra of the WT and mutant AtUGlyAH enzymes.** Spectra were measured in an 1-mm path-length cuvette with Chirascan™-plus CD spectrometer (AppliedPhotophysics), using an enzyme concentration of 0.5 mg/mL in 50 mM Tris-HCl (pH 7.4).

DISCUSSION

The presence of a gene encoding the UGlyAH enzyme, and its biochemical function in the ureide pathway was identified in both *A. thaliana* and *E. coli* (14, 15). However, the Research Centers for Structural Genomics had earlier determined the crystal structure of YlbA, an *E. coli* ortholog of UGlyAH (Fig. 4; PDB ID 1RC6), and three structural homologs from other bacteria (Fig. 5; PDB ID 1SFN, 1SEF, and 1SQ4). These four proteins exhibit structural features highly homologous with those of AtUGlyAH; a bi-cupin monomer and an overall octameric conformation, with Z-scores of 28.3–34.8, sequence identities of 26–40%, and rmsd values of 1.4–2.5 Å for the monomer, as defined by the DALI program (44). Although the metal-binding site was not recognized in those structures, high similarities in both structure and sequence, including residues for metal-coordination and the substrate-binding site (Fig. 5), suggest that these three structural homologs are also members of the UGlyAH family. In particular, the cupin fold in this family of proteins is associated with a broad-range of enzyme activities and is widely distributed in the protein kingdom (41).

Structural and functional analysis of AtUGlyAH revealed several functional consequences. First, AtUGlyAH is an Mn^{2+}-dependent enzyme (14). It is evident from the structure of the binary complex that the Mn^{2+} ion acts as a molecular anchor for the substrate; it positions the incoming (*S*)-ureidoglycine into the productive binding mode for catalysis (Fig. 12). Secondly, the resulting binding mode serves as the molecular basis of the catalytic enantioselectivity. More specifically, a molecular face with a hydrogen atom attached to the chiral carbon atom is vulnerable to enzyme catalysis. Of the neighboring residues, Tyr-287 and

Lys-291 are essential for enzyme activity, possibly by dictating the orientation of the ureido and carboxyl groups of the substrate, respectively (Table 5 and Fig. 12). Binding of the (*S*)-configuration of ureidoglycine is further stabilized by Glu-235, due to possible interactions between the amino group of a reaction center carbon and the carboxyl group of Glu-235 (Fig. 13). In addition, the side chain carboxyl group on Glu-235 appears to be crucial in catalysis, because the E235Q mutant still containing Mn^{2+} ion (Table 4) is catalytically incompetent.

Potential catalytic mechanisms of AtUGlyAH were considered by examining structural and functional analyses, as well as inversion of configuration at the chiral carbon center between (*S*)-ureidoglycine and (*S*)-ureidoglycolate (Fig. 1) (14). A possible direct attack of the general base-catalyzed water molecule at the chiral carbon appears to be highly unlikely; such an S_N2-type reaction requires a water molecule which presents on the side opposite to the leaving amino group, with a geometrical requirement for an inline attack on the chiral carbon center. The presence of the water molecule is implausible, mainly due to steric hindrance in that area which is densely populated with His-241, Leu-289, and Lys-291 (Fig. 12). Alternatively, allanturate (10) could be an intermediate in catalysis (Fig. 16). Glu-235, with the probable low p*K* value of the hydrogen on the Mn^{2+}-bound amide nitrogen, could facilitate abstraction of the proton from the amide nitrogen and subsequent transfer of this proton to the amino leaving group in the form of ammonia, which is analogous to the proposed function of the metal-bound aspartate in adenine deaminase (45). Subsequently, the resulting planar intermediate allanturate could be subject to a nucleophilic attack by the general base-catalyzed water molecule, yielding (*S*)-ureidoglycolate. Specifically, it is likely that the *re* face of the reaction center carbon in the intermediate remains

occluded by Glu-235, as observed in the binary complex (Fig. 12). Therefore, the general base-catalyzed water molecule could reach the intermediate only from the *si* face of the planar intermediate, which is consistent with a previous observation of an inversion of configuration at the chiral carbon atom after catalysis (Fig. 1) (14).

However, this analysis does not reveal any obvious general base to deprotonate the incoming water molecule in the immediate vicinity of the chiral carbon atom of (*S*)-ureidoglycine (Figs. 12 and 13), but provides two candidates, His-221 and Tyr-287. Both are located on the side opposite the leaving amino group of the substrate, 5.3–5.4 Å from the chiral carbon atom. The presence of a water molecule connecting His-221 and the chiral carbon atom (Fig. 13), as well as the low catalytic activity of the H221A mutant (Fig. 14*D*), raises the possibility that His-221 acts as the proposed general base. However, His-221 is not a strictly conserved residue in the UGlyAH family (Fig. 5), in which the mechanistic features should remain identical. In contrast, Tyr-287 is invariant in the family, but its side chain hydroxyl group should be deprotonated to act as a general base. The following observations suggest that a hydrogen bonding network between Tyr-252 and Tyr-287, as well as the ureido group, might participate in the deprotonation of Tyr-287 and facilitate its functional role in catalysis: (i) the Y287F mutant is an inactive enzyme; (ii) the two hydroxyl groups in Tyr-287 and Tyr-252 are separated by ~3.7 Å; and (iii) the hydroxyl group in Tyr-252 is involved in catalysis rather than substrate binding, given that the Y252F mutation largely affects the k_{cat}, but not the K_m, value. Therefore, Tyr-287 is a candidate general base. However, Tyr-287 became almost buried after binding of (*S*)-ureidoglycine. Lack of the amino group in the intermediate could cause its binding mode to differ from that in the enzyme–

substrate complex, possibly by inclining towards Glu-235, exposing Tyr-287 to the solvent. Further investigation is necessary for conclusive identification of the general base in AtUGlyAH.

In this study, a crystal structure of the ligand-free form of AtUGlyAH and its complex with (*S*)-ureidoglycine were determined. The structures revealed the overall features of both the monomeric and octameric conformations, which are applicable to other UGlyAH family proteins, as well as the binding mode of (*S*)-ureidoglycine in the Mn^{2+} coordination shell. The enzyme assay, which employed an *E. coli* (*S*)-ureidoglycolate dehydrogenase as a coupled enzyme, allows for measurement of the steady-state kinetic parameters of AtUGlyAH. These results provide molecular insights into the structure of UGlyAH and early events in the catalytic release of ammonia from the nitrogen-rich metabolite, which is the crucial step of the ureide pathway.

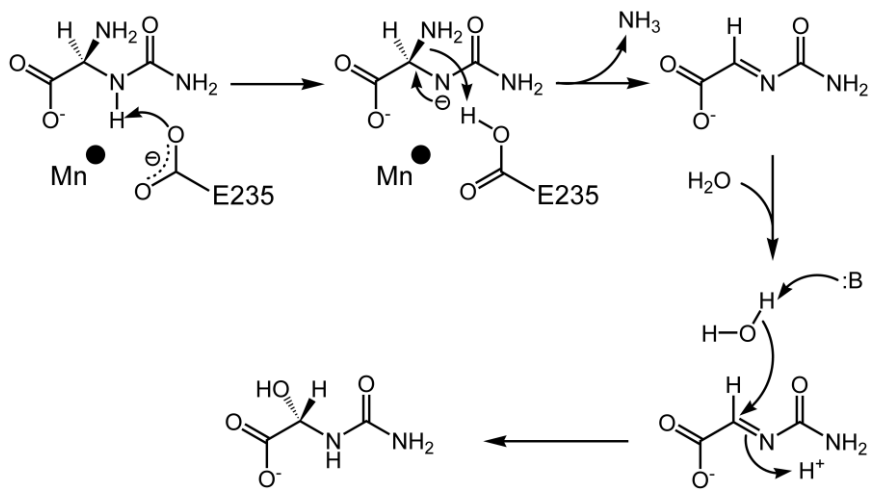


FIGURE 16. **Proposed mechanism for AtUGlyAH.** Details for the proposed mechanism are described in text. In the current analysis, His-221 and Tyr-287 were proposed as the possible general base to deprotonate the incoming water molecule.

CHAPTER III.

Structural and Functional Insights into the Substrate Specificity of (*S*)-Ureidoglycolate Amidohydrolase from *Arabidopsis thaliana* and Its Comparison with Allantoate Amidohydrolase

MATERIALS AND METHODS

1. Construct

The genes used in this study were as follows: allantoinase from *E. coli* (gene accession no. NC_010473:477702–479063), AAH from *A. thaliana* (gene accession no. NM_118126) and *E. coli* strain *K-12* substrain *DH10b* (gene accession no. NC_010473:482612–483847), UGlyAH from *A. thaliana* (gene accession no. NM_117809), AtUAH (gene accession no. NM_123726), and allantoicase from *Pseudomonas syringae* pv. tomato strain DC3000 (gene accession no. NC_004578:4136051–4137046).

Full-length genes were used for constructing expression vectors for bacterial genes. N-terminal truncated variants, from which the signal peptide and an additional N-terminal region (AtAAH, residues 1–75; AtUGlyAH, residues 1–35; AtUAH, residues 1–49) had been removed, were used for genes from *A. thaliana*. The genes were amplified using a pair of sequence-specific primers (Table 6). The resulting PCR products were restricted by NdeI and XhoI enzymes and introduced into a modified pET-28a vector (Merck) containing a *Tobacco etch virus* protease cleavage site designed for the removal of N-terminal His_5-tag. Mutants of AtUAH and EcAAH were also prepared using the QuickChange method (Agilent Technologies) with mutagenic primers (Table 6).

TABLE 6. Primers for AtUAH, EcAAH and other enzymes

AtAAH(Δ75)	
Forward	GGAATTCCATATGGGAATCGGTGAAACAGAAG
Reverse	AACCGCTCGAGTTACATTTGAGACTCTAGAAAC
AtUGlyAH(Δ35)	
Forward	GGAATTCCATATGAAGACGAACCCCATTTATTG
Reverse	AACCGCTCGAGTCACAATGGATTTCGATTCAC
AtUAH(Δ49)	
Forward	GGAATTCCATATGTTCGGTTCCATCAATCTCG
Reverse	AACCGCTCGAGTCAGTCTAAAGAGAGCTTAGC
E183Q	CTTGTTCACATCT**CAA**GAACCGACCCGATTTG
E183D	CTTGTTCACATCT**GAC**GAACCGACCCGATTTG
E183A	CTTGTTCACATCT**GCA**GAACCGACCCGATTTG
H138A	CTGTAGCCACTGGTTCT**GCC**ATCGATGCTATTCC
D149N	TATTCCGGGAAATAT**AAC**GGTGTTGTCGGTGTT
D149A	CATATTCCGGGAAATAT**GCT**GGTGTTGTCGGTG
E184A	CTTGTTCACATCTGAA**GCA**CCGACCCGATTTG
H254A	GCTTTTCTAGAGTTG**GCT**ATTGAACAGGGACC
H448A	CTACAAAGGATACAGC**GCT**AAACCTGAAGAATAC
H290Q	GAAGGCAATGGAGGT**CAA**GCTGGTGCTGTAC
H290N	GAAGGCAATGGAGGT**AAT**GCTGGTGCTGTAC
H290A	GAAGGCAATGGAGGT**GCT**GCTGGTGCTGTAC
N340D	CCTGGGGCAATC**GAC**AGCATTCCAAGCAAATC
N340A	CCTGGGGCAATC**GCC**AGCATTCCAAGCAAATC
R353K	CTTGAAATCGATACA**AAG**GATATCGATGAAGCG
R353A	CTTGAAATCGATACA**GCG**GATATCGATGAAGCG
D351A	CCCATCTTGAAATC**GCT**ACAAGGGATATCGATG
Y423F	GATGATCAGCCGAGCT**TTT**CATGATTCTCTCTTC
Y423G	GATGATCAGCCGAGCT**GGT**CATGATTCTCTCTTC
Y423A	GATGATCAGCCGAGCT**GCT**CATGATTCTCTCTTC
Q257E	GAGTTGCATATTGAA**GAG**GGACCTATACTTGAAG
Q257N	GAGTTGCATATTGAA**AAC**GGACCTATACTTGAAG
Q257A	GAGTTGCATATTGAA**GCG**GGACCTATACTTGAAG
H424N	ATCAGCCGAGCTTAT**AAC**GATTCTCTCTTCATG
H424A	ATCAGCCGAGCTTAT**GCG**GATTCTCTCTTCATG
EcAAH	
Forward	GGAATTCCATATGATTACACATTTCCGTCAAG
Reverse	AACCGCTCGAGTTATTTCTGCCAGGCAAG
E126A	TGGCGATGGCAGAA**GCT**GAAGGCAGCCGCTTC

Sequences are written from 5' to 3'. For mutagenic primers, forward versions are listed. Restriction sites used for cloning are underlined, and the boldfaced-underlines show the mutated sequences.

2. Protein Expression and Purification

For structural studies, SeMet-substituted, N-terminal His-tagged AtUAH was expressed in *E. coli* BL21 (DE3) cells (Merck). The cells had been grown at 37℃ in M9 minimal medium until the optical density at 600 nm reached 0.6. The culture was then supplemented with a mixture containing lysine, threonine, phenylalanine, leucine, isoleucine, and valine to inhibit endogenous methionine biosynthesis (46). After 30 min, the SeMet-substituted recombinant AtUAH protein was expressed for 18 h at 20℃ with the addition of 0.25 mM of SeMet and 0.5 mM isopropyl-L-thio-β-D-galactopyranoside. The cells were harvested and sonicated in buffer A (50 mM NaH_2PO_4, pH 7.4, and 200 mM NaCl) and the supernatant was obtained by centrifugation at 30,000 × *g* for 1 h at 4℃. The enzyme was purified using an immobilized metal affinity column (GE Healthcare) that had been equilibrated with buffer A. The eluent was collected using buffer A with 500 mM imidazole. After the pooled elution fractions were dialyzed against buffer B (50 mM NaH_2PO_4, pH 7.4, and 100 mM NaCl), the N-terminal His-tag was removed by incubating the dialyzed samples with *Tobacco etch virus* protease for 12 h at 4℃. The His-tag-free SeMet–AtUAH was further purified by immobilized metal affinity chromatography followed by gel filtration chromatography using a Superdex-200 column (GE Healthcare) with buffer C (50 mM Tris, pH 7.4).

Other proteins including native, WT AtUAH and its various mutant enzymes, the EcAAH(E126A) mutant, and other enzymes used in this study were expressed in Luria–Bertani medium and purified as described above. Except for when used in structural studies, all enzymes were purified without removing the N-terminal His-tag.

3. Crystallization

Four different structures are presented in this study: three structures of AtUAH, each with a distinct ligation state, and one structure of EcAAH in a complex with allantoate.

To crystallize AtUAH, SeMet–AtUAH was concentrated to ~10 mg/mL, and 1 mM $MnCl_2$ and 5 mM dithiothreitol were added prior to crystallization. Crystallization was performed at 295 K using the sitting-drop vapor-diffusion method in a crystallization buffer of 0.1 M MES pH 6.5, 12% (w/v) polyethylene glycol 20000 (Fig. 17). Under these conditions, an electron density corresponding to the product glyoxylate was clearly identified in the vicinity of the active site. Therefore, this structure is referred to as the AtUAH–glyoxylate complex, although crystallization was performed in the absence of any ligands. Crystals of the AtUAH binary complex with its substrate or reaction intermediate were obtained by co-crystallization of a catalytically incompetent AtUAH(E183A) mutant with a cognate substrate, (*S*)-ureidoglycolate. In particular, Glu-183 is strictly conserved in the peptidase M20 family (InterPro accession IPR002933) and has been proposed as a catalytic base (47, 48). Crystals of the AtUAH(E183A)–ureidoglycolate complex were obtained from a crystallization solution containing 0.2 M calcium acetate hydrate, 20% (w/v) polyethylene glycol 3350, 1 mM $MnCl_2$, and 5 mM (*S*)-ureidoglycolate. A proposed reaction intermediate, (*S*)-hydroxyglycine, was identified in the structure of an AtUAH(E183A)–hydroxyglycine complex whose crystal had been grown from a crystallization solution of 0.1 M MES (pH 6.5), 12% (w/v) polyethylene glycol 20000, 1 mM $MnCl_2$, and 2 mM (*S*)-ureidoglycolate.

A catalytically inactive EcAAH(E126A) mutant was also crystallized with a

cognate substrate allantoate (EcAAH(E126A)–allantoate complex). Specifically, Glu-126 in EcAAH corresponds to Glu-183 in AtUAH. Preliminary analysis of enzyme activity indicated that divalent metal ions such as Mn^{2+}, Ni^{2+}, Co^{2+} are required for maximum activity (Fig. 18). Crystals of the EcAAH(E126A)–allantoate complex were therefore obtained from a solution of 0.1 M sodium citrate (pH 5.6), 20% (w/v) polyethylene glycol 4000, 20% (v/v) isopropanol, 1 mM $MnCl_2$, and 5 mM allantoate.

Substrates including allantoate for EcAAH and (*S*)-ureidoglycolate for AtUAH were prepared as follows. First, a solution of 20 mM allantoate was prepared from a racemic mixture allantoin using the enzyme allatoinase, as described in the MATERIALS AND METHODS of Chapter I. Next, the enzyme allantoicase (15) was employed to produce (*S*)-ureidoglycolate directly from the prepared allantoate. After each step, the reaction mixture was filtered through an Amicon Ultra Centrifugal Filter Device (Millipore) with a 3-kDa molecular weight cutoff.

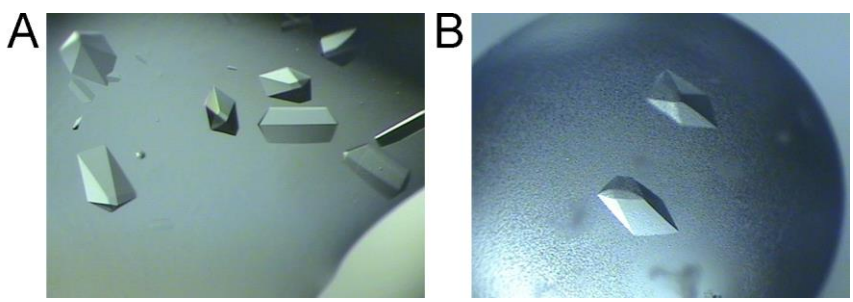


FIGURE 17. **Crystals of AtUAH.** *A*, crystals of SeMet-substituted AtUAH were obtained using a crystallization buffer of 0.1 M MES pH 6.5, 12% (w/v) polyethylene glycol 20000. *B*, crystals of native AtUAH were also grown in a crystallization buffer of 0.1 M MES pH 6.5, 12% (w/v) polyethylene glycol 20000.

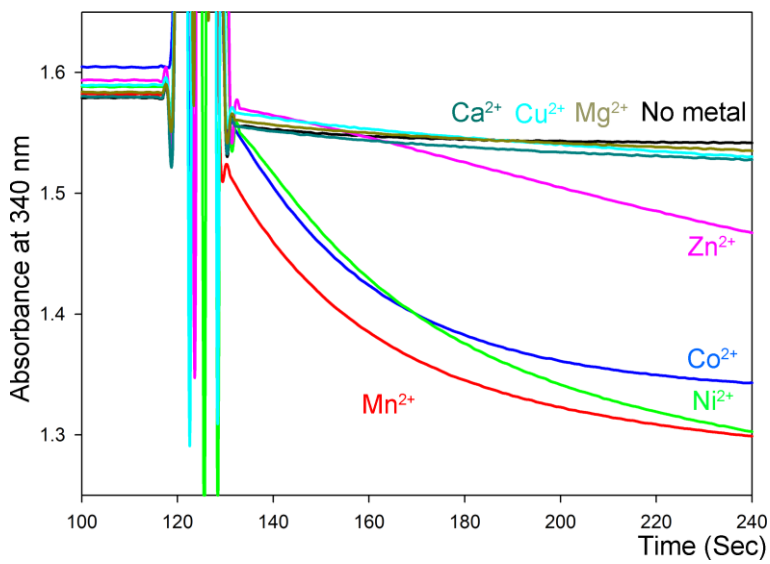


FIGURE 18. **Metal-dependency of EcAAH.** The metal-dependent reaction profile of EcAAH is shown. In the assay, 30 μM allantoate was used as a substrate. The solution contained 100 μM of each divalent metal. In the control, no metal was added in the reaction mixture.

4. Data Collection and Structure Determination

Multiwavelength anomalous dispersion data for the SeMet-substituted AtUAH–glyoxylate complex were collected to a resolution of 2.05 Å on Beamline AR-NW12A at the Photon Factory (Tsukuba, Japan) (Table 7). Data with a higher resolution of 1.45 Å were later collected using crystals of native AtUAH on Beamline BL1A at the Photon Factory. Single-wavelength data were collected on Beamline 7A of the Pohang Accelerator Laboratory (Pohang, Korea) at resolutions of 1.90 Å for the AtUAH(E183A)–ureidoglycolate complex, 1.90 Å for the AtUAH(E183A)–hydroxyglycine complex, and 2.20 Å for the EcAAH(E126A)–allantoate complex. All data were collected at 100 K and processed using the program HKL2000 (31). Solutions of 20–30% (v/v) ethylene glycol and 25% (v/v) glycerol were used as cryoprotectants for AtUAH crystals and EcAAH crystals, respectively. Crystals of AtUAH belonged to the $P2_12_12_1$ space group, while crystals of EcAAH belonged to the $P2_12_12$ space group. Both crystals contained two monomers of enzyme in an asymmetric unit.

The structure of AtUAH–glyoxylate was initially determined at a resolution of 2.05 Å using multiwavelength anomalous dispersion data and the software programs SOLVE (32) and RESOLVE (33, 34) for phasing and density modification, respectively. Data acquired at a resolution of 1.45 Å were later used to refine the model of AtUAH–glyoxylate. Molecular model building and refinement were carried out using COOT (35) and PHENIX (37) software packages. After several iterative manual inspections and refinement processes, the model was built including Asn-54 to Asp-476 for each monomer. Specifically, rigid-body refinements were followed by simulated annealing. During refinement, water molecules, whose refined temperature factors were less than 50 $Å^2$, were assigned

based on possible hydrogen bonds to the enzyme or nearby water molecules. In the final refinement stage, TLS refinement was performed using multiple TLS groups identified automatically in PHENIX. In both subunits, the electron density of glyoxylate was identified in the vicinity of the bimetal reaction center. The final model of the AtUAH–glyoxylate complex was used as a starting model for refinement of the AtUAH(E183A)–ureidoglycolate and AtUAH(E183A)–hydroxyglycine complexes. During refinement, an electron density corresponding to (*S*)-ureidoglycolate or hydroxyglycine was observed in the bimetal center.

For the EcAAH(E126A)–allantoate complex, molecular replacement was performed using a structure of EcAAH that had been previously determined (16) (PDB ID 1Z2L) as a search model in PHENIX. Specifically, an initial trial using the whole molecule of 1Z2L as a search model failed. Eventually, two monomers were able to be located in an asymmetric unit by using two segments of 1Z2L as search models: one segment for a catalytic domain (Ile-2 to Asn-208, Pro-329 to Lys-411) and the other for a dimerization domain (Gly-212 to Glu-325). Subsequent refinement revealed the electron density for a hinge region between the two segments, as well as allantoate located in the active site. The EcAAH(E126A)–allantoate complex is also observed as a closed conformation, which differs from the open conformation of a previously characterized EcAAH structure (16). After several cycles of iterative manual inspection and refinement, a model for the EcAAH(E126A)–allantoate complex was built including residues Met-1 to Lys-411 in each monomer.

The stereochemistry of the refined models was evaluated using a MolProbity (38). Details on data collection and refinement are described in Table 8. Structural comparisons and analyses were carried out using the CCP4 suite program (39).

Figures depicting molecular structure were prepared using PyMOL (The PyMOL Molecular Graphics System).

TABLE 7. Data collection for SeMet-substituted AtUAH

	WT AtUAH SeMet		
Data collection	Peak	Edge	Remote
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions			
a (Å)	69.4	69.5	69.5
b (Å)	88.5	88.6	88.6
c (Å)	162.3	162.4	162.5
Wavelength (Å)	0.97923	0.97940	0.96417
Resolution (Å)	50-2.05(2.12-2.05)[a]	50-2.05(2.12-2.05)	50-2.04(2.11-2.04)
Total reflections	876430	879856	889262
Unique reflections	62374	62595	63419
Redundancy	14.1(13.8)	14.1(13.8)	14.0(13.6)
R_{sym} or R_{merge}[b] (%)	9.3(27.5)	9.6(31.0)	10.2(35.6)
I / σI	29.8(8.3)	29.4(7.4)	28.5(6.1)
Completeness (%)	98.7(96.7)	98.6(96.3)	98.0(90.0)
FOM[c]		0.74	

[a]Numbers in parentheses refer to data in the highest resolution shell.

[b]$R_{merge} = \Sigma|I_h - \langle I_h \rangle| / \Sigma I_h$, where I_h is the observed intensity and $\langle I_h \rangle$ is the average intensity.

[c]FOM: figure of merit after RESOLVE (33, 34)

TABLE 8. Data collection and refinement statistics for AtUAH and EcAAH

PDB ID	WT AtUAH Glyoxylate 4PXE	AtUAH E183A Ureidoglycolate 4PXB	AtUAH E183A Hydroxyglycine 4PXC	EcAAH E126A Allantoate 4PXD
Data collection				
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12$
Cell dimensions				
a (Å)	69.9	70.0	70.1	106.8
b (Å)	89.8	89.5	89.7	171.9
c (Å)	163.9	163.1	164.1	43.2
Wavelength (Å)	1.1	0.97934	0.97934	0.97934
Resolution (Å)	50-1.45(1.50-1.45)[a]	50-1.90(1.97-1.90)	50-1.90(1.97-1.90)	50-2.20(2.28-2.20)
Total reflections	1593063	1124546	580198	599186
Unique reflections	182920	80858	83293	41528
Redundancy	8.7(8.7)	13.9(13.7)	7.0(6.8)	14.4(14.6)
R_{sym} or R_{merge}[b] (%)	11.4(86.8)	16.4(68.8)	13.3(66.0)	14.9(74.1)
I / σI	20.7(2.2)	16.7(4.2)	14.2(2.9)	19.7(4.5)
Completeness (%)	100(100)	99.8(99.5)	99.9(99.8)	99.9(100)
Refinement				
Resolution (Å)	46.7-1.45	30.6-1.90	29.8-1.90	31.3-2.20
No. Reflections	182800	80780	83205	41466
R_{work}[c] / R_{free}[d]	16.2/17.6	16.2/18.9	16.8/19.5	16.8/21.8
No. Atoms	7741	7449	7545	6911
Protein[e]	6430	6396	6396	6408
Manganease	4	4	4	2
Ligand	10	18	12	24
Water	1297	1031	1133	477
B-factors (Å^2)	19.4	18.9	22.0	20.3
Protein	18.5	17.1	19.7	24.2
Manganease	18.4	13.2	19.7	14.3
Ligand	14.6	18.4	21.1	14.8
Water	32.2	29.1	30.7	29.5
r.m.s. deviations				
Bond lengths (Å)	0.006	0.007	0.007	0.007
Bond angles (°)	1.130	1.089	1.066	1.114
Ramachandran plot				
Favored (%)	97.8	96.9	97.5	98.4
Allowed (%)	2.2	3.1	2.5	1.6
Outlier (%)	0	0	0	0

[a]Numbers in parentheses refer to data in the highest resolution shell.

[b]$R_{merge} = \Sigma|I_h - \langle I_h \rangle| / \Sigma I_h$, where I_h is the observed intensity and $\langle I_h \rangle$ is the average intensity.

[c]$R_{work} = \Sigma \, ||F_{obs}| - k|F_{cal}|| / \Sigma |F_{obs}|$

[d]R_{free} is the same as R_{obs} for a selected subset (10%) of the reflections that was not included in prior refinement calculations.

[e]Ordered residues: AtUAH (Asn-54 to Asp-476 in both subunits), and EcAAH (Met-1 to Lys-411 in both subunits).

5. Enzyme Assays

Steady-state kinetic analyses of WT AtUAH and its mutant enzymes were performed at 30°C using a UV-visible spectrophotometer (V-650; Jasco). In these analyses, AtUAH-dependent ammonia release was monitored using a coupled assay with an NADPH-dependent glutamate dehydrogenase (40). In brief, the decrease in absorbance at 340 nm was measured after the addition of a known concentration of (*S*)-uriedoglycolate. The change in absorbance is due to the oxidation of NADPH in the reaction mixture (2 mL), which contained the enzyme of interest, 0.2 M Tris–HCl (pH 8.5), 0.3 mM NADPH, 2.5 mM α-ketoglutarate, 10 units of glutamate dehydrogenase from *Proteus* sp. (Sigma-Aldrich), and 100 μM $MnCl_2$. Further experiments validated that concentrations of glutamate dehydrogenase and NADPH in the reaction mixture were in excess. The substrate (*S*)-uriedoglycolate was prepared by converting 20 mM (*S*)-allantoin to (*S*)-ureidoglycolate in the presence of allantoicase (6.4 μM) at 30°C for 2 h. The resulting (*S*)-ureidoglycolate solution was stored at –70℃ prior to its use in enzyme assays. The initial enzymatic velocity was determined by measuring the linear decrease in absorbance at 340 nm for 30 s. Since two molecules of ammonia are released from one molecule of (*S*)-ureidoglycolate during the reaction with AtUAH (Figs 1 and 2), the initial velocity was calculated as half of the NADPH concentration that was consumed per minute, assuming a molar extinction coefficient of 6220 $M^{-1} \cdot cm^{-1}$ at 340 nm. K_m and V_{max} values were obtained using SigmaPlot (Systat Software), and k_{cat} values were computed by dividing V_{max} by the AtUAH concentration used.

RESULTS

1. Overall Structure of AtUAH

Three different structures of AtUAH, each in a different ligation state, are described in this study (Table 8). These three structures are all in the $P2_12_12_1$ space group and the two monomers of AtUAH that exist in an asymmetric unit are related by a non-crystallographic twofold symmetry. The overall structural features of AtUAH and its active site can be described based on the AtUAH–glyoxylate complex. Detailed features of the active site are also presented with regard to the binding of substrate, intermediate, and product.

Monomeric AtUAH (Asn-54 to Asp-476) is composed of 13 α-helices, 12 β-strands, and 2 short 3_{10}-helices (Fig. 19). It folds largely into two structural domains (Fig. 20): a catalytic domain (residues 54–275, 392–476), and a dimerization domain (residues 276–391) that is inserted between β6 and α11 of the catalytic domain. The two structural domains are connected by a so-called hinge region (residues 273–275 and 392–394) and the functional roles of each domain are discussed below. This overall structure and conformation of monomeric AtUAH are typical for the amidase, hydantoinase/carbamoylase families (InterPro accession IPR010158). A homologous structure search using DALI (44) revealed a *Z*-score of more than 42 (16, 47, 48), or more than 20 when considering the broader peptidase M20 family (InterPro accession IPR002933) (49–53). Enzymes in this family hydrolyze the amide bond of compounds containing carbamoyl groups.

The catalytic domain exhibits an α/β/α-folded architecture. Specifically, eight strands are found in the center of the catalytic domain arranged as β1, β2, β4, β3, β5, β12, β6, and β11, in which β2, β6, and β11 are antiparallel to the remaining five

strands. This central β-sheet is sandwiched between a layer of four α-helices (α3, α4, α13, α11) and a layer of two α-helices (α5, α12). Note that α5 and α12 face the dimerization domain. An antiparallel four-stranded β-sheet from β7 to β10 forms one facet of the dimerization domain and faces the catalytic domain, with two long α-helices (α9, α10) and one short 3_{10}-helix located on the surface of AtUAH. This dimerization β-sheet connects the catalytic domain by the hinge region. Two protruding loops across from the hinge region are located at the tip of the dimerization domain. One loop connecting β7 and α9 is referred to as the H-loop in this study and contains His-290. The other loop between β8 and β9 is referred to as the N-loop, which harbors Asn-340 (Fig. 20).

Dimerization of AtUAH is achieved by extensive interactions, mainly between the dimerization domains of each subunit (Fig. 21). In particular, two subunits related by a non-crystallographic twofold symmetry are packed in such a manner that the dimerization β-sheets from both subunits are juxtaposed in an antiparallel manner, apparently resulting in a continuous, antiparallel, eight-stranded β-sheet in the dimer. The dimer interface is mediated largely by β8 and α9, and by β8* and α9* (hereafter, an asterisk denotes an element or residue from another subunit). The protruding H-loop and N-loop from one subunit penetrate into another subunit and provide additional structure to the cleft located at the interface between the catalytic and dimerization domains, forming a pocket that serves as the active site (Fig. 21).

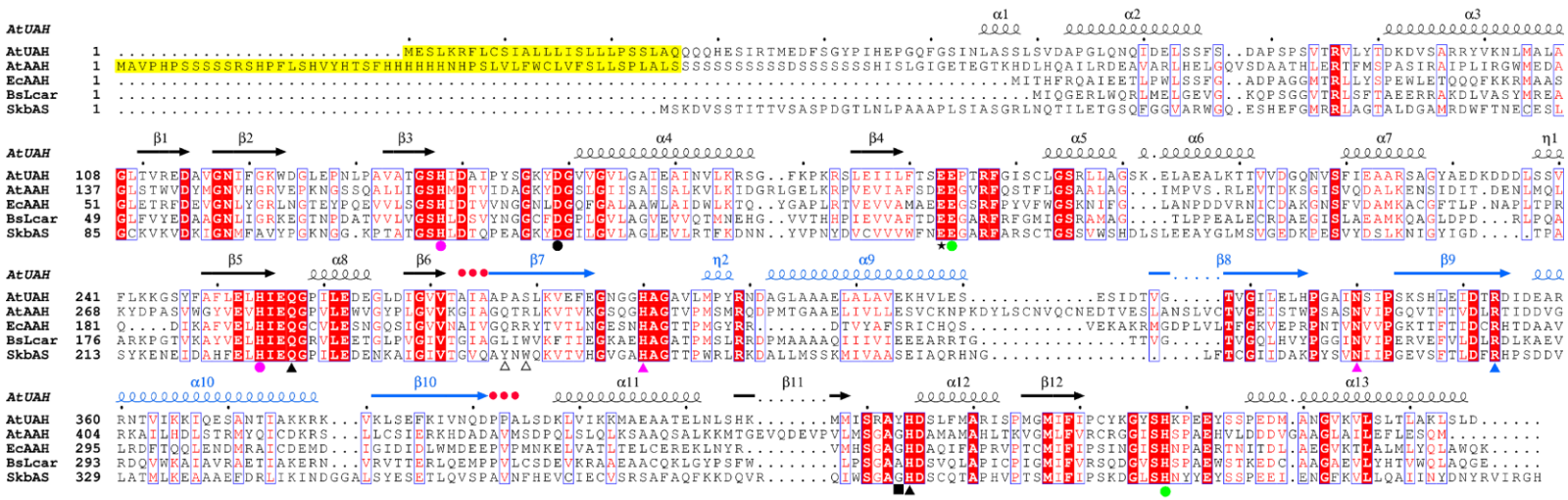

AtUAH
AtUAH 1 ..........MESLKRFLCSIALLLISLLLPSSLAQQQHESIRTMEDFSGYPIHEPGQFGSINLASSLSVDAPGLQNQIDELSSFS..DAPSPSVIRVLYTDKDVSARRYVKNLMALA
AtAAH 1 MAVPHPSSSSSRSHFFLSHVYHTSFHHHHHHNHPSLVLFWCLVFSLLSPLALSSSSSSSSSSDSSSSSSSHISLGIGETEGTKHDLHQAILRDEAVARLHELGQVSDAATHLERTFMSPASIRAIPLIRGWMEDA
EcAAH 1 .....MITHFRQAIEETLPWLSSFG..ADPAGGMTRLLYSPEWLETQQQFKKRMAAS
BsLcar 1 .....MIQGERLWQRLMELGEVG..KQPSGGVTRLSFTAEERRAKDLVASYMREA
SkbAS 1 .....MSKDVSSTITTVSASPDGTLNLPAAAPLSIASGRLNQTILETGSQFGGVARWGQ.ESHEFGMRRLAGTALDGAMRDWFTNECESL
AtUAH
AtUAH 108 GLTVREDAVGNIFGKWDGLEPNLPAVATGSHIDAIPYSGKYDGVVGVLGAIEAINVLKRSG..FKPKRSLEIILFTSEEPTRFGISCLGSRLLAGSK.ELAEALKTTVVDGQNVSFIEAARSAGYAEDKDDDLSSV
AtAAH 137 GLSTWVDYMGNVHGRVEPKNGSSQALLIGSHMDTVIDAGKYDGSLGIISAISALKVLKIDGRLGELKRPVEVIAFSDEEGVRFQSTFLGSAALAG...IMPVS.RLEVTDKSGISVQDALKENSIDIT.DENLMQL
EcAAH 51 GLETREDEVGNLYGRLNGTEYPQEVVLSGSHLDTVVNGGNLDGQFGALAAWLAIDWLKTQ..YGAPLRTVEVVAMAEEEGSRFPYVFWGSKNIFG...LANPDDVRNICDAKGNSFVDAMKACGFTLP.NAPLTPR
BsLcar 49 GLFVYEDAAGNLIGRKEGTNPDATVVLVGSHLDSVYNGGCFDGPLGVLAGVEVVQTMNEHG..VVTHHPIEVVAFTDEEGARFRFGMIGSRAMAG....TLPPEALECRDAEGISLAEAMKQAGLDPD...RLPQA
SkbAS 85 GCKVKVDKIGNMFAVYPGKNGG.KPTATGSHLDTQPEAGKYDGILGVLAGLEVLRTFKDNN..YVPNYDVCVVVWFNEEGARFARSCTGSSVWSHDLSLEEAYGLMSVGEDKPFSVYDSLKNIGYIGD....TPA
AtUAH
AtUAH 241 FLKKGSYFAFIELHIEQGPILEDEGLDIGVVTAIAAPASLKVEFEGNGGHAGAVLMPYRNDAGLAAAELALAVEKHVLES.....ESIDTVG....TVGILELHPGAINSIPSKSHLEIDTRDIDEAR
AtAAH 268 KYDPASVWGYVEVHIEQGPVLEWVQGYPIGVVKGIAGQTRLKVTVKGSQGHAGTVPMSMRQDPMTGAAELIVLLESVCKNPKDYLSCNVQCNEDTVESLANSLVGTVGEISTWPSASNVIPGQVTFTVDLRTIDDVG
EcAAH 181 Q....DIKAFVELHIEQGCVLESNGQSIGVVNAIVGQRRYTVTLNGESNHAGTTPMGYRR.......DTVYAFSRICHQS..........VEKAKRMGDPLVLTFGKVEPRPNTVNVVPGKTTFTIDCRHTDAAV
BsLcar 176 ARKPGTVKAYVELHIEQGRVLEETGLPVGIVTGIAGLIWVKFTIEGKAEHAGATPMSLRRDPMAAAAQIIIVIEEEARRTG.............TTVG....TVGQLHVYPGGINVIPERVEFVDLRDLKAEV
SkbAS 213 SYKENEIDAHFELHIEQGPILEDENKAHGIVTGVQAYNWQKVTVHGVGAHAGTTPWRLRKDALLMSSKMIVAASEIAQRHNG.................LFTCGIIDAKPYSVNAIPGEVSFTLDFRHPSDDV
AtUAH
AtUAH 360 RNTVIKKIQESANTTAKKRK...VKLSEFKIVNQDPPALSDKLVIKKMAEAATELNLSHK.......MMISRAYHDSLFMARISPMGMIFIPCYKGYSHKPEEYSSPEDM.ANGVKVLSLTLAKLSLD......
AtAAH 404 RKAILHDLSTRMYQICDKRS...LLCSIERKHDADAVMSDPQLSLQLKSAAQSALKKMTGEVQDEVPVLMSGAGHDAMAMAHLTKVGMLFVRCRGGISHSPAEHVLDDDVGAACLAILEFLESQM.........
EcAAH 295 LRDFTQQLENDMRAICDEMD...IGIDIDLWMDEEPVPMNKELVATLTELCEREKLNYR........VMHSGAGHDAQIFARPVPTGMIFIPSINGISHNPAERTNITDLLAEGVKTLALMLYQLAWQK.....
BsLcar 293 RDQVWKAIAVRAETIAKERN...VRVTTERLQEMPPVLCSDEVKRAAEAACQKLGYPSFW.......LPSGAAHDSVQLAPICPIGMIFVRSQDGVSHSPAEWSTKEDC.AAGAEVLYHTVWQLAQGE.....
SkbAS 329 LATMLKEAAAEFDRLIKINDGGALSYESETLQVSPAVNFHEVCIECVSRSAFAQFKKDQVR......QLWSGAGHDSCQTAPHVPTSMIFIPSKDGLSHNYYEYSSPEEI.ENGFKVLQALIINYDNYRVIRGH

FIGURE 19. **Sequence alignment of AtUAH.** The amino acid sequences of AtUAH are compared with those of *A. thaliana* and *E. coli* AAH and their structural homologs: EcAAH (PDB ID 1Z2L; *Z*-score, 42.4; sequence identity, 30%; rmsd, 5.4 Å), BsLcar for L-carbamylose from *Geobacillus stearothermophilus* (3N5F; 48.3; 37%; 4.6 Å), SkβAS for β-alanine synthase from *Saccharomyces kluyveri* (2V8G; 47.9; 30%; 1.9 Å). Sequences of endoplasmic reticulum-targeting signals are marked in yellow (residues 1–25 for AtUAH, 1–53 for AtAAH). Highly conserved residues are shown in red and boxed in blue. Strictly conserved residues are shown on a red background. The secondary structural elements defined in the AtUAH–glyoxylate binary complex are shown for the corresponding AtUAH sequences, with the catalytic and dimerization domains in black and blue, respectively. Hinge regions are indicated by red closed circles. Closed circles in different color codes represent residues involved in the bimetal coordination shell: residues ligating M1, M2, and bridging residues are shown in magenta, green, and black, respectively. A black asterisk indicates the catalytic base (47, 48). Closed triangles in black represent residues involved in the (*S*)-ureidoglycolate binding site, and those in magenta indicate residues from a neighboring subunit. Open triangles show the QXR motif only in AAH. A black rectangle denotes the tyrosine residue with stacking interactions with the carboxyl groups of (*S*)-ureidoglycolate, (*S*)-hydroxyglycine, and glyoxylate in UAH. This figure was prepared using ESPript (42).

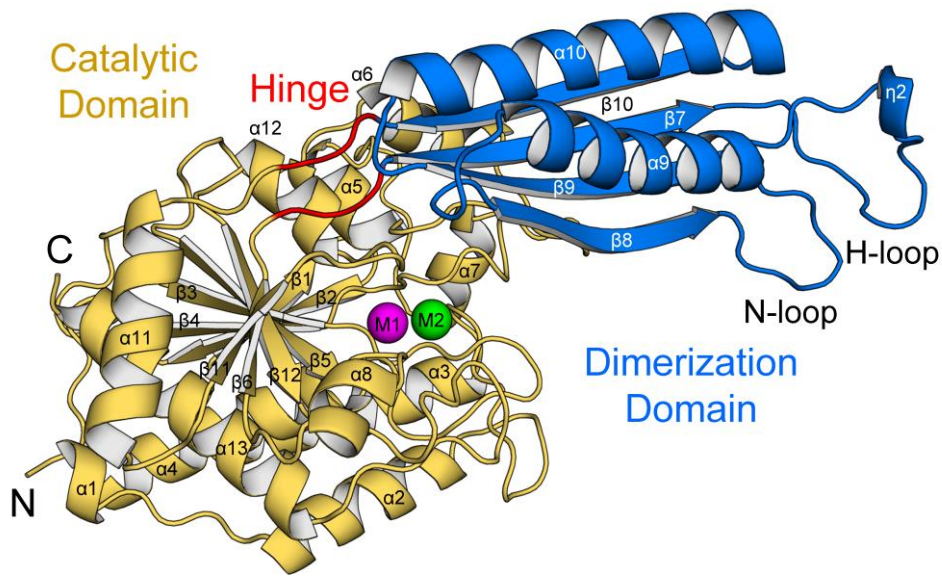


FIGURE 20. **Overall structure of AtUAH monomer.** The monomer structure of AtUAH is shown in different color codes. Magenta and green spheres represent the two metal ions in the active site.

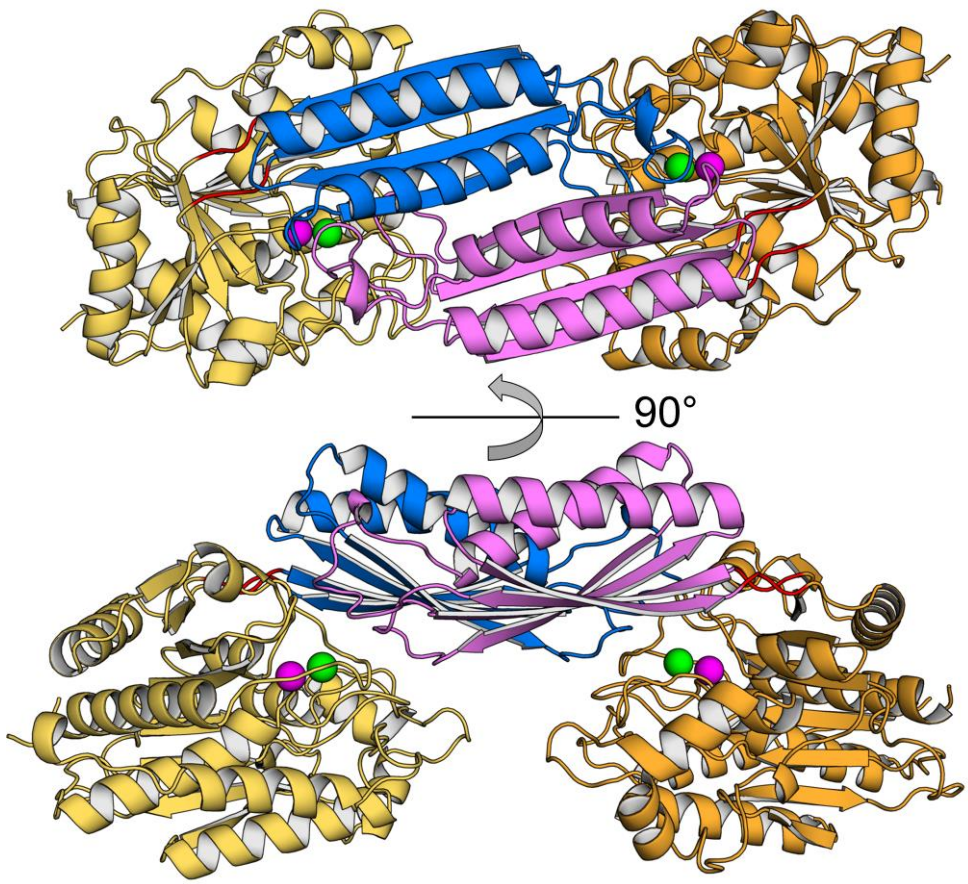


FIGURE 21. **The dimer of AtUAH.** The dimer in an asymmetric unit is shown in two different orientations. The color codes in Fig. 20 are used in one monomer, and the dimerization domain of the other monomer is indicated in magenta.

2. Active Site with a Bimetal Center

Identification of the gene for AtUAH was initially facilitated by a high sequence homology with AAH, in which a bimetal reaction center was previously observed in its crystal structure (Fig. 19) (15, 16). Accordingly, preliminary analyses of AtUAH activity indicated that a Mn^{2+} ion is required for maximum catalytic activity (Fig. 22). Based on these functional analyses, crystallization of AtUAH was performed in the presence of 1 mM $MnCl_2$.

The bimetal reaction center of AtUAH is indeed located at the pocket enclosed by the catalytic domain, dimerization domain, and other structural elements from a different subunit (Figs. 20 and 21). Five residues originating from the catalytic domain and four water molecules participate directly in the bimetal coordination shell within a distance of 2.4 Å (Fig. 23). Metal-1 (M1, magenta) and metal-2 (M2, green), which are possibly occupied by Mn^{2+} ions, are separated by a distance of 3.8 Å, where M2 is closer to the domain interface. In the bimetal coordination shell, Asp-149 and a water molecule bridge the two metal ions, while M1 and M2 form octahedral and hemi-octahedral coordinations, respectively. In the coordination shell of M1, an equatorial plane is formed by a water molecule (WatEQ1), a bridging water (WatBR), Asp-149, and His-254. Axial ligands to M1 are His-138 and a water molecule (WatAX1) *trans* to His-138, which results in an octahedral coordination. In contrast, a vacancy exists in the M2-coordination for an equatorial ligand (WatEQ2; open red circle in Fig. 23). Currently, the equatorial ligands to M2 consist of Asp-149, His-448, and WatBR. Its axial coordination is mediated by Glu-184 and a water molecule (WatAX2) *across* from Glu-184, resulting in a hemi-octahedral coordination shell.

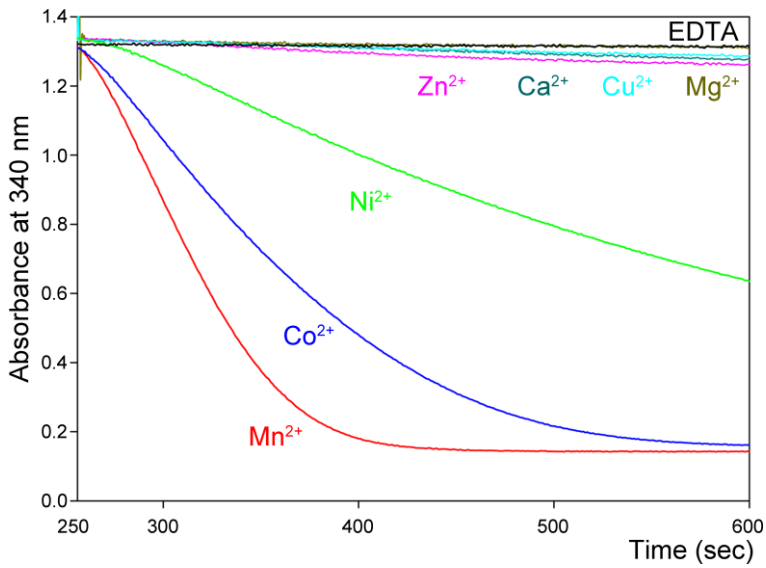


FIGURE 22. **Metal-dependency of AtUAH.** The metal-dependent reaction profile of AtUAH is shown. In the assay, 75 μM (*S*)-ureidoglycolate was used as a substrate. The solution contained 100 μM of each divalent metal and 10 mM EDTA was used as a control. Due to different concentrations of substrate, the degree of absorbance reduction at 340 nm are different from Fig. 2.

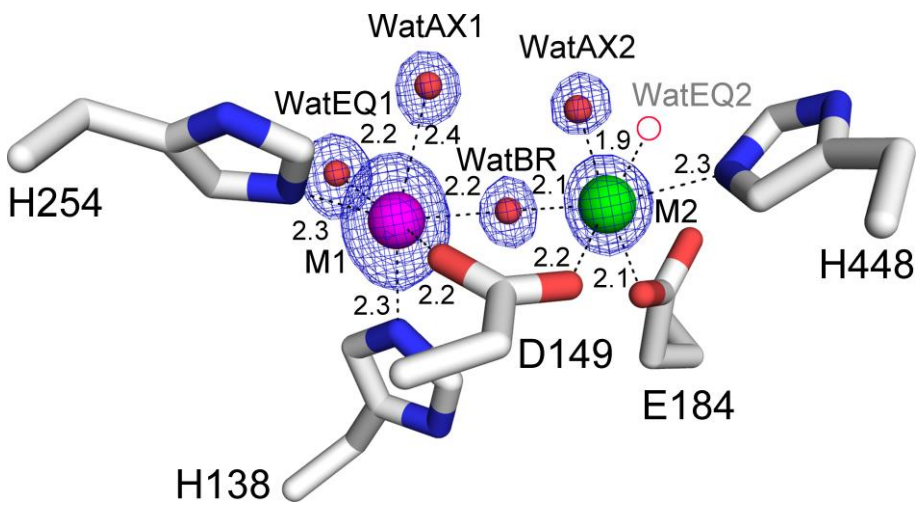


FIGURE 23. **Bimetal reaction center in AtUAH.** The bimetal reaction center, consisting of two metal ions and nearby residues involved in the metal-coordination shells, in the AtUAH–glyoxylate binary complex is shown. Interatomic distances are indicated in Å. Water molecules are indicated by red spheres, except for the missing WatEQ2. The metal ions and water molecules are overlaid with an F_o–F_c electron density map contoured at 10 σ.

3. Binding Mode of (S)-Ureidoglycolate in a Binary Complex with AtUAH(E183A)

The three binary complexes of AtUAH(E183A), with either (*S*)-uridoglycolate, (*S*)-hydroxyglycine, or glyoxylate, are structurally identical within an rmsd of 0.16–0.22 Å for 422 or 423 Cα atoms in the monomer. Therefore, this report focuses on the binding environment of each ligand to infer the substrate specificity of AtUAH.

The active site includes a bimetal center and is composed of residues from the catalytic and dimerization domains, as well as H- and N-loops from another subunit. The crystal structure of the AtUAH(E183A)–ureidoglycolate complex revealed the binding mode of ureidoglycolate in an (*S*)-configuration, consistent with the enantioselectivity of AtUAH (Figs. 24 and 25) (15). (*S*)-Ureidoglycolate is anchored to a bimetal center in each subunit with its ureido group and the hydroxyl group at the chiral carbon coordinated to the two metal ions. Specifically, the positions of the amino group and the carbonyl oxygen in the ureido group are assigned based on the binding environment with surrounding residues. The carboxyl group is far from a bimetal center, forming a dihedral angle of 112.8° (±2.5) with the ureido group. The binding of (*S*)-ureidoglycolate is achieved by replacing the two axial water molecules (*i.e.*, WatAX1 and WatAX2) in the bimetal center with the nitrogen and carbonyl oxygen atoms of the ureido group, respectively (Figs. 23 and 24). In addition, the hydroxyl oxygen attached to the chiral center of (*S*)-ureidoglycolate provides a ligand for the vacant equatorial position of M2, completing the octahedral coordination of M2.

The binding of (*S*)-ureidoglycolate in the complex is further stabilized by various interactions with nearby residues (Fig. 25). The most notable of these is a

possible stacking interaction between the carboxyl group of the substrate and the aromatic ring of Tyr-423, which dictates the relative orientation of (*S*)-ureidoglycolate in the active site. In addition, the negative charge on the carboxyl group of (*S*)-ureidoglycolate is possibly neutralized by Arg-353, which is held in its position by Asp-351. The carboxyl group is further stabilized by a hydrogen bonds with the main chain nitrogen of Tyr-423 (3.0 Å distant), and the side chain residues His-290* (3.5 Å) and Asn-340* (3.0 Å). The hydroxyl group of (*S*)-ureidoglycolate that participates in the M2 coordination is also located within 3.0 Å of the side chains of Glu-184 and Asn-340*. The ureido tail of (*S*)-ureidoglycolate also forms a possible bidentate hydrogen bond with Gln-257 (2.8 Å) and is further stabilized by interacting with the main chain carbonyl oxygen of Tyr-423 (3.1 Å) and WatBR (2.8 Å), as well as with the side chains of His-424 (3.1 Å) and His-290* (2.6 Å).

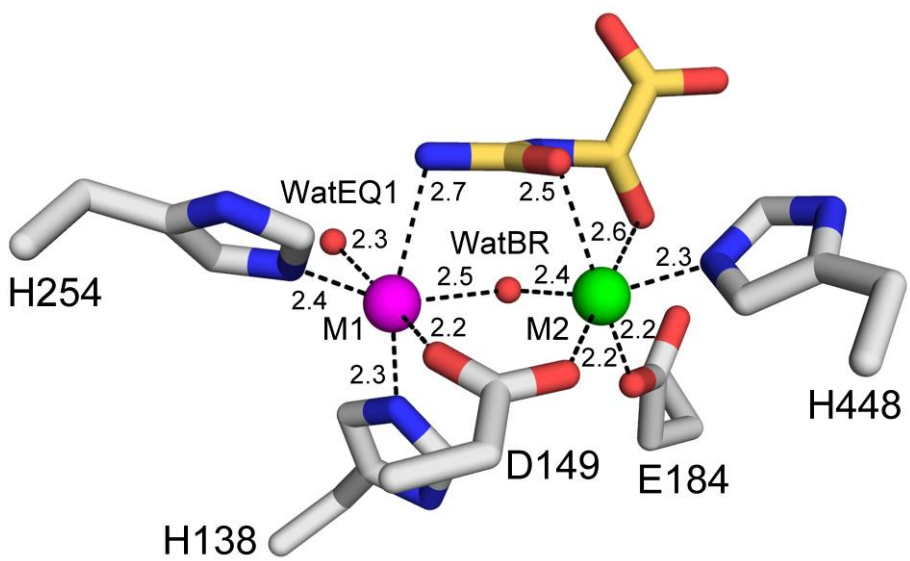


FIGURE 24. **The metal coordination shell in the AtUAH(E183A)-(*S*)-ureidoglycolate complex.** The binding mode of (*S*)-ureidoglycolate is shown in the AtUAH(E183A)–ureidoglycolate complex. Interatomic distances are indicated in Å . The axial water molecules (WatAX1 and WatAX2) and the missing equatorial water (WatEQ2) in the AtUAH–glyoxylate binary complex (Fig. 23) were replaced by atoms from (*S*)-ureidoglycolate, thus completing the octahedral coordination of each metal ion.

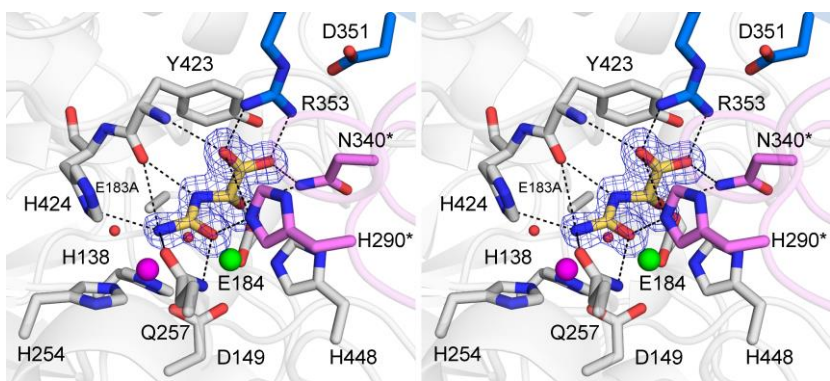


FIGURE 25. **The binding mode of (*S*)-ureidoglycolate in the binary complex with AtUAH(E183A).** The active site of the AtUAH(E183A)–ureidoglycolate complex is displayed in a stereoscopic view. Possible hydrogen bonds within 3.6 Å are indicated by dashed lines and the modeled (*S*)-ureidoglycolate is overlaid with an F_o–F_c electron density map contoured at 4 σ. Different color codes are used for residues from the catalytic domain (gray), dimerization domain (blue), and the adjacent monomer (magenta).

4. Binding Mode of the Intermediate and Product in the Binary Complex with AtUAH

A proposed reaction intermediate, (*S*)-hydroxyglycine, was unexpectedly identified in the crystal structure of AtUAH(E183A) when co-crystallized with (*S*)-ureidoglycolate. Crystallization conditions of this intermediate differed from those of the AtUAH(E183A)–ureidoglycolate complex (see MATERIALS AND METHODS). (*S*)-hydroxyglycine was observed in the active site of AtUAH(E183A)–hydroxyglycine, with an unambiguous density corresponding to a hydroxyl group on the chiral carbon atom (Fig. 26). The binding mode of (*S*)-hydroxyglycine is almost identical to that of (*S*)-ureidoglycolate (Fig. 25), except for absence of those interactions mediated by the ureido group in (*S*)-ureidoglycolate. Thus, axial water molecules in WatAX1 and WatAX2 coordinate M1 and M2, respectively. The hydroxyl group of (*S*)-hydroxyglycine is also bound in the M2-coordination shell as an equatorial ligand as in the AtUAH(E183A)–ureidoglycolate complex. Other interactions associated with the carboxyl group of (*S*)-ureidoglycolate are maintained with (*S*)-hydroxyglycine.

In the crystal structure of WT AtUAH complexed with glyoxylate, an electron density corresponding to the sp^2 planar form of glyoxylate was observed in the active site (Fig. 27). Both the binding site and mode of binding are reminiscent of those of (*S*)-hydroxyglycine. The only difference is that the carbonyl group of glyoxylate replaces the amino group of (*S*)-hydroxyglycine and is therefore within a hydrogen bonding distance of 2.8 Å to the main chain carbonyl oxygen of Tyr-423. The detailed features of the bimetal center binding are described above. The carboxyl group is common to all three of the evaluated complexes, and the interactions associated with this moiety are maintained, suggesting that those

interactions are crucial for substrate specificity. In particular, Tyr-423 acts as a molecular pillow and allows stacking interactions with the carboxyl group of the various ligands (Figs. 25, 26 and 27).

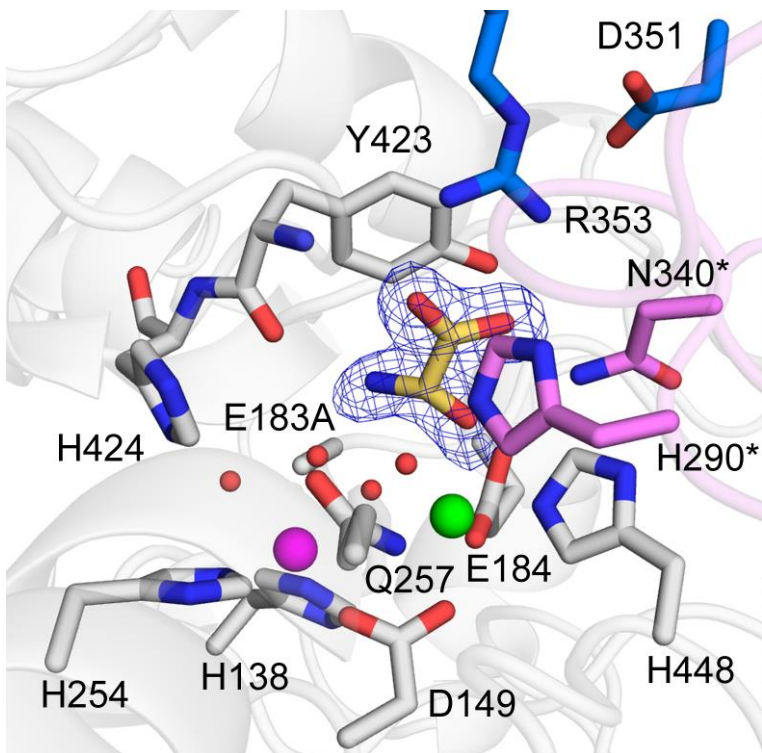


FIGURE 26. **The binding modes of the reaction intermediate in the binary complex with AtUAH(E183A).** An orientation similar to that shown in Fig. 25 is presented for the AtUAH(E183A)–hydroxyglycine complex. The modeled (*S*)-hydroxyglycine is overlaid with an F_o–F_c electron density map contoured at 4 σ.

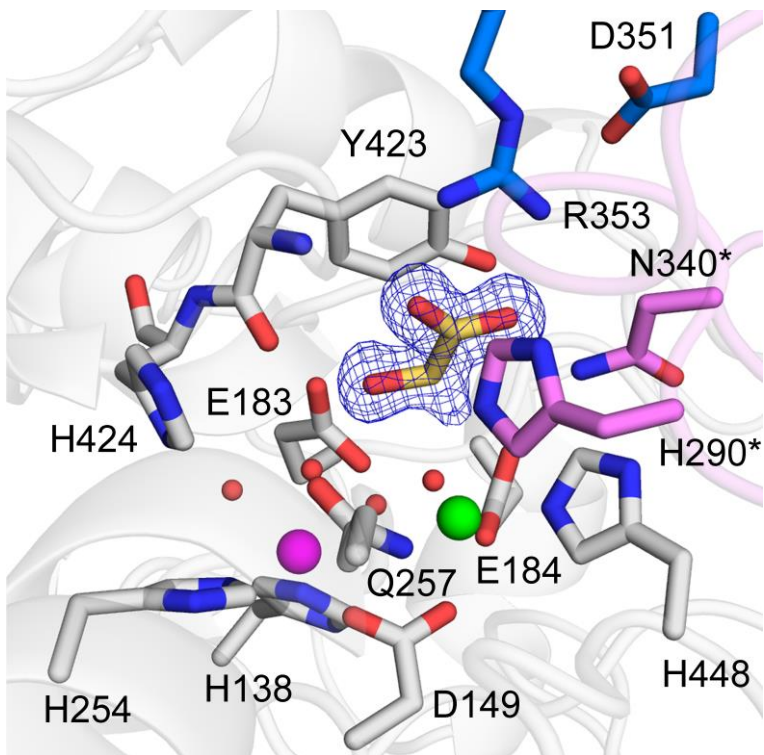


FIGURE 27. **The binding modes of product in the binary complex with AtUAH.** The active site of the AtUAH–glyoxylate is shown with an F_o–F_c electron density map of glyoxylate contoured at 4 σ.

5. Binding Mode of Allantoate in the Closed Conformation of the EcAAH Binary Complex

The crystal structure of EcAAH complexed with allantoate has been previously characterized using the WT enzyme (16) (PDB code: 1Z2L; hereafter this known EcAAH structure is referred to as EcAAH(1Z2L)). Compared with AtUAH, the binding site of allantoate in EcAAH(1Z2L) differs significantly from that of (*S*)-ureidoglycolate, (*S*)-hydroxyglycine, and glyoxylate. AtUAH and AAH are homologous in sequence (30% identity; Fig. 19) and structure (*Z*-score of 42.4), including the bimetal center. The substrates allantoate and (*S*)-ureidoglycolate, for AAH and UAH, respectively, differ by the presence of an additional ureido group in allantoate. Therefore, the structure of EcAAH complexed with allantoate was reexamined using a catalytically incompetent EcAAH(E126A) mutant to validate the binding site of allantoate.

The conformation of the EcAAH(E126A)–allantoate complex in this study differs from that of EcAAH(1Z2L), with an rmsd of 5.0 Å for 411 Cα atoms in the monomer. Specifically, a large movement of the catalytic domain occurs relative to the dimerization domain between the two structures (Fig. 28). The EcAAH(E126A)–allantoate complex adopts a closed conformation while EcAAH(1Z2L) presents an open conformation. An unambiguous electron density for allantoate, including two ureido groups, was identified in the active site (Fig. 29). However, under the experimental conditions, only one metal ion, corresponding to M2 in AtUAH, was identified. In general, except for the *pro-S* ureido group, the binding mode of allantoate is similar to that of (*S*)-ureidoglycolate in the AtUAH(E183A)-ureidoglycolate complex (Fig. 25). One of the ureido groups in allantoate is anchored to M2, but the other *pro-S* ureido group,

a chemical moiety absent in (*S*)-ureidoglycolate, points toward the innermost side of the active site. This *pro-S* ureido group could be accommodated in EcAAH by the fact that a residue structurally equivalent to Tyr-423 in AtUAH is represented by a glycine, *i.e.*, Gly-357, in EcAAH. Without this sequence change, steric hindrance in EcAAH would interfere with the additional ureido group of allantoate. The main chain nitrogen and carbonyl oxygen atoms of Gly-357 also mediate possible hydrogen bonds with the carbonyl oxygen of the *pro-S* ureido group. The *pro-S* ureido group is further stabilized by interactions with Gln-213 and Arg-215, which are strictly conserved as the QXR motif in AAH enzymes (Fig.19) (15). Except for these unique interactions associated with the *pro-S* ureido group, the interactions between allantoate and EcAAH are very similar to those between (*S*)-ureidoglycolate and AtUAH.

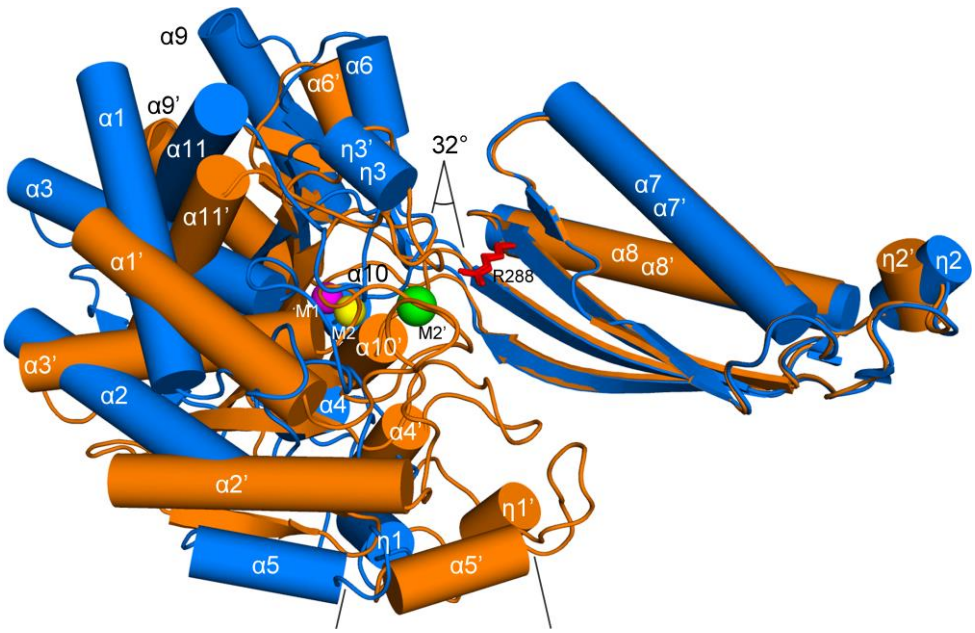


FIGURE 28. **Conformational changes in the EcAAH(E126A)–allantoate complex.** EcAAH(1Z2L; blue) and the EcAAH(E126A)–allantoate complex (brown) are compared. Note that large conformational changes in a catalytic domain occur in rigid-body movement relative to the dimerization domain. The di-zinc metal center of EcAAH(1Z2L) are represented by magenta and yellow spheres, while metal-2 (M2') of EcAAH(E126A)–allantoate is indicated by a green sphere. A superscript comma is used to label the secondary structure elements of the EcAAH(E126A)–allantoate complex.

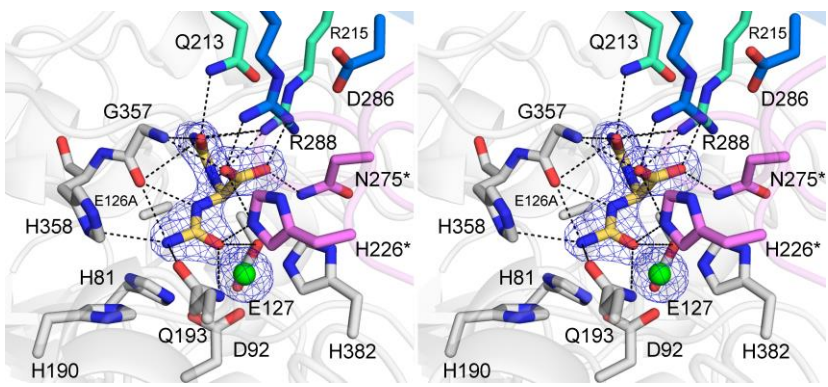


FIGURE 29. **The binding mode of allantoate in the EcAAH(E126A)–allantoate complex.** The active site of the EcAAH(E126A)–allantoate binary complex is shown in a stereoscopic view. The modeled allantoate and the metal ion are overlaid with an F_o–F_c electron density map contoured at 4 σ and 10 σ, respectively. Details of the figure are essentially identical to those of Fig. 25 and glutamine and arginine residues of the QXR motif are colored in cyan.

6. Functional Analyses

Steady-state kinetic analyses of WT AtUAH and various mutant enzymes were performed to characterize the functional role(s) of the active site residues. Briefly, the release of ammonia as a reaction product was monitored by a coupled reaction in which an NADPH-dependent glutamate dehydrogenase incorporates ammonia into α-ketoglutarate with a concomitant formation of $NADP^+$, causing changes in absorbance at 340 nm.

Based on structural analyses, 13 residues in AtUAH were selected for site-directed mutagenesis. The kinetic parameters are summarized in Table 9. Mutation of residues directly coordinating to the bimetal center, such as H138A, D149A, D149N, E184A, H254A, and H448A, caused the enzyme to become catalytically incompetent, suggesting an essential role for each of these residues. Other residues possibly involved in ligand binding were also subject to mutation. In particular, Glu-183, Arg-353, and Asp-351 were crucial for AtUAH activity, consistent with their proposed structural roles, either by interacting directly with an incoming substrate, as in the cases of Glu-183 and Arg-353, or by stabilizing Arg-353 to mediate interactions with a ligand, as in the case of Asp-351 (Fig. 25 and Table 9). Mutants including Y423F, Q257N, H424N, and H290Q exhibited some enzymatic activity. Specifically, while the Y423A and Y423G mutant enzymes were completely inactive, the K_m value of Y423F is approximately twofold greater than that of the WT enzyme. The k_{cat} value of Y423F was also ~41% that of the WT enzyme. These data indicate that the suggested stacking interaction between the carboxyl group of the ligand and Tyr-423 plays a pivotal role in substrate binding. Other residues interacting with the substrate were also crucial for AtUAH activity (Fig. 25). The K_m value of Q257N is ~64-fold higher than, and its k_{cat} value is

~3.5%, that of the WT enzyme. Similar to Gln-257, residues such as His-290, Asn-340, and His-424 also play major roles in determining enzyme activity, possibly by facilitating binding of the substrate. The K_m values of H290Q and H424N are 75- and 3-fold higher, respectively, than that of WT AtUAH, while the K_m value of N340A was more than 10 mM. The H290Q and H424N mutants also had k_{cat} values that were ~27% and ~16% that of the WT enzyme, resulting in catalytic efficiencies of less than ~6% that of WT AtUAH.

TABLE 9. Kinetic parameters of AtUAH and its mutants[a]

	K_m (μM)	k_{cat} (sec^{-1})	k_{cat}/K_m ($\mu M^{-1}sec^{-1}$)
WT	11.29(0.49) [b]	1.43(0.019)	0.127
Y423F	20.62(1.39)	0.59(0.013)	0.0284
H424N	32.81(1.37)	0.23(0.0036)	0.00704
Q257N	721.35(110.96)	0.05(0.0031)	0.00008
H290Q	844.36(78.21)	0.37(0.014)	0.00043
N340A	$> 10^4$	N/A	N/A

[a]Mutants as follows were inactive: H138A, D149A, D149N, E184A, H254A, H448A, H290N, H290A, E183Q, E183D, E183A, R353K, R353A, D351A, N340D, Y423A, Q257E, Q257A.
[b]Values in parentheses are standard errors estimated by the program SigmaPlot.

DISCUSSION

Enzymes in the ureide pathway, including AAH, UGlyAH, and UAH, catalyze consecutive enzyme reactions (15). Since this pathway constitutes a major purine catabolism route in plants, as well as in some bacteria and fungi, strict requirements for substrate specificity are important for recovering nitrogen from allantoin. In plants, the three aforementioned enzymes are co-localized in the endoplasmic reticulum, and AAH and UAH are highly homologous in their sequences, and both contain a bimetal reaction center. Given their structural similarities and localization, substrate specificity is a major concern.

Following the structural and enzyme kinetic analyses of AtUAH, and a structural comparison with the EcAAH–allantoate binary complex, two notable features were identified in the active site of both enzymes: the size of the active site and the absence of a QXR motif in UAH. The active site of AtUAH is smaller than that of EcAAH, mainly due to the presence of Tyr-423 in AtUAH. Tyr-423 should induce steric hindrance that would prevent entrance of the *pro-S* ureido group of allantoate into the active site, thereby serving as a key structural element that defines the substrate specificity of UAH for (*S*)-ureidoglycolate. Accordingly, this particular tyrosine residue is strictly conserved in members of the UAH family, while a structurally equivalent glycine is present as an invariant residue in the AAH family (Fig. 19). The glycine is able to accommodate the *pro-S* ureido group of allantoate. However, mutagenesis study suggests other roles of Tyr-423 beyond as just a steric barrier preventing allantoate from occupying the active site of UAH. The observation that AtUAH(Y423A) mutant was completely inactive toward (*S*)-ureidoglycolate (Table 2) implicates a more subtle role of Tyr-423, because the

enlarged active site in AtUAH(Y423A) mutant should accommodate the substrate in the active site. In contrast, Y423F mutant recovered enzyme activity, although the efficiency (k_{cat}/K_m) was about sixfold lower than that of WT AtUAH (Table 9). Given these observations, we concluded that the aromatic ring of Tyr-423 constrains the torsional angle of the C-N bond in (*S*)-ureidoglycolate connecting the carboxyl group and the ureido group, leading to a conformer of (*S*)-ureidoglycolate optimal for interacting with Arg353 and other nearby residues (Fig. 25). It is possible that the absence of Tyr-423 could result in a non-productive binding mode of (*S*)-ureidoglycolate and its carboxyl group may not form interactions effectively with Arg353. Consistent with this proposal, the AtUAH R353A and R353K were totally inactive, indicating that distance between Arg353 and the carboxyl group in (*S*)-ureidoglycolate was also crucial in activity. Furthermore, the inactive D351A mutant (Table 9) suggests that Asp351 plays a role in positioning the side chain of Arg353 in a catalytically active orientation (Fig. 25). Therefore, structural constraints by Tyr-423 and the resulting interaction geometry between the carboxyl group of substrate and Arg353 appears to dictate the productive binding mode of substrate and serves as a determinant for substrate specificity. This proposal that substrate specificity in UAH is a function of interactions more complex than those conferred by a single active site residue was also evaluated. There was also an attempt to confer an AAH-like activity on AtUAH by redesigning the size of the active site, but without success. Specifically, the AtUAH Y423G mutant was still inert to allantoate (Fig. 30), suggesting that modifying the substrate specificity of UAH cannot be achieved simply by relieving the steric hindrance of Tyr-423 to the incoming *pro-S* ureido group in allantoate. In addition, the EcAAH–allantoate binary complex contains a QXR motif present

only in members of the AAH family (Figs. 19 and 29), which results in unique interactions associated with the *pro-S* ureido group in allantoate. Previous structural analyses have suggested that the QXR motif is required for the binding of the sulfate ion, which was thought to act as an allosteric effector for stabilizing substrate binding (15, 16). The current study, however, revealed that the QXR motif is directly involved in substrate binding through interactions with the *pro-S* ureido group in allantoate (Fig. 29). Therefore, the QXR motif also helps determine the substrate specificity of AAH.

Recently, gene context analyses of the ureide pathway unraveled a functional divergence within the UGlyAH family (Fig. 1) (54), which contains two groups of enzymes having distinct substrate specificity for allantoate and (*S*)-ureidoglycine, respectively. In the study, differences in the active site size were identified by variations in the chemical identity of residues proximal to the *pro-S* ureido group of allantoate, eventually resulting in different substrate specificity. Furthermore, a single mutation was insufficient to convert the functional features of those enzymes.

Ligand-induced conformational changes are well characterized in the amidase, hydantoinase/carbamoylase families, to which AAH and UAH belong. In contrast to the structure of EcAAH(1Z2L), the EcAAH binary complex adopts a closed conformation that results from a 32° rigid body movement of the catalytic domain (Fig. 28). These conformational changes entail an increase in the buried surface area from 5790 to 6140 $Å^2$ (55). More importantly, the change also moves residues in the active site closer to the substrate. In particular, Arg-288 in the open conformation of EcAAH(1Z2L) (16), which is structurally equivalent to Arg-353 in UAH, locates more than ~ 12 Å from a bimetal reaction center. The binding of allantoate brings the catalytic domain into the active site, and Arg-288 is

repositioned to ~7 Å from a bimetal reaction center.

From a mechanistic perspective, AtUAH likely employs a reaction mechanism essentially identical to those of other structural homologs in the peptidase M20 family (16, 48–53, 56). Specifically, the first step in the reaction mechanism is the activation of a water molecule that bridges two metal ions in the bimetal reaction center. This bridged water molecule is reportedly deprotonated by a strictly conserved glutamate residue, corresponding to Glu-183 of AtUAH. The resulting hydroxide attacks the carbonyl carbon of the ureido group that is anchored to the metal ions, forming a tetrahedral oxyanionic intermediate. Through a subsequent electron rearrangement, the scissile C–N amide bond in a carbamoyl group is cleaved, generating (*S*)-hydroxyglycine and carbamate (Fig. 31). Since carbamate is known to be labile (4, 56) as well as (*S*)-hydroxyglycine, these two compounds might undergo spontaneous decomposition producing one carbon dioxide, two molecules of ammonia, and one glyoxylate.

Three crystal structures of AtUAH were characterized in the current study, each in a distinct ligation state. A crystal structure of EcAAH complexed with allantoate was also elucidated. The binding modes of the (*S*)-ureidoglycolate substrate, a reaction intermediate, and a glyoxylate product were identified in a bimetal reaction center. In each of the three structures of AtUAH, Tyr-423, which is conserved in the UAH family, causes structural constraints of the ligands and results in their productive binding in the active site by interacting with the nearby residues. In contrast, a glycine residue structurally equivalent to Tyr-423 in UAH is present in AAH, which accommodates the *pro-S* ureido group of allantoate. The binding of the *pro-S* ureido group is mediated by the QXR motif unique to the AAH family. These differences in the active site size and the interaction

environments act as structural determinants for substrate specificity of UAH.

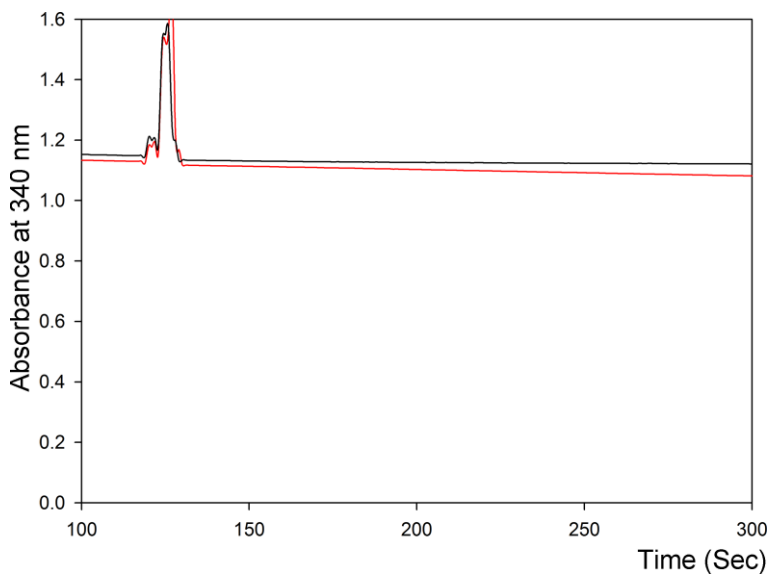


FIGURE 30. **No AAH activity of AtUAH Y423G mutant.** The progress curves for AtUAH Y423G mutant (red) and buffer as a control (black) are shown. In this experiment, 2mM allantoate was used to investigate whether AtUAH Y423G gains AAH activity.

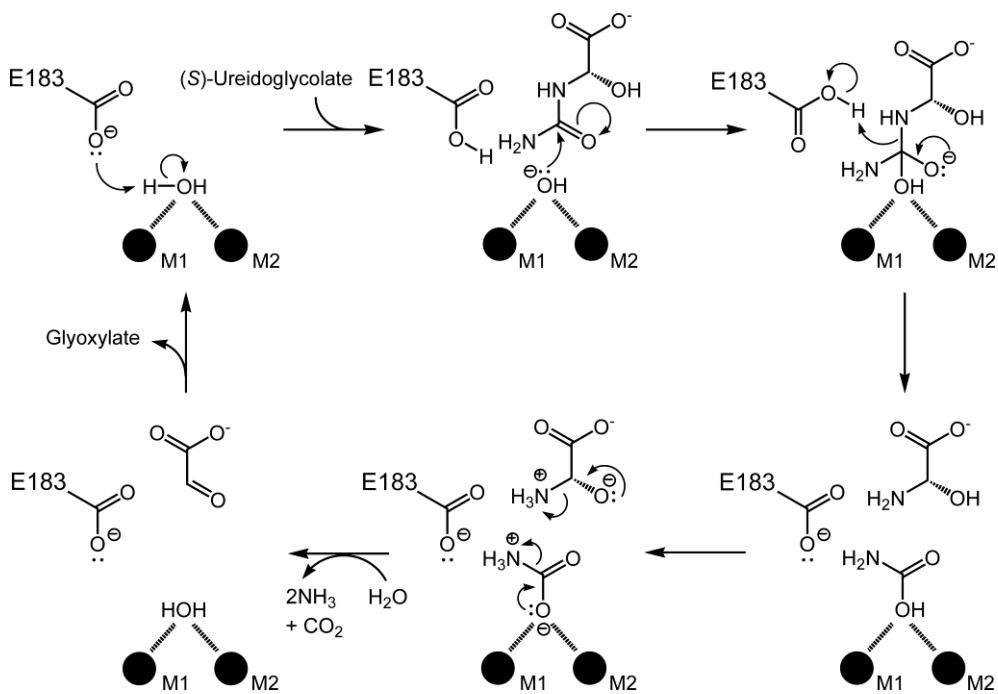


FIGURE 31. **Proposed mechanism for AtUAH.** Glu-183 is the proposed catalytic base. Enzyme reaction begins with the deprotonation of a bridging water molecule bridging by this conserved glutamate residue. (*S*)-hydroxyglycine and carbamate probably undergo spontaneous decomposition thus producing one moleculae of glyoxylate and two molecules of ammonia.

CHAPTER IV. CONCLUSIONS

The ureide pathway as a metabolic route for purine catabolism plays an important role in nitrogen remobilization. In the second phase of the pathway, the purine ring nitrogen atoms are released as ammonia by three consecutive enzymatic reactions consisted of AAH, UGlyAH and UAH. In this study, two enzymes, UGlyAH and UAH were focused and their crystal structures were determined. UGlyAH converts (*S*)-ureidoglycine into (*S*)-ureidoglycolate using Mn^{2+} ion as a cofactor. The crystal structure of the ligand-free form of AtUGlyAH revealed that Mn^{2+} ion binding active site is located in the C-domain of monomeric bi-cupin fold and homo-octamer constitutes the functional unit of the enzyme. The crystal structure of AtUGlyAH in complex with substrate (*S*)-ureidoglycine revealed its unique binding mode of (*S*)-ureidoglycine in the Mn^{2+} coordination shell which dictates the enantioselectivity of the reaction. Based on the structural analysis, substrate binding and putative catalytic residues were subjected to site-directed mutagenesis and their functional roles were investigated. The enzyme assay coupled with *E. coli* (*S*)-ureidoglycolate dehydrogenase enabled the measurement of the steady-state kinetic parameters of AtUGlyAH. The results of this study provide the molecular insights into the structure of AtUGlyAH and its mechanism. UAH is the last enzyme of the pathway in plants which catalyzes the conversion of (*S*)-ureidoglycolate into glyoxylate. UAH is similar to AAH in sequence and structure. But, both enzymes display strict substrate specificities. Three crystal structures of AtUAH, each with a distinct ligation state as well as a crystal structure of EcAAH

in complex with its substrate were determined. Those crystal structures revealed the binding mode of each ligand in the bimetal active site. Structural analyses elucidated that all the four structures adopt the closed conformation thus enabling the direct comparison between the active site of UAH and AAH. The strictly conserved residue, Tyr-423 of AtUAH could be pinpointed as the key residue defining the active site size while the tyrosine residue of UAH is changed to glycine in AAH allowing the accommodation of allantoate with the *pro-S* ureido group. Furthermore, the role of the conserved QXR motif could be redefined as a structural element directly interacting with the *pro-S* ureido group of allantoate. This study provides the molecular insights into the structure of AtUAH and the reason of substrate specificity between two homologous enzymes, UAH and AAH.

REFERENCES

1. Atkins, C. A., and Smith, P. M. C. (2000) in *Prokaryotic Nitrogen Fixation: A Model System for Analysis of a Biological Process* (Triplett, E. W., ed.) pp. 559-587, Horizon Scientific Press, Wymondham, UK
2. Smith, P. M., and Atkins, C. A. (2002) Purine biosynthesis. Big in cell division, even bigger in nitrogen assimilation. *Plant Physiol.* **128**, 793-802
3. Boldt, R., and Zrenner, R. (2003) Purine and pyrimidine biosynthesis in higher plants. *Physiol. Plant* **117**, 297-304
4. Werner, A. K., and Witte, C. P. (2011) The biochemistry of nitrogen mobilization. Purine ring catabolism. *Trends Plant Sci.* **16**, 381-387
5. Daniel-Vedele, F., Filleur, S., and Caboche, M. (1998) Nitrate transport: a key step in nitrate assimilation. *Curr. Opin. Plant Biol.* **1**, 235-239
6. Ludewig, U., Neuhauser, B., and Dynowski, M. (2007) Molecular mechanisms of ammonium transport and accumulation in plants. *FEBS Lett.* **581**, 2301-2308
7. Lea, P. J., and Miflin, B. J. (1974) An alternative route for nitrogen assimilation in plants. *Nature* **251**, 680-685
8. Melo-Oliveira, R., Cinha-Oliveria, I., and Coruzzi, G.M. (1996). Arabidopsis mutant analysis and gene regulation define a non-redundant role for glutamate dehydrogenase in nitrogen assimilation. *Proc. Natl. Acad. Sci. USA* **96**, 4718-4723
9. Zrenner, R., Stitt, M., Sonnewald, U., and Boldt, R. (2006) Pyrimidine and purine biosynthesis and degradation in plants. *Annu. Rev. Plant Biol.* **57**, 805-

836

10. Vogels, G. D., and Van der Drift, C. (1976) Degradation of purines and pyrimidines by microorganisms. *Bacteriol. Rev.* **40**, 403-468
11. Loh, K. D., Gyaneshwar, P., Markenscoff Papadimitriou E., Fong, R., Kim, K. S., Parales, R., Zhou, Z., Inwood, W., and Kustu, S. (2006) A previously undescribed pathway for pyrimidine catabolism. *Proc. Natl Acad. Sci. USA* **103**, 5114-5119
12. Kim, K. S., Pelton, J. G., Inwood, W. B., Andersen, U., Kustu, S., and Wemmer, D. E. (2010) The Rut pathway for pyrimidine degradation. Novel chemistry and toxicity problems. *J. Bacteriol.* **192**, 4089-4102
13. Ramazzina, I., Folli, C., Secchi, A., Berni, R., and Percudani, R. (2006) Completing the uric acid degradation pathway through phylogenetic comparison of whole genomes. *Nat. Chem. Biol.* **2**, 144–148
14. Serventi, F., Ramazzina, I., Lamberto, I., Puggioni, V., Gatti, R., and Percudani, R. (2010) Chemical basis of nitrogen recovery through the ureide pathway. Formation and hydrolysis of *S*-ureidoglycine in plants and bacteria. *ACS Chem. Biol.* **5**, 203-214
15. Werner, A. K., Romeis, T., and Witte, C. P. (2010) Ureide catabolism in *Arabidopsis thaliana* and *Escherichia coli*. *Nat. Chem. Biol.* **6**, 19-21
16. Agarwal, R., Burley, S. K., and Swaminathan, S. (2007) Structural analysis of a ternary complex of allantoate amidohydrolase from *Escherichia coli* reveals its mechanics. *J. Mol. Biol.* **368**, 450-463
17. Cendron, L., Berni, R., Folli, C., Ramazzina, I., Percudani, R., and Zanotti, G.

(2007) The structure of 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase provides insights into the mechanism of uric acid degradation. *J. Biol. Chem.* **282**, 18182-18189

18. Kim, K., Park, J., and Rhee S. (2007) Structural and functional basis for (*S*)-allantoin formation in the ureide pathway. *J. Biol. Chem.* **282**, 23457-23464
19. Kim, K., Kim, M. I., Chung, J., Ahn, J., and Rhee, S. (2009) Crystal structure of metal-dependent allantoinase from *Escherichia coli. J. Mol. Biol.* **387**, 1067-1074
20. Ramazzina, I., Cendron, L., Folli, C., Berni, R., Monteverdi, D., Zanotti, G., and Percudani, R. (2008) Logical identification of an allantoinase analog (puuE) recruited from polysaccharide deacetylases. *J. Biol. Chem.* **283**, 23295-23304
21. Hennebry, S. C. (2009) Evolutionary changes to transthyretin. Structure and function of a transthyretin-like ancestral protein. *FEBS J.* **276**, 5367-5379
22. French, J. B., and Ealick, S. E. (2010) Structural and mechanistic studies on *Klebsiella pneumoniae* 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase. *J. Biol. Chem.* **285**, 35446-35454
23. Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G. E., Rasche, N., Lüder, F., Weckwerth, W., and Jahn, O. (2007) Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* **19**, 3170-3193
24. Lamberto, I., Percudani, R., Gatti, R., Folli, C., and Petrucco, S. **(**2010) Conserved alternative splicing of *Arabidopsis* transthyretin-like determines

protein localization and *S*-allantoin synthesis in peroxisomes. *Plant Cell* **22**, 1564-1574

25. Kim, M. I., Shin, I., Cho, S., Lee, J., and Rhee, S. (2012) Structural and functional insights into (*S*)-ureidoglycolate dehydrogenase, a metabolic branch point enzyme in nitrogen utilization. *PLoS One* **7**, e52066
26. Percudani, R., Carnevali, D., and Puggioni, V. (2013) Ureidoglycolate hydrolase, amidohydrolase, lyase: how errors in biological databases are incorporated in scientific papers and vice versa. *Database (Oxford)* **2013**, bat071
27. Ramazzina, I., Costa, R., Cendron, L., Berni, R., Peracchi, A., Zanotti, G., and Percudani, R. (2010) An aminotransferase branch point connects purine catabolism to amino acid recycling. *Nat. Chem. Biol.* **6**, 801-806
28. Winkler, R. G., Blevins, D. G., and Randall, D. D. (1988) Ureide catabolism in soybeans : III. Ureidoglycolate amidohydrolase and allantoate amidohydrolase are activities of an allantoate degrading enzyme complex. *Plant Physiol.* **86**, 1084-1088
29. Wells, X. E., and Lees, E. M. (1991) Ureidoglycolate amidohydrolase from developing French bean fruits (*Phaseolus vulgaris* [L.].). *Arch. Biochem. Biophys.* **287**, 151-159
30. Werner, A. K., Sparkes, I. A., Romeis, T., and Witte, C. P. (2008) Identification, biochemical characterization, and subcellular localization of allantoate amidohydrolases from Arabidopsis and soybean. *Plant Physiol.* **146**, 418-430
31. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data

collected in oscillation mode. *Methods Enzymol.* **276**, 307-326

32. Terwilliger, T. C., and Berendzen, J. (1999) Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**, 849-861
33. Terwilliger, T. C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 965-972
34. Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 38-44
35. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
36. Brünger, 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., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR system. A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905-921
37. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX. A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213-221
38. Chen, V. B., Arendall III, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010)

MolProbity. All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12-21

39. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235-242
40. Muratsubaki, H., Satake, K., and Enomoto, K. (2006) Enzymatic assay of allantoin in serum using allantoinase and allantoate amidohydrolase. *Anal. Biochem.* **359**, 161-166
41. Agarwal, G., Rajavel, M., Gopal, B., and Srinivasan, N. (2009) Structure-based phylogeny as a diagnostic for functional characterization of proteins with a cupin fold. *PLoS ONE* **4**, e5736
42. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript. Analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-308
43. Gravenmade, E. J., Vogels, G. D., and Van der Drift, C. (1970) Hydrolysis, racemization and absolute configuration of ureidoglycolate, a substrate of allantoicase. *Biochim. Biophys. Acta* **198**, 569-582
44. Holm, L., and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233**, 123-138
45. Kamat, S. S., Bagaria, A., Kumaran, D., Holmes-Hampton, G. P., Fan, H., Sali, A., Sauder, J. M., Burley, S. K., Lindahl, P. A., Swaminathan, S., and Raushel, F. M. (2011) Catalytic mechanism and three-dimensional structure of adenine

deaminase. *Biochemistry* **50**, 1917-1927

46. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105-124
47. Lundgren, S., Andersen, B., Piskur, J., and Dobritzsch, D. (2007) Crystal structures of yeast beta-alanine synthase complexes reveal the mode of substrate binding and large scale domain closure movements. *J. Biol. Chem.* **282**, 36037-36047
48. Martinez-Rodriguez, S., Garcia-Pino, A., Las Heras-Vazquez, F. J., Clemente-Jimenez, J. M., Rodriguez-Vico, F., Garcia-Ruiz, J. M., Loris, R., and Gavira, J. A. (2012) Mutational and structural analysis of L-N-carbamoylase reveals new insights into a peptidase M20/M25/M40 family member. *J. Bacteriol.* **194**, 5759-5768
49. Rowsell, S., Pauptit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M., and Brick, P. (1997) Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure* **5**, 337-347
50. Jozic, D., Bourenkow, G., Bartunik, H., Scholze, H., Dive, V., Henrich, B., Huber, R., Bode, W., and Maskos, K. (2002) Crystal structure of the dinuclear zinc aminopeptidase PepV from *Lactobacillus delbrueckii* unravels its preference for dipeptides. *Structure* **10**, 1097-1106
51. Unno, H., Yamashita, T., Ujita, S., Okumura, N., Otani, H., Okumura, A., Nagai, K., and Kusunoki, M. (2008) Structural basis for substrate recognition and hydrolysis by mouse carnosinase CN2. *J. Biol. Chem.* **283**, 27289-27299

52. Bitto, E., Bingman, C. A., Bittova, L., Houston, N. L., Boston, R. S., Fox, B. G., and Phillips, G. N., Jr. (2009) X-ray structure of ILL2, an auxin-conjugate amidohydrolase from *Arabidopsis thaliana*. *Proteins* **74**, 61-71

53. Nocek, B. P., Gillner, D. M., Fan, Y., Holz, R. C., and Joachimiak, A. (2010) Structural basis for catalysis by the mono- and dimetalated forms of the dapE-encoded N-succinyl-L,L-diaminopimelic acid desuccinylase. *J. Mol. Biol.* **397**, 617-626

54. Puggioni, V., Dondi, A., Folli, C., Shin, I., Rhee, S., and Percudani, R. (2014) Gene context analysis reveals functional divergence between hypothetically equivalent enzymes of the purine-ureide pathway. *Biochemistry* **53**, 735-745

55. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774-797

56. Lundgren, S., Gojkovic, Z., Piskur, J., and Dobritzsch, D. (2003) Yeast beta-alanine synthase shares a structural scaffold and origin with dizinc-dependent exopeptidases. *J. Biol. Chem.* **278**, 51851-51862

ABSTRACT IN KOREAN (국문요약)

식물에서 Ureide pathway는 퓨린 고리에 존재하는 질소 원자를 암모니아로 변환하는 대사경로로서 질소 회수에 중요한 역할을 담당한다. 본 연구에서는 X선 결정학을 이용하여 이 대사경로에 존재하는 두 효소 단백질의 구조를 규명하였다. 하나는 (*S*)-ureidoglycine aminohydrolase (UGlyAH)이고 다른 하나는 (*S*)-ureidoglycolate amidohydrolase (UAH)로서 두 효소 모두 애기장대로부터 유래한 것이다. UGlyAH 는 (*S*)-ureidoglycine을 이 대사경로 최후의 기질인 (*S*)-ureidoglycolate로 변환하며 이 과정에서 한 분자의 암모니아를 배출한다. 본 연구에서는 애기장대에서 유래한 UGlyAH (AtUGlyAH)의 구조와 기능연구를 통하여 다음과 같은 사실을 밝혔다. 먼저 기질이 결합되어 있지 않은 AtUGlyAH의 결정구조를 분석한 결과 monomer로서 베타 배럴로 이루어진 bi-cupin 폴드로 되어 있으며 octamer가 기능단위이고 Mn^{2+} 이온의 결합위치 또한 규명하였다. (*S*)-Ureidoglycine과 효소의 복합체 구조를 통하여 이 효소가 지니는 입체이성질체 선택적인 반응의 이유가 Mn^{2+} 이온이 분자 닻 (molecular anchor)의 역할을 하며 기질과 결합하는

고유의 결합 모드에 있음을 밝혔다. 구조정보를 바탕으로 한 반응속도 연구를 통하여 Mn^{2+} 이온과 (*S*)-ureidoglycine의 결합부위 주변의 아미노산 잔기들의 기능을 규명하였다. 이러한 연구 결과는 이 효소 단백질의 삼차원 구조와 가능한 촉매 작용기작에 대해 분자 수준에서의 통찰력을 제공해준다. 이 대사 경로의 마지막 단계에서, UAH는 (*S*)-ureidoglycolate를 glyoxylate와 두 분자의 암모니아로 변환하는 화학반응을 촉매한다. UAH는 이 대사경로의 상류에 존재하는 allantoate amidohydrolase (AAH)와 구조와 아미노산 서열에 있어서 매우 유사하다. UAH와 AAH는 amidase라는 비슷한 기능을 하지만 상호간의 엄격한 기질 특이성을 보이며 퓨린의 분해대사 경로에서 조화롭게 각자의 역할을 감당한다. 본 연구에서는 애기장대에서 유래한 UAH (AtUAH)의 기질, 중간체, 산물과의 각각의 복합체 결정 구조와 대장균에서 유래한 AAH (EcAAH)의 allantoate와의 복합체 결정 구조를 보고한다. 각각의 UAH 복합체 결정구조는 bimetal 센터에서 각 리간드의 결합모드를 closed conformation 상태에서 보여준다. 리간드는 두 금속 이온의 coordination shell에 직접 참여하는 동시에 주변의 아미노산 잔기에 의해 안정화 된다. UAH와 유사한 기질 결합 부위를 보여주는 AAH는

UAH와는 대조적으로 기질인 allantoate에 존재하는 추가적인 ureido group 때문에 UAH 보다 더 큰 활성 부위를 요구한다. AAH와 UAH의 구조와 기능연구는 두 효소 모두 리간드의 결합에 반응하여 열린 상태에서 닫히는 상태의 conformation의 변화를 경험한다는 것을 보여주었고, UAH와 AAH에서의 활성부위의 크기 차이와 각 기질이 상호작용하는 활성부위 내 환경이 구조적으로 매우 유사한 두 효소의 기질 특이성을 결정짓는 요소임을 보여준다.

주요어: 효소 반응속도론, 금속효소, 단백질 구조, 질소 대사, 퓨린 분해, Ureide pathway, 기질 특이성

학번: 2008-30927