Australian National University

Structure-function study of the enzymes of the cyanuric acid catabolic pathways

A Thesis Submitted

for the Degree of Doctor of Philosophy

of

The Australian National University

Research School of Chemistry

June 2018

Declaration

This research constitutes an original work carried out by myself during my PhD from February 2015 to June 2018, unless otherwise stated. Results and methods published or obtained by another person have been duly acknowledged.

The work presented herein has not been submitted as part of any other degree.

Lygie Esquirol

June 2018

Acknowledgements

I would like to thank sincerely Dr Colin Scott, Dr Matt Wilding and Dr Carol Hartley for their support and encouragements during my PhD, for being patient, positive and for providing support and precious advice during my weekly meetings.

I would like to express my appreciation to Pr Chris Easton for co-supervising my work and for being there to discuss my issues each time I needed it. Thanks a lot to Dr Hideki Onagi for helping me find ways to monitor all my many water-sensitive unstable substrates and products.

A big thank you to Nigel French for teaching me how to express and purify proteins. Without his insight, AtzE would probably still be a recalcitrant enzyme impossible to produce heterelogously!

Thanks a lot to Dr Janet Newman and Dr Tom Peat for being amazingly enthusiastic, persistent and so efficient at crystallising and solving the structures of my enzymes and uncovering the identity of unexpected small proteins.

Thanks a lot to Dr Jian-Wei Liu for providing help with proteomics to confirm the identity of my purified enzymes when I needed it and for revealing the true name of my many unknown bacterial strains.

Thanks you to Dr Chris Coppin for sharing his knowledge of molecular cloning and library screening.

Thanks to all my colleagues and students for discussion and support: Dr Karine Caron, Dr Sahil Balotra, Dr Andrew Warden, Dr Madhura Shettigar, Fizza Modh Pushri, James Antoney, Dr Brendon Lee, Matthias Nachtschatt, Ellen Walsh, Blair Ney, Dr Sarah Rottet, Dr Mihir Shah, Dr Phil Robbins.

Finally thanks to the “coffee team” for the much needed daily coffee and puzzle break.

I gratefully acknowledge the funding received towards my PhD from the government and from the ANU Research School of Chemistry (RSC): the Australian Postgraduate Award (APA) stipend, the RSC Supplementary scholarship and the Alan Sargeson Merit Scholarship in Chemical Sciences.

List of publications

Esquirol L*, Peat TS*, Wilding M, Lucent D, French NG, Hartley CJ, Newman J, Scott C. (2018) Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841. PLoS One 13(2): e0192736. https://doi.org/10.1371/journal.pone.0192736.

Lygie Esquirol*, Thomas S. Peat, Matthew Wilding, Jian-Wei Liu, Nigel G. French, Carol J. Hartley, Hideki Onagi, Thomas Nebl, Christopher J. Easton, Janet Newman, and Colin Scott (2018). An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme. J. Biol. Chem. 2018 293: 7880-7891. Published 05/2018. doi:10.1074/jbc.RA118.001996 http://www.jbc.org/content/293/20/7880

Esquirol L*, Peat TS*, Wilding M, Hartley CJ, Newman J, Scott C (2018) A novel decarboxylating amidohydrolase involved in avoiding metabolic dead ends during cyanuric acid catabolism in Pseudomonas sp. strain ADP. PLoS ONE 13(11): e0206949. Published 11/2018. https://doi.org/10.1371/journal.pone.0206949.

*Authors contributed equally

Contents

Abbreviation

µM	Micromolar
Å	Ångström
AH	Allophanate amidohydrolase
AtzA	Atrazine chlorohydrolase
AtzB	Hydroxy-atrazine *N*-ethylaminohydrolase
AtzC	*N*-ethylaminohydrolase
AtzD	Cyanuric acid amidohydrolase
AtzE	1-Carboxybiuret hydrolase
AtzF	Allophanate hydrolase
AtzR	Cyanuric acid-responsive regulator regulator
BA	Barbituric acid
BAH	Barbituric acid amidohydrolase
BiuH	Biuret amidohydrolase
CA	Cyanuric acid
ca.	*Circa*
CAH	Cyanuric acid amidohydrolase
DNA	Deoxyribonucleic acid
DSF	Differential scanning fluorometry
DTT	Dithiothreitol
E. coli	*Escherichia coli*
E.U.	European Union
EDTA	Ethylenediaminetetraacetic acid
EPA	Environmental protection agency
FAO	Food and agriculture organisation
GatCAB	Glutamyl-tRNA amidotransferase
GDH	Glutamate dehydrogenase
GuaD	Guanine deaminase
H-bond	Hydrogen bond
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
IPTG	Isopropyl β–D-1-thiogalactopyranoside
IS	Insertion sequence
kb	Kilobase pairs
k_{cat}	Catalytic constant
kDa	KiloDalton
K_M	Michaelis constant
LB	Luria-Bertani
LC-MS	Liquid chromatography mass spectrophotometry
mer	Mercury-resistance gene
MES	2-(*N*-Morpholino)ethanesulfonic acid
NA	Not available
NADH	Nicotinamide adenine dinucleotide hydrogen
ND	Non determined
NDSB	Non-detergent sulfobetaine
NicF	Maleamate amidohydrolase

OriV	Origin of replication
PDB	Protein data bank
PncA	Nicotinamidase
RDX	Royal demolition explosive
RMSD	Root-mean-square deviation
RU	Repeat unit
RutB	Ureidoacrylate peracid amidohydrolase
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SeMet	seleno-*L*-methionine
spont.	Spontaneous
T_m	Melting temperature
tnpA	Transposase gene
tra	Transfer operon
trb	Transfer operon
TriA	Melamine deaminase
TrzA	Melamine deaminase/ atrazine dechlorohydrolase
TrzB	Hydroxy-atrazine N-ethylaminohydrolase
TrzC	*N*-ethylaminohydrolase
TrzD	Cyanuric acid amidohydrolase
TrzF	Allophanate amidohydrolase
TrzN	Atrazine chlorohydrolase

Abstract

Pseudomonas sp. strain ADP1 mineralises the persistent *s*-triazine herbicide atrazine. It has been used as a model organism to study evolution of xenobiotic degradation pathways in bacteria since its discovery in 1995. Prior to this study it was thought that the degradation of atrazine involved six enzymes: AtzA, AtzB, AtzC, which transform atrazine into cyanuric acid, and AtzD, AtzE and AtzF, which mineralise cyanuric acid to ammonia and carbon dioxide. The genes a*tzD*, *atzE* and *atzF* are organised in an operon that is induced by nitrogen starvation and presence of cyanuric acid. The cyanuric acid catabolism operon of *Pseudomonas* sp. strain ADP1 has been intensively studied; however, the *s*tudy of the cyanuric acid degradation pathways in other bacteria has revealed an unexpected diversity in their composition.

The work presented in this thesis has focused on characterising the penultimate step of the atrazine degradation pathway: biuret hydrolysis. This step is performed by quite different enzymes in *Pseudomonas* sp. strain ADP1 and *Rhizobium leguminasorum bv. viciae 3841.* The amidase AtzE was thought to be the biuret hydrolase of *Pseudomonas*, whereas in *Rhizobium* an isochorismatase-like enzyme, BiuH, performs that function. In Chapter 2, we solve the structure of BiuH and investigate its catalytic mechanism by site-directed mutagenesis.

Chapter 3 focuses on the study of AtzE, which had not been characterised previously because it could not be produced in *E. coli*. The structure of native AtzE isolated from *Pseudomonas* sp. strain ADP1 revealed the presence of an unexpected complex between AtzE and a small 68 amino acids protein,

subsequently named AtzG. AtzG is encoded by a previously unreported gene located in the cyanuric acid catabolism operon between *atzD* and *atzE*. The co-expression of AtzE and AtzG in *E. coli* allowed production of stable and active AtzEG. Biochemical characterisation revealed AtzEG does not hydrolyse biuret, as previously thought, but instead catalyses the hydrolysis of the water-sensitive product of AtzD, 1-carboxybiuret. Interestingly, the AtzEG complex was found to be structurally similar to the GatCAB complex, a glutamyl-tRNA amidotransferase, suggesting the AtzEG complex might have evolved from the GatCA complex.

Examination of the cyanuric acid operon also revealed the presence of another small open reading frame located between *atzE* and *atzF*, renamed *atzH*. Proteomics confirmed that AtzH was expressed along with the rest of the cyanuric acid catabolism operon. A crystal structure of AtzH was obtained and biochemical characterisation confirmed it played a role in the degradation of cyanuric acid (Chapter 4). Docking studies and mutagenesis suggested a catalytic role for AtzH: an allophanate-forming, decarboxylating 1,3-dicarboxyurea amidohyrdolase.

Two different cyanuric acid degradation strategies have clearly evolved in *Pseudomomas* and *Rhizobium*. The *Pseudomomas* strategy is to use enzymes to limit the formation of metabolic ‘dead-ends’ that spontaneously form from unstable metabolic intermediates. By contrast, the *Rhizobium* pathway relies on these spontaneous reactions to produce the substrate of BiuH.

1 Introduction

1.1 *s*-Triazine usage has led to the evolution of new metabolic pathways in microorganisms

The anthropogenic symmetrical triazines (*s*-triazines) form a family of polyvalent compounds that are used around the world. The first *s*-triazine compound synthesised was cyanuric acid (1,3,5-triazinane-2,4,6-trione) in the 1800s (1). All of the *s*-triazine compounds share a common core structure of a six membered, aromatic ring containing three nitrogen and three carbon atoms (Fig. 1-1). Commonly used *s*-triazine compounds include: broad-spectrum herbicides (such as atrazine, simazine, propazine, cyanazine, ametryn and terbutylazine), melamine (used in the production of certain plastics and polymers) (2), the military explosive RDX (Royal Demolition eXplosive) (3) and the disinfectant and swimming pool chlorine stabiliser cyanuric acid (4) (Fig. 1-1). Although *s*-triazines are predominantly synthetic, cyanuric acid is also produced naturally. It has been found to be the main component in the mineral Joanneumite, which was discovered recently in Chile (5), and it is also formed as a product of oxidative and ionizing radiation damage of purines (6).

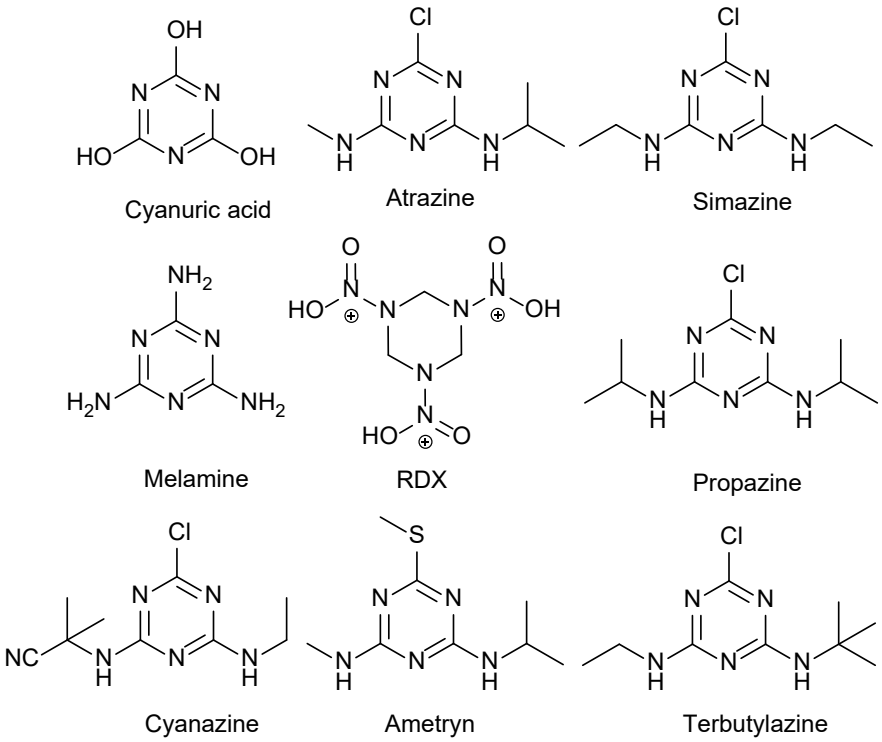


Figure 1-1 s-Triazine compounds cited in this work. *Structures of: cyanuric acid a disinfectant and swimming pool chlorine stabiliser (4), the herbicides atrazine, simazine, of melamine, a compound used in the production of certain plastics and polymers (2), of the military explosive RDX (3) and of the herbicides propazine, cyanazine, ametryn and terbutylazine.*

Among the most abundant group of *s*-triazines deliberately introduced into the environment are the herbicides. The *s*-triazine herbicides are broad-range pre- and post-emergence herbicides, which disrupt electron transport during photosynthesis by binding to the plastoquinone-binding protein in photosystem II and preventing quinone binding (7-9). The *s*-triazine herbicide atrazine, synthetised in the 1950s, has been one of the most used herbicides worldwide (10, 11) behind glyphosate (12). Based on FAO and EPA data, in 2012 around 1.3 billion tonnes of herbicides were used around the world and atrazine accounted for around 80,000 to 90,000 million tonnes (12, 13). The use of

atrazine was found to increase the yield efficiency of maize by 6% and across all type of crops by an average of 3-4% (14).

Despite the undeniable advantages conferred by their use, there are a number of concerns regarding the continued use of *s*-triazine herbicides. These herbicides are quite persistent, with a half-life in the soil ranging from two months to one year (15). They have also been detected in both surface and ground waters beyond regulatory limits, leading to potential human exposure (16-22). Toxicological studies have suggested that the *s*-triazines are possible carcinogens, teratogens and an endocrine disrupters (23-31). These concerns led to the discontinuation of the use of atrazine in the E.U. in 2003 (20), but they are still used widely elsewhere (11). There are also potential environmental effects of atrazine use, including off-target herbicidal activity, with some atrazine-mediated impacts found on lichen, microalgae and benthic diatom communities (8, 32-34).

In addition to the immediate effects of *s*-triazine exposure, the introduction of *s*-triazines to the environment has resulted in the introduction of new selection pressures. This has led to the acquisition of new biochemical functions in soil and aquatic microorganisms, including the evolution of new metabolic pathways and enzymatic activities (15, 35-39).

1.2 Discovery *of s*-triazine-degrading microorganisms and their catabolic pathways

Work towards the discovery of *s*-triazine degrading strains of soil microorganism started in the 1960s, and continues today. In 1950, *s*-triazines degradation products were shown to be present in soil that had been exposed to *s*-triazine herbicides, albeit it was debated at the time whether these products were the result of a microbial catabolic activity (40) or the result of a slow abiotic degradation (41, 42). Between 1960 and 1970, several fungal species that could transform atrazine were identified (Table 1-1). It was later demonstrated that their degradation abilities were mediated by non-specific enzymes, such as cytochromes P450 (15, 43). Between the 1960s and the 1980s, a number of bacterial species were shown to transform or degrade atrazine and other *s*-triazines (Table 1-1), albeit they could only catalyse a small number of transformative steps and these compounds were not completely catabolised by any single bacterial species. In 1984, the first pure culture that could completely degrade an *s*-triazine (melamine), *Pseudomonas* sp. strain NRRL B-12228 (44-46), was isolated by Cook and co-workers. Although Cook *et al.* were able to isolate the enzymes that were responsible for the degradation of melamine ammeline, ammelide and cyanuric acid in 1981, it took a further ten years before the genes encoding the *s*-triazine catabolic enzymes (*triA*, *trzB*, *trzC* and *trzD*), were cloned by Eaton *et al.* (47, 48) (Table 1-2).

In 1994, *Pseudomonas* sp. YAYA6 was isolated by Yanzekontchou and Gschwind (49); it was the first isolated bacterial strain able to entirely degrade atrazine to carbon dioxide and ammonia. The following year, Mandelbaum *et al.* discovered another *Pseudomonas* strain able to entirely degrade atrazine: *Pseudomonas* sp. strain ADP (50), which later became the model organism used to study *s*-triazine degradation. Since then, a large number of *s*-triazine catabolising bacterial strains have been isolated (Table 1-1).

Table 1-1 Isolated s-triazine degrading strains. *ND: not determined.*

Ref.	year	Microorganism	Gene(s) present	Substrate(s)
(40)	1961	Mixed culture *Penicillium* *Aspergillus* *Streptomyces*	ND	simazine
(51)	1963	*Aspergillus fumigatus*	ND	simazine
(52)	1970	*Aspergillus* *Fusarium* *Penicillium* *Trichoderma*	ND	atrazine, simazine, propazine
(53)	1982	*Nocardia*	ND	atrazine
(45)	1984	*Rhodococcus corallinus* *Pseudomonas* sp. strain NRRL B-12228	ND	melamine, simazine, atrazine
(46)	1985	*Pseudomonas* strain D (= *Pseudomonas* sp. strain NRRL B-12228) *Klebsiella pneumoniae*	ND	cyanuric acid
(47)	1991	*Pseudomonas* sp. strain NRRL B-12227	*triA, atzD*	melamine, cyanuric acid
(54)	1986	2 strains *Pseudomonas*	ND	atrazine
(55)	1993	*Acinetobacter calcoaceticus* *Pseudomonas alcaligenes/Agrobacterium sp.* *Pseudomonas putida/ xanthomonas maltophilia*	ND	atrazine

(56)	1994	*Rhodococcus strain B-30*	ND	atrazine, propazine, simazine
(49)	1994	*Pseudomonas* sp. YAYA6	ND	atrazine
(50)	1995	*Pseudomonas* sp. strain ADP	ND	atrazine
(43)	1995	*Rhodococcus* NI86/21	P-450	atrazine
(57)	1997	*Rhizobium* sp. PATR	*atzA*	atrazine
(58)	1997	*Klebsiella terragena*	ND	melamine
(59)	1999	*Pseudomonas* sp. strain NRRLB-12227	*trzD*	cyanuric acid
(60)	2000	*14 strains, Pseudoaminobacter* C147	*atzA, B, C*	atrazine
(61)	2000	*Nocardia Nocardioides* sp. strain C190	*trzN*	atrazine
(62)	2001	*Comamonas acidovorans*	ND	atrazine
(63)	2001	*Penicillium steckii* DS6F *Moraxella ovis* N5C	ND	simazine
(64)	2001	*Chelobacter heintzii* *Aminobacter aminovorans* *Stenotrophomonas maltophilia* *Arthrobacter crystallopoietes*	*atzA, B, C, trzD* " " *atzB, C*	atrazine
(65)	2002	*Arthrobacter aurescens* TC1	*trzN, atzB, C*	atrazine, broad range
(66)	2003	*Nocardioides sp. SP12*	*trzN, atzB, C*	atrazine
(67)	2003	*Arthrobacter* sp. AD1	*AtzA*	atrazine
(68)	2004	*Acinetobacter* A6	ND	atrazine, broad range
(69)	2004	*Agrobacterium radiobacter* *Bradyrhizobium japonicum*	ND	simazine
(70)	2005	*Arthrobacter nicotinovorans HIM*	*atzA*, *B*, *C*	atrazine, simazine, propazine, broad range
(71)	2006	*Nocardioides*	*trzN*	atrazine
(72)	2006	*Micrococcus* sp. strain MF-1	ND	melamine
(73)	2007	*Arthrobacter* sp. strain MCM B-436	*trzN, atzB, C, D*	atrazine
(74)	2007	*6 strains* *Nocardioides* *Arthrobacter*	*trzN, atzC*	atrazine
(75)	2007	*Nocardioides sp. NEA-A*	*atzA, B, C, D, E, F*	atrazine
		Agrobacterium sp. NEA-D	*atzA, B, C, D, E, F*	
		Shinorhizobium sp. NEA-B *Polaromonas sp. NEA-C*	*trzN, B, C*	

		Nocardioides sp. 1D	*trzN, B, C*	
		Arthrobacter sp. 2B	*trzN*	
		Arthrobacter sp. 3A	*trzN*	
			trzN	
(76)	2007	*Arthrobacter ATZ1*	trzN, atzC	atrazine
		Arthrobacter ATZ2	trzN, atzB,	atrazine
		Ochrobactrum CA1	C	cyanuric acid
		Pseudomonas CA2	*trzD*	cyanuric acid
			trzD	
(77)	2007	*Variovorax* sp. MD1/MD2	*atzA*, *B*	atrazine
(78)	2007	*Methyloversalis* CDB21	*atzA, B, C, D, E, F*	simazine
(73)	2007	*Arthrobacter* sp strain MCM B-436	*trzN, atzB, C, D*	atrazine
(79)	2008	*Pseudomonas* sp. MHP41	*atzA, B, C, D, E, F*	simazine
(80)	2008	*Stenotrophomonas* P51	*atzA, D*	simazine
		Stenotrophomonas C53	*atzA, D*	
		Arthrobacter P52	*atzD*	
(81)	2008	*Arthorbacter* sp. AD26	*trzN, atzB, C*	atrazine
(82)	2009	*Alcaligenes faecalis*	ND	atrazine
		Klebsiella ornithinolytica	*atzA*	
		Bacillus megaterium	ND	
		Agrobacterium tumefaciens	*atzA*	
(83)	2009	*Arthrobacter* sp. GZK-1	ND	atrazine terbuthylazine
(84)	2010	54 strains *Arthrobacter*	*trzN*	NA
		28 strains *Nocardioides*	*trzN*	
		Ancylobacter T10AII	*atzA*, *B*, *C*, *D*, *E*, *F*	
(85)	2010	*Nocardioides* sp. strain DN36	*trzN*, *atzB*, *C*	atrazine simazine propazine
(86)	2010	*Arthrobacter* sp. strain KU001	*trzN*, *atzB*, *C*	
(87)	2010	*Klebsiella* sp. A1	ND	atrazine
		Comamonas sp. A2	*atzA, B, C, D, E, F*	atrazine
(88)	2011	*Arthorbacter* sp. strain DNS10	*trzN*, *atzB*, *C*	atrazine
(89)	2011	*Arthrobacter* sp. TES6	*trzN*, *atzB*, *C*	atrazine
(90)	2012	*Rhodococcus* sp. JN201860	*trzA*	melamine
(91)	2012	*Nocardioides* sp. strain ATD6	*triA*	melamine
(92)	2012	*Rhizobium* sp. C14	*AtzA, B, C*	atrazine
		Acinetobacter lwoffii C1		
(93)	2012	*Enterobacter cloacae*	ND	atrazine
		Burkholderia capacia sp.		

(94)	2012	*Arthrobacter* sp. strain DAT1	*trzN, atzB, C*	atrazine
(95)	2013	*Arthrobacter*	*trzN, atzB, C*	atrazine
(96)	2013	*Nocardioides* sp. EAA-3 *Nocardioides* sp. EAA-44	*trzN, atzB, C*	atrazine
(97)	2014	*Alcaligenes* sp. strain EGD-AK7	*atzA, B, C, D, E, F*	
		Arthrobacter sp. strain AK-YN10	*trzN, atzB, C*	
(98)	2014	*Bacilus subtilis* HB-6	*trzN,B,C*	atrazine
(95)	2014	*Pseudomonas aeruginosa*	*atzA, B, C, D, E, F*	atrazine
(99)	2014	*Arthrobacter* sp. SD3-25	*trzN, atzB, C*	simazine
(100)	2015	*Sterptomyces* sp. Atz2	ND	atrazine
(101)	2015	*Arthrobacter* sp. MCO *Arthrobacter* sp. CSP *Microbacterium* sp. ZEL	Not found	melamine, cyanuric acid
(102)	2016	*Variovorax* sp. *Schlesneria* *Arthorbacter* spp	*trzN, atzA, B, C, trzD*	atrazine
(103)	2016	*Shewanella* sp. YJY4	*atzA, B,C*	atrazine
(104)	2017	*Ensifer* sp. strain CX-T	*atzA, B, C, D, E, F*	atrazine
(105)	2017	*Pseudomonas* sp. ZXY-1	ND	atrazine
(106)	2018	*Citricoccus* sp. strain TT3	*trzN, atzB, C*	atrazine

(ND: not determined, CA: cyanuric acid)

The atrazine degradation pathway is composed of a seven steps process and can be separated in two parts: the 'upper degradation pathway', which transforms particular *s*-triazines to cyanuric acid, and the 'lower degradation pathway' which mineralises cyanuric acid to ammonia and carbon dioxide (Fig. 1-2).

In *Pseudomonas* sp. strain ADP, six enzymes have been identified that are required for *s*-triazine catabolism (Section 1.4). Wackett et *al.* cloned the genes coding for the enzymes of the first three steps of the degradation pathway, required to transform atrazine to cyanuric acid: *atzA* (107), *atzB* (108) and *atzC* (109) (Fig. 1-2, Table 1-2, 3). The *atz* genes were shown to be localised on an IncPβ plasmid, pADP1. In 2001, pADP1 was sequenced by Martinez *et al.* (110). In addition to *atzA*, *atzB* and *atzC*, Martinez *et al.* also identified three new genes *atzD*, *atzE* and *atzF*, which encodes enzymes required to degrade cyanuric acid to ammonia and carbon dioxide (Fig. 1-2, Table 1-2, 1-3). Interestingly, de Souza *et al.* (111) identified *atzA*, *B*, *C* across five different strains of bacteria (*Pseudomonas*, *Alcaligenes*, *Clavibacter Ralstonia* and *Agrobacterium*) and concluded these genes were both transmissible and widespread.

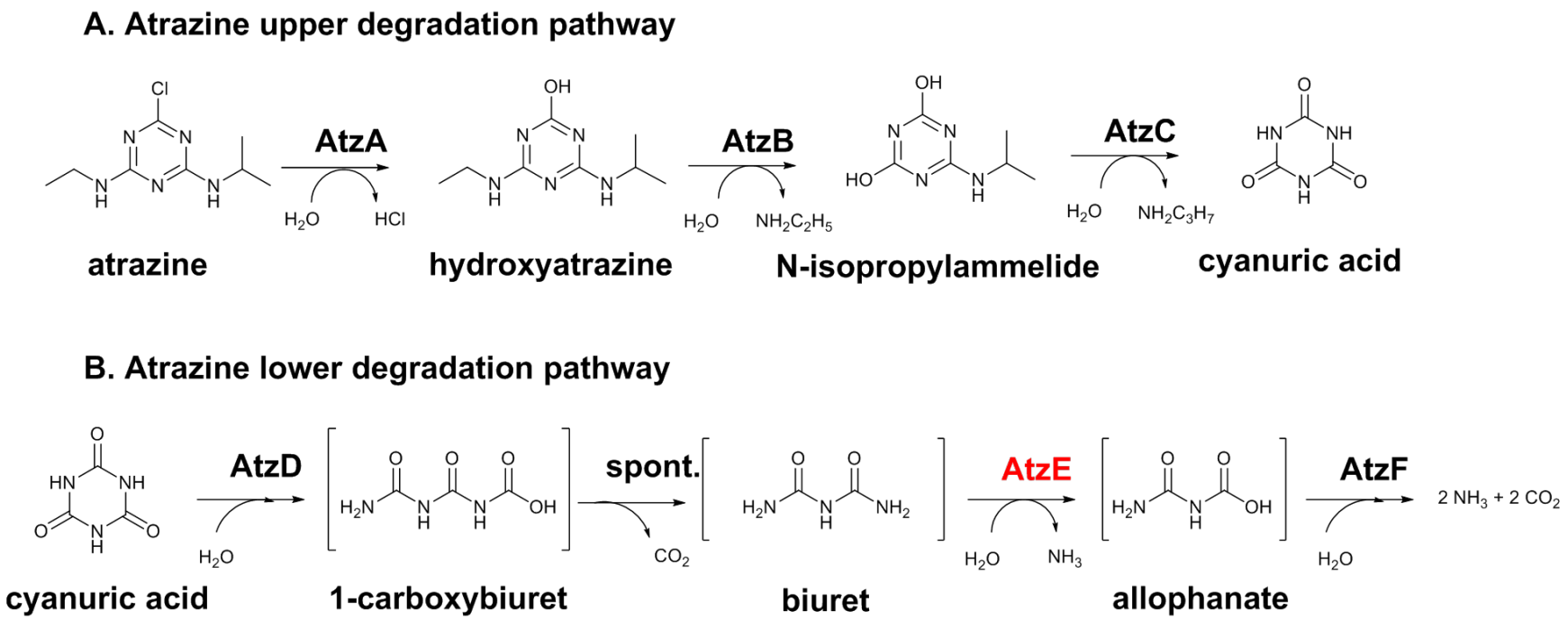


Figure 1-2 Atrazine upper and lower degradation pathways from* Pseudomonas *sp. strain ADP. A. *The 'upper pathway' composed of: the dechlorination of atrazine (or other choro-s-triazines), followed by the deamination of the ethylamine side-chain of 2-hydroxy-atrazine and the deamination of* N*-isopropylammelide leading to the production of cyanuric acid (107-109, 112);* ***B.*** *The 'lower pathway' involves the hydrolytic ring-opening of cyanuric acid ring leading to the production of 1-carboxybiuret, which is thought to spontaneously decarboxylate to biuret in physiological conditions (113, 114). Biuret is then deaminated to allophanate, which is then degraded to ammonia and carbon dioxide (115, 116). The enzymes responsible for each step are shown and AtzE is shown in red as it has not been biochemically or structurally characterised yet. 'Spont.' indicates a solvent-mediated hydrolysis.*

Although the *Pseudomonas* sp. strain ADP pathway is the best characterised, several alternative enzymes have been adopted for *s*-triazine catabolism in other bacteria. In 2000, *Nocardioides* sp. strain C190 was demonstrated to perform the dechlorination of atrazine, the first step in the degradation pathway, without *atzA*. Topp *et al.* cloned, expressed and characterised the enzyme responsible for this activity: TrzN (61). Subsequently, *trzN* was found to be present in other *Nocardia* (66), in *Arthrobacter* species (65, 73, 74, 81, 88, 94, 117) and in *Bacillus* species (98), revealing a previously unknown diversity in the *s*-triazine degradation pathways of bacteria. Cameron *et al.* identified and characterised a *biuH* gene from *Rhizobium leguminosarum* bv. *viciae* 3841, encoding for a cysteine hydrolase enzyme from the isochorismatase family, that was able to degrade biuret (118), showing that the lower degradation pathway also possesses some variability (Section 1.3 and 1.4).

Table 1-2 Genes (A) and genomes (B) involved in s-triazine herbicide catabolism.

A.

Ref.	year	microorganism	Genes
(47, 48)	1991	*Pseudomonas* sp. strain NRRL B-12228/27/99 *Klebsiella pneumoniae*	*triA, trzB, trzC*, *trzD*
(119)	1995	*Pseudomonas* sp. strain ADP	*atzA*
(120)	1995	*Rhodococcus corallinus*	*trzA*
(108)	1997	*Pseudomonas* sp. strain ADP	*atzB*
(109)	1998	*Pseudomonas* sp. strain ADP	*atzC*
(111, 121)	1998	*Pseudomonas* sp. strain CP1 *Clavibacter michiganense* ATZ1	*atzA*, *atzB, atzC*
(61, 122)	2000/2	*Nocardioides* sp. strain C190	*trzN*
(123-128)	2003/5/7/9/10	*Pseudomonas* sp. strain ADP	*atzR, atzD, atzE, atzF*
(129)	2004	*Arthrobacter aurescens* TC1	*trzN, atzB, atzC*

B.

Ref.	year	Microorganism
(110)	2001	*Pseudomonas* sp. strain ADP p-ADP1 plasmid
(130)	2006	*Arthrobacter aurescens* TC1
(131)	2006	*Rhizobium leguminosarum* bv. *viciae* 3841
(132)	2013	*Alcaligenes* sp. Strain HPC1271
(133)	2013	*Arthrobacter* sp.FB24
(97)	2014	*Alcaligenes* sp. strain EGD-AK7 *Arthrobacter* sp. strain AK-YN10
(134)	2016	*Pseudomonas* sp. strain ADP

Table 1-3 Biochemical (A) and structural (B) characterisation of the s-triazine enzymes.

A. Biochemical characterisation

Ref	Year	microorganism	Enzyme
(135)	1982	*Pseudomonas* strain A D F, 2 strains *Klebsiella pneumoniae*	Not named
(43)	1995	*Rhodococcus* sp. strain Ni86/21	Cytochrome P450
(59)	1999	*Pseudomonas* sp. strain NRRLB-12227	trzD
(107, 136)	1996/ 2002	*Pseudomonas* sp. strain ADP *Rhodococcus corallinus*	AtzA TrzA
(137)	2002	*Pseudomonas* sp. strain ADP	AtzC
(115, 116)	2005	*Pseudomonas* sp. strain ADP	AtzF
(138)	2006	*Enterobacter cloacae* strain 99	TrzF
(139)	2006	*Arthrobacter aurescens* TC1	TrzN
(112)	2007	*Pseudomonas* sp. strain ADP	AtzB
(118)	2011	*Rhizobium leguminosarum* bv. *viciae* strain 3841	BiuH

B. Structural characterisation

Ref	year	microorganism	Enzyme
(140)	2010	*Arthrobacter aurescens* TC1	TrzN
(141, 142)	2014/15	*Pseudomonas* sp. strain ADP	AtzF
(143, 144)	2013/14	*Azorhizobium caulinodans* ORS571	AtzD
(145)	2013	*Pseudomonas* sp. strain ADP	AtzD
(146)	2015	*Pseudomonas* sp. strain ADP	AtzC
(147)	2015	*Pseudomonas* sp. strain ADP	AtzA
(148)	2017	*Acidovorax citrulli*	TrzD
(149)	2017	*Frankia*	AtzD

1.3 *s*-Triazine catabolic pathway(s)

1.3.1 The *s*-triazine catabolic pathway of *Pseudomonas* sp. strain ADP

In *Pseudomonas* sp. strain ADP the hydrolytic enzymes required for the transformation of atrazine to cyanuric acid (i.e., the 'upper pathway') are: the atrazine chlorohydrolase AtzA enzyme (E.C. 3.8.1.8) (107, 108), the hydroxyatrazine amidohydrolase AtzB enzyme (EC 3.5.99.3) (112), and the isopropylammelide amidohydrolase AtzC enzyme (EC 3.5.99.4)(107-109, 112, 137) (Fig. 1-2, 1-4). The steps of the 'lower pathway' are performed by: the cyanuric acid amidohydrolase AtzD enzyme (E.C. 3.5.2.15) (113, 114), the biuret amidohydrolase AtzE enzyme (EC 3.5.1.54) and the allophanate amidohydrolase AtzF enzyme (3.5.1.84) (116). The last three steps are the hydrolysis of the cyanuric acid ring leading to the production of 1-carboxybiuret, thought to spontaneously decarboxylate to biuret in the presence of water (113, 114). Biuret is deaminated to produce allophanate, which is hydrolysed to form ammonia and carbon dioxide (115, 116). The genes encoding all of the atrazine catabolic enzymes of *Pseudomonas* sp. strain ADP are located on the pADP1 catabolic plasmid (110) (Fig. 1-3).

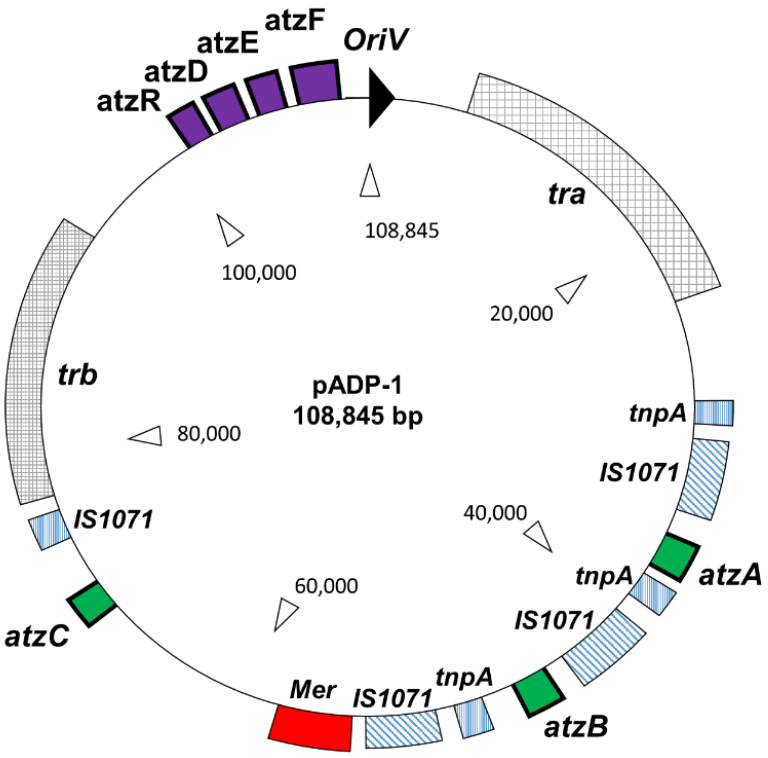


***Figure 1-3 Map of the catabolic plasmid pADP-1 from* Pseudomonas *sp. strain ADP (reproduced based on Martinez* et al. (110)).** *The genes required for atrazine catabolism are represented in green for the 'upper pathway' and in purple for the 'lower pathway'; OriV marks the origin of replication; tra and trb are transfer operons allowing plasmid conjugation; tnpA represents transposases and IS1071 represents insertion sequences allowing transposition of specific portion of DNA; mer is a mercury resistance operon.*

The pADP1 sequence revealed the presence of the *tra* and *trb* operons, which encode proteins essential for plasmid conjugation (110). In 1998, pADP has been shown to be transmissible under laboratory conditions (111), and (as noted above) the *atzA*, *atzB* and *atzC* genes have been found in a variety of bacteria and geographic locations (75, 77, 84). The *atzA*, *atzB*, and *atzC* genes were found to be monocistronic, constitutively expressed, not co-located (Fig. 1-3), and flanked by transposases and insertion sequences (110). This suggested that these genes are mobile and explained how they can be found independently from one another in different bacterial strains (15, 110, 150)(Table 1-1).

The *atzD*, atz*E* and atz*F* genes were shown to be organised as an operon (110). The regulation of this operon has been studied in detail by Govantes *et al. (123-128, 151)*. The operon was found to be regulated by a LysR-type transcription regulator named AtzR (124, 125). The induction of the operon was found to be upregulated under nitrogen-limiting conditions. Nitrogen starvation in *Pseudomonas* sp. strain ADP is sensed *via* the general nitrogen control circuit (GlnK-NtrC) and regulates transcription of the cyanuric mineralization operon *via* upregulation of *atzR* expression from a σ^{54}-dependent promoter under nitrogen limiting conditions (123, 126, 128). AtzR then upregulates the expression of the the cyanuric acid mineralisation operon, even in the absence of cyanuric acid (125, 128); however, the presence of cyanuric acid further increases the level of transcription of the operon by inducing a conformational change in the regulator, that induces DNA bending and enhances the rate of transcription (124).

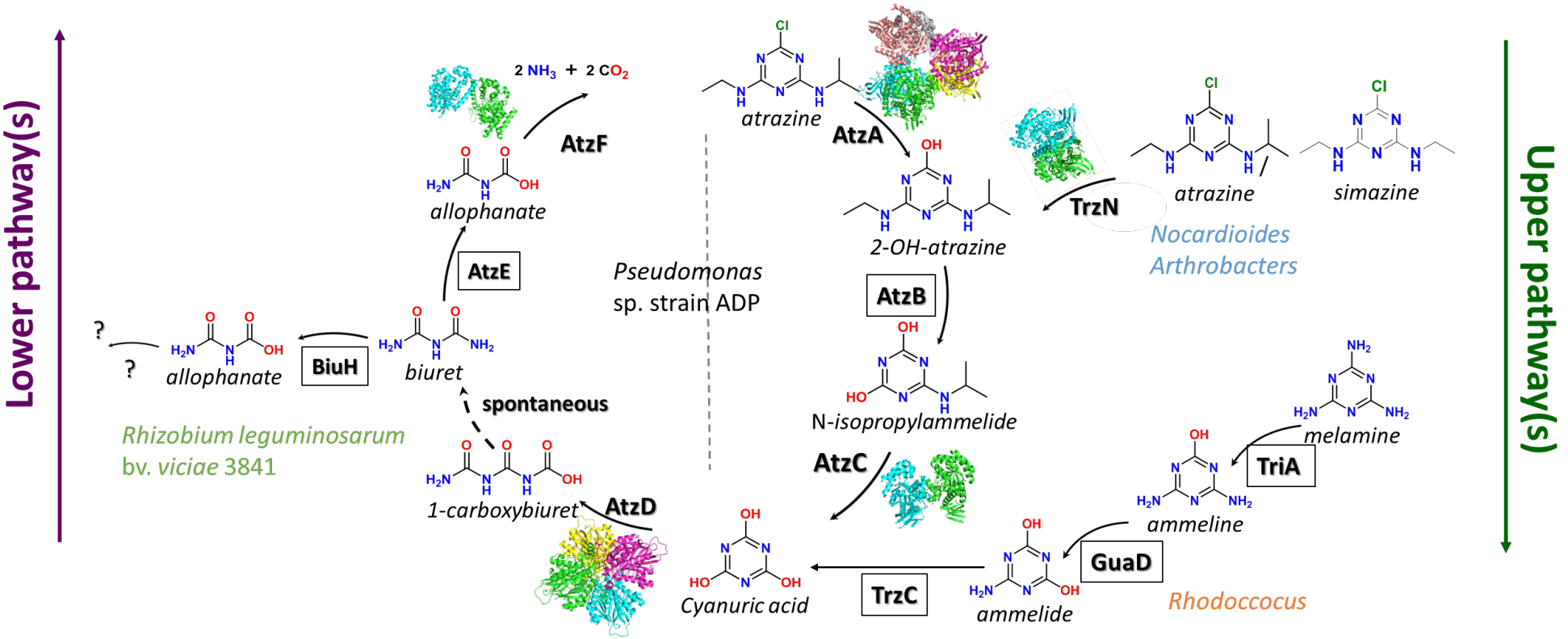


Figure 1-4* s-*Triazine herbicide catabolic pathways. *The* Pseudomonas *sp. strain ADP catabolic pathway is shown by the central pathway. Pathways that have been identified from other bacteria, such as* Nocardioides, Arthrobacters, Rhodoccocus *or* Rhizobium, *are shown around. The 'upper pathway' comprises of the degradation of atrazine, simazine or melamine leading to the production of cyanuric acid. The 'lower pathway' comprises of the degradation of cyanuric acid to ammonia and carbon dioxide. Where available, protein structures are shown in cartoon representation next to their names. Where no structure is available, the enzyme names is shown inside a black box.*

1.3.2 Variations to the *s*-triazine catabolic pathway identified in other bacteria

Although less extensively studied than the *Pseudomonas* sp. strain ADP pathway, several variations on the model *s*-triazine catabolism pathway have been described in other bacterial strains (15). As shown in the Table 1-1, *s*-triazine degrading bacteria can contain a variety of triazine catabolism genes: *atzA* or *trzN* only, complete '*Pseudomonas*-type' pathways (*atzA*, *atzB*, *atzC*, *atzD*, *atzE*, *atzF*), the '*Pseudomonas* upper degradation-type' pathway (*atzA*, *atzB*, *atzC*), or a combination of *trzN*, *atzB* and *atzC* with or without *atzD* or its homologue *trzD*.

The *trzN* gene encodes a physiological analogue of the enzyme encoded by *atzA* and is frequently found in Gram-positive bacteria such as *Nocardioides* and *Arthrobacter* (61, 122). TrzN dehalogenates atrazine simazine, propazine (152), and some naturally occurring TrzN variants can also hydrolyse less reactive leaving groups, such as the thiomethyl substituent of ametryn (140, 152). The structural and mechanistic differences between TrzN and AtzA are explored in more detail in Section 1.4.

The lower pathway also has some plasticity. *TrzD* homologues have been found to encode a cyanuric acid hydrolases. Depending on the *trzD* homologue, sequence identity between TrzD and AtzD has been found to vary between 40-60 %, but *trzD* has never been found to be part of a cyanuric acid degradation operon (59, 114). This suggests that the polycistronic organisation of cyanuric acid catabolism genes in *Pseudomonas* sp. strain ADP is not a ubiquitous configuration for these genes (118, 153) (Fig. 1-4). A*tzD* homologues have also

been found in association with genes from pathways other than the *Pseudomonas*'; for example, in *Rhizobium leguminosarum* bv. *viciae* 3841, an *atzD* homologue is found immediately upstream of a gene encoding a biuret hydrolase (*biuH*). This pathway differs from the *Pseudomonas*' because BiuH belongs to the isochorismatase family, unlike AtzE that belongs to the amidase family. Additionally, there is no equivalent of *atzF* found downstream of *biuH*.

Interestingly, the melamine catabolic pathway parallels the atrazine catabolic pathway with three hydrolytic steps that sequentially deaminate melamine to ammeline, ammeline to ammelide and ammelide to cyanuric acid (Fig. 1-4). *TriA*, a melamine deaminase, was first identified in 1991 from a *Pseudomonas* sp. strain NRRL B-12227 (47). TriA shares 98% sequence identity with AtzA, but was found to be only a melamine deaminase, unable to degrade atrazine. A second melamine deaminase *TrzA,* 50% identical to AtzA, was identified from a *Rhodococcus coralinus* NRRL-B-15444R strain in 1994. TrzA can perform both the dechlorination of atrazine and the deamination of melamine (154).

s-Triazine catabolic genes are usually plasmid-borne (15), but some studies also report atrazine catabolic genes present in bacterial chromosomes. Cai *et al.* found the presence of *atzA* in the bacterial chromosome of *Arthrobacter* sp. AD1 strain (67). Devers *et al.* also observed variation in gene compositions and locations when they sampled 17 atrazine degrading bacteria from soil (75). The most frequent pattern of genes they observed were: *atzABC-trzD*, *atzABCDEF* and *trzN-atzBC*. In some cases genes were found on one plasmid, in others they were spread across multiple plasmids and they were sometimes found on the chromosome (75). Plasmids were also found to vary in size. Finally, the gene

copy number also varies. For example, there are multiple copies of *trzN* in the *Arthrobacter aurescens* strain, which contains six identical tandem repeats of *trzN* on a 16 kb region of the pTC1 plasmid (130). As yet, the effect of *trzN* gene dosage in this strain has not been investigated.

1.4 Enzymatic basis of the *s*-triazine catabolism

1.4.1 Structure-function of the metalloproteins amidohydrolases of the upper pathway

AtzA (E.C. 3.8.1.8), TrzN (E.C. 3.8.1.8), AtzB (EC 3.5.99.3) and AtzC (EC 3.5.99.4) are all metal-dependant amidohydrolase family enzymes that share a common fold known as the $(\alpha/\beta)_8$ barrel (140, 146, 147, 155) (Fig. 1-5). The substrate specificity of these enzymes depends on the steric, conformational and electronic restrictions imposed by the eight loops that form the active site, located at the end of the eight β-strands (155). Although they catalyse a broad range of reactions, the amidohydrolase mechanism usually involves the activation of a nucleophilic water molecule through the complexation with the metal centre, followed by nucleophilic attack on the substrate by the activated water. In some cases, a proton transfer from an active site amino acid to the substrate is also required to activate the substrate (155). The characteristics of each of these four enzymes are described in the Table 1-4.

Table 1-4 Characteristics of the amidohydrolases of the upper pathway.

E.C. No.	Enzyme	Physiological Function	Kinetic Values[1]	Approx. Size	PFam Family	Accession[2]	Mechanism	Refs
3.8.1.8	TrzN	*s*-Triazine hydrolase	K_M = 20 k_{cat} = 2 k_{cat}/K_M = 1 x 10^5	Dimer (2 x 54,000 Da)	Amidohydrolase	**UniProt:** A1RCJ9, Q8VS01 **PDB:** 4l9x, 4lh8, 5hmd, 5hme, 5hmf	Zn^{2+}-dependent hydrolase	(122, 139, 140, 152, 156, 157)
3.8.1.8	AtzA	*s*-Triazine chlorohydrolase	K_M = ND k_{cat} = ND K_M/k_{cat} = 1.5 x 10^4	Hexamer (6 x 52,000 Da)	Amidohydrolase	**UniProt:** P72156 **PDB:** 4v1x, 4v1y	Fe^{2+}-dependent hydrolase	(119, 136, 147, 158, 159)
3.5.99.3	AtzB	Hydroxydechloro-atrazine ethylaminohydrolase	K_M = 20 k_{cat} = 3 k_{cat}/K_M = 1.5 x 10^5	Dimer (2 x 50,000 Da)	Amidohydrolase	**UniProt:** P95442 **PDB:** NA	Fe^{2+}/Zn^{2+}-dependent hydrolase	(108, 112, 160)
3.5.4.42	AtzC	*N*-Isopropylammelide isopropylamino-hydrolase	K_M = 400 k_{cat} = 13 k_{cat}/K_M = 3.3 x 10^4	Tetramer (4 x 45,000 Da)	Amidohydrolase	**UniProt:** O52063 **PDB:** 2qt3, 4cqb, 4cqc, 4cqd, 5akq	Zn^{2+}-dependent hydrolase	(109, 146)

[1]Values are given for the physiological substrate only. Units for kinetic values are as follows: K_M, μM; k_{cat}, s^{-1}; k_{cat}/K_M, $s^{-1}.M^{-1}$

[2]Accession numbers are given for the 'archetypal' enzymes only

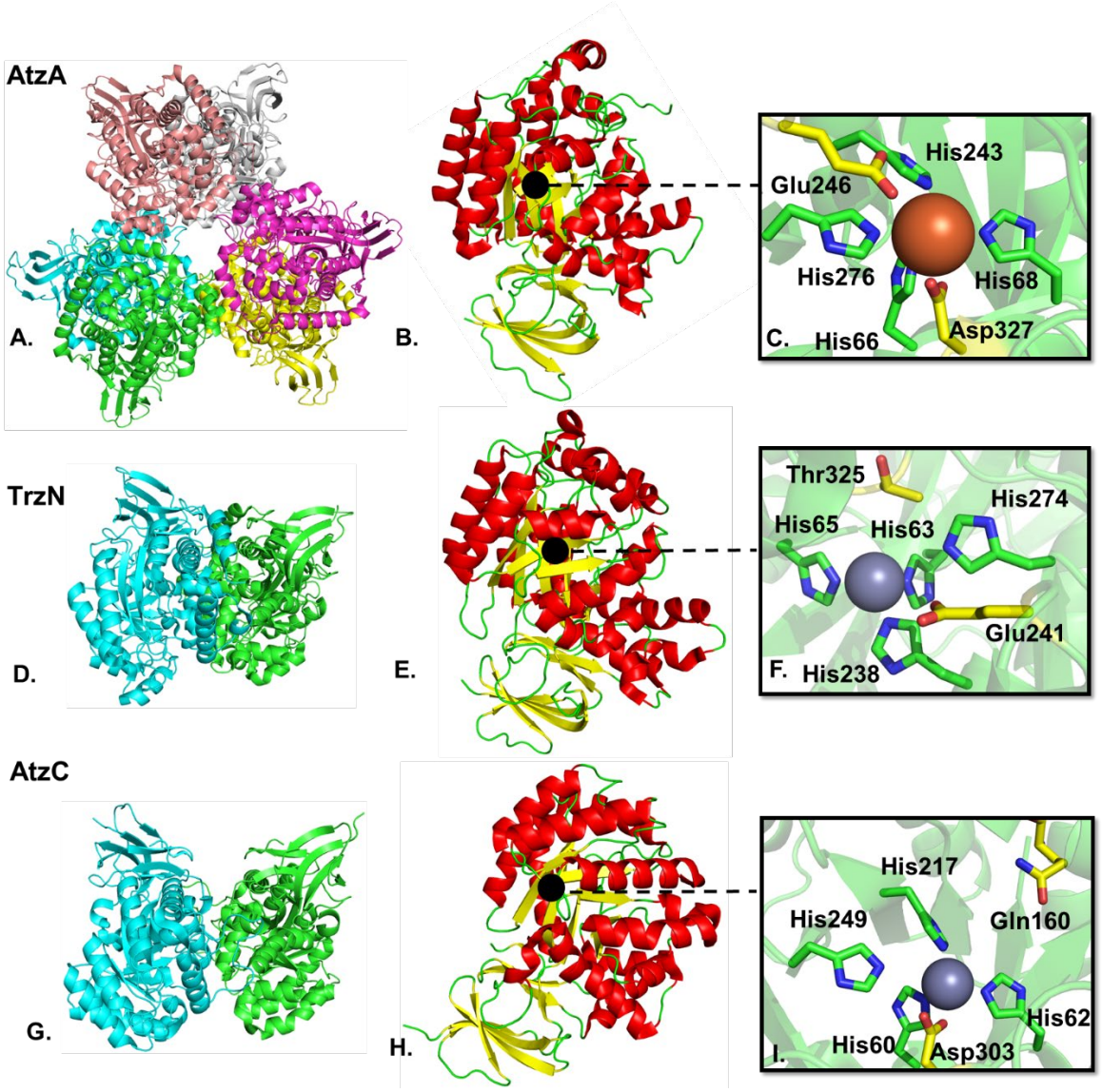


Figure 1-5 Structures of the s-triazine degrading enzymes of the upper pathway. A. *Cartoon representation of the AtzA hexamer (formed by a trimer of dimers (147)); **B.** AtzA monomer; **C.** Zoom onto the AtzA active site with bound Fe^{2+}; **D.** TrzN dimer; **E.** TrzN monomer; **F.** Zoom on the TrzN active site with bound Zn^{2+}; **G.** AtzC dimer; **H.** AtzC monomer; **I.** Zoom on the AtzC active site with bound Zn^{2+}. For the quaternary structures, each monomer is shown in a different colour. For the tertiary structures of the monomers, loops are in green, α-helices are in red and β-strands are in yellow. The black dot indicates the location of the active site, metal cations are represented as spheres (grey for $Zn^{2+}$and orange for Fe^{2+}). The amino acids involved in metal coordination are represented in green sticks for histidine, and the conserved Asp/Thr and Glu residues are represented in yellow sticks. PDB code used are: 4v1y for AtzA, 4l9x for TrzN and 5akq for AtzC.*

AtzA is a hexameric, Fe^{2+}-dependent metalloenzyme (107, 119). The k_{cat}/K_M of this enzyme has been reported at 1.5 x 10^4 $sec^{-1}.M^{-1}$ (159). The K_M of AtzA exceeds the water solubility of atrazine (~153 µM) and as a consequence the K_M and k_{cat} of AtzA cannot be directly measured (Table 1-4). The active site of AtzA is buried at the end of a long hydrophobic channel. A Fe^{2+}, coordinated by four histidine residues (His66, His68, His243 and His276) and an aspartate (Asp327), form the metal centre (147). No crystal structure with atrazine in the active site has been obtained, nor could the substrate be docked into the active site, suggesting a poor complementarity between the active site and its substrate. Peat *et al.* concluded that the high K_M of AtzA for atrazine could be explained by the lack of complementarity. AtzA also has a low affinity for iron, and this was shown to be due to unusually long bond lengths between the metal and the histidine ligands. Overall the unusual bonding pocket appears “ill-adapted” for its function (147, 159). (Fig. 1-5)

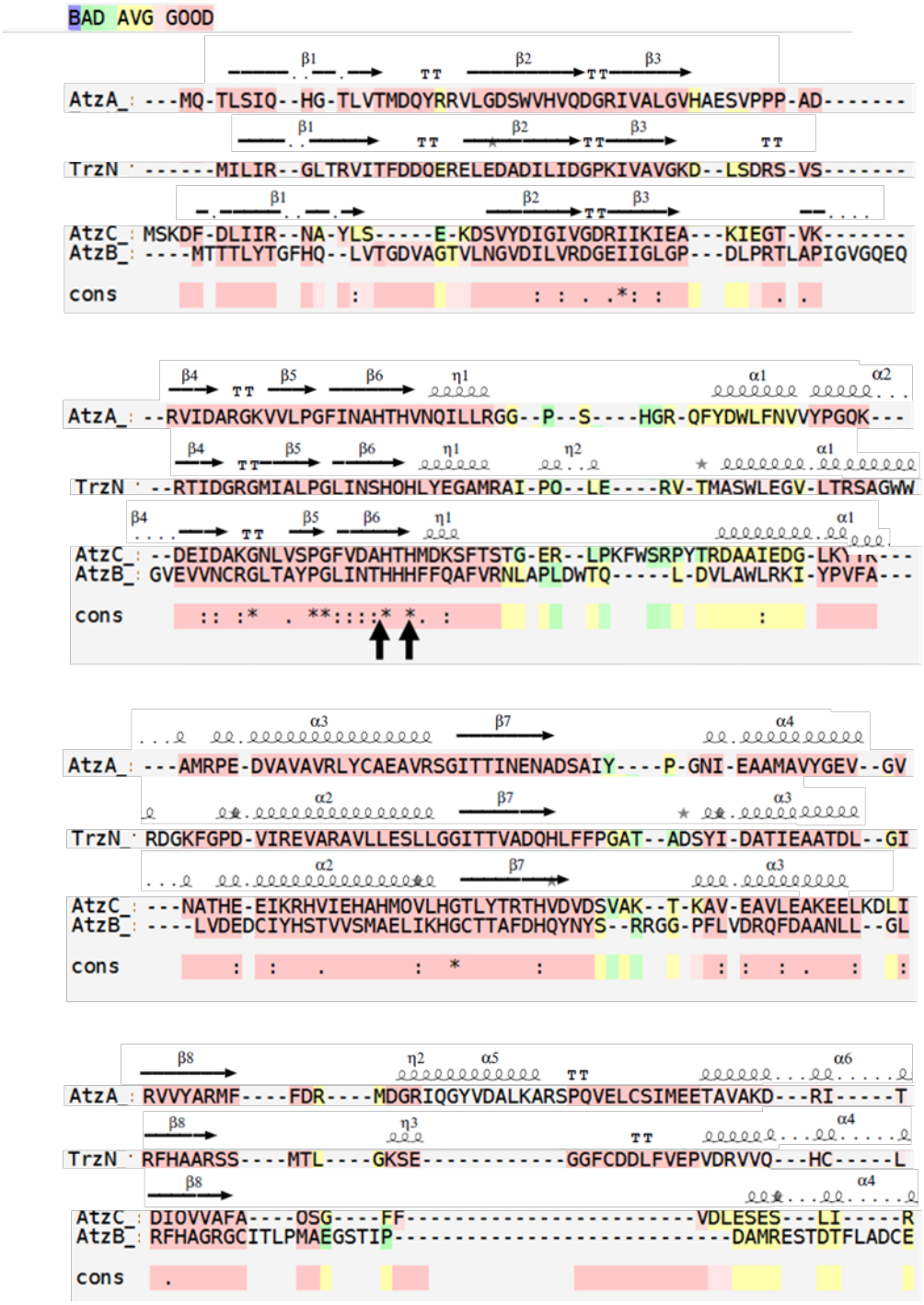
TrzN is a dimeric, zinc-dependent amidohydrolase performing the same physiological function as AtzA. Its metal centre is composed of a Zn^{2+} coordinated by four histidines (His63, His65, His238 and His274) and, unusually, a threonine (Thr325) residue forms a bond to the metal *via* a bridging water molecule. (140, 152).

Some naturally occurring TrzN variants are strict chlorohydrolases (e.g. the *Nocardioides* sp. strain AN3 TrzN) (157, 161, 162), while others (such as the *Nocardioides* sp. strain MTD22 TrzN) are capable of hydrolyzing non-halogen substituents, including the *S*-methyl and *O*-methyl groups from ametryn and

ametron, respectively. Interestingly, this biochemical phenotype is controlled by a small number of amino acid substitutions in the active site: Thr/Pro214, His/Tyr215 and Gln/Glu241 (157). The strict dehalogenase TrzN contains the Gln241 mutation, whereas the broader substrate range TrzN enzyme contains a Glu241. Glu241 residue is thought to protonate the atrazine ring, which then protonates the poorly reactive leaving groups (e.g. $–SCH_3$). Sugrue *et al*. showed that the presence of the two hydrophobic residues Pro214 and Tyr215 contributed to the creation of a microenvironment, increasing the pK_a of the glutamine side-chain and making Glu241 a more reactive catalytic residue, allowing the efficient hydrolysis of ametryn (157). Laboratory evolved AtzA variants that can protonate unreactive leaving groups (such as amines and –SCH3 groups) have been produced, but such variants have reduced dehalogenase activities (158, 163). The active site of TrzN is a better fit for its substrate than is the case for AtzA and atrazine (140, 147), explaining why the K_M of TrzN is lower for atrazine (20 µM) than that of AtzA and the k_{cat}/K_M higher (*ca*. 1 x 10^5 $s^{-1}.M^{-1}$)(159). (Fig. 1-5).

The two next enzymes in the pathway are AtzB and AtzC. AtzB, a metal-dependent amidohydrolase, is a hydroxyatrazine *N*-ethylaminohydrolase. AtzC, a zinc-dependent tetramer, catalyzes the hydrolysis of the *N*-isoproylamine group from *N*-isopropylammelide (k_{cat} = 13 sec^{-1}, K_M = 400 µM) (137). The Zn^{2+} metal centre of AtzC is coordinated by four histidines (His60, His62, His217 and His249) and an Aspartate, Asp303. In AtzC, while the water molecule is activated by the metal ion, the substrate is protonated by Gln160, leading to the protonation of the leaving group (146) (Fig 1-5). AtzC has a relaxed substrate specificity, removing sidechains from a range of substituted ammelides (137). As yet, there is no X-ray structure for AtzB, but it also belongs to the amidohydrolase family, forms a

homodimer and has been demonstrated to contain a metal ion, that could be either Fe^{2+} and/or Zn^{2+} (112). Like AtzC, AtzB has been shown to hydrolyse the side chains from a range of di-*N*-alkyl-2-hydroxy-*s*-triazines (112). Interestingly, AtzB has also been shown to possess chlorohydrolase activity, albeit ten-fold lower than its physiological deaminase activity with k_{cat}/K_M values of 1.5×10^5 and 1.6×10^4 $M^{-1}.s^{-1}$ for deamination and dechlorination, respectively (112). Sequence alignment of AtzB with the other metallo-enzymes TrzN, AtzA and AtzC shows conservation of the histidine binding metals and of the asparagine/threonine and suggests His75, His77, His249, His284 and Asp336 might be the amino acids involved in metal binding in AtzB (Fig. 1-6).


BAD AVG GOOD
AtzA_: ---MQ-TLSIQ--HG-TLVTMDQYRRVLGDSWVHVQDGRIVALGVHAESVPPP-AD-------
TrzN : ------MILIR--GLTRVITFDDQERELEDADILIDGPKIVAVGKD--LSDRS-VS-------
AtzC_: MSKDF-DLIIR--NA-YLS-----E-KDSVYDIGIVGDRIIKIEA---KIEGT-VK-------
AtzB_: ----MTTTLYTGFHQ--LVTGDVAGTVLNGVDILVRDGEIIGLGP---DLPRTLAPIGVGQEQ
cons
AtzA_: --RVIDARGKVVLPGFINAHTHVNQILLRGG--P--S----HGR-QFYDWLFNVVYPGQK---
TrzN : --RTIDGRGMIALPGLINSHQHLYEGAMRAI-PQ--LE----RV-TMASWLEGV-LTRSAGWW
AtzC_: --DEIDAKGNLVSPGFVDAHTHMDKSFTSTG-ER--LPKFWSRPYTRDAAIEDG-LKYIK---
AtzB_: GVEVVNCRGLTAYPGLINTHHHFFQAFVRNLAPLDWTQ-----L-DVLAWLRKI-YPVFA---
cons
AtzA_: ---AMRPE-DVAVAVRLYCAEAVRSGITTINENADSAIY----P-GNI-EAAMAVYGEV--GV
TrzN : RDGKFGPD-VIREVARAVLLESLLGGITTVADQHLFFPGAT--ADSYI-DATIEAATDL--GI
AtzC_: ---NATHE-EIKRHVIEHAHMQVLHGTLYTRTHVDVDSVAK--T-KAV-EAVLEAKEELKDLI
AtzB_: ----LVDEDCIYHSTVVSMAELIKHGCTTAFDHQYNYS--RRGG-PFLVDRQFDAANLL--GL
cons
AtzA_: RVVYARMF----FDR----MDGRIQGYVDALKARSPQVELCSIMEETAVAKD---RI-----T
TrzN : RFHAARSS----MTL----GKSE-----------GGFCDDLFVEPVDRVVQ---HC-----L
AtzC_: DIQVVAFA----QSG----FF----------VDLESES---LI-----R
AtzB_: RFHAGRGCITLPMAEGSTIP----------DAMRESTDTFLADCE
cons

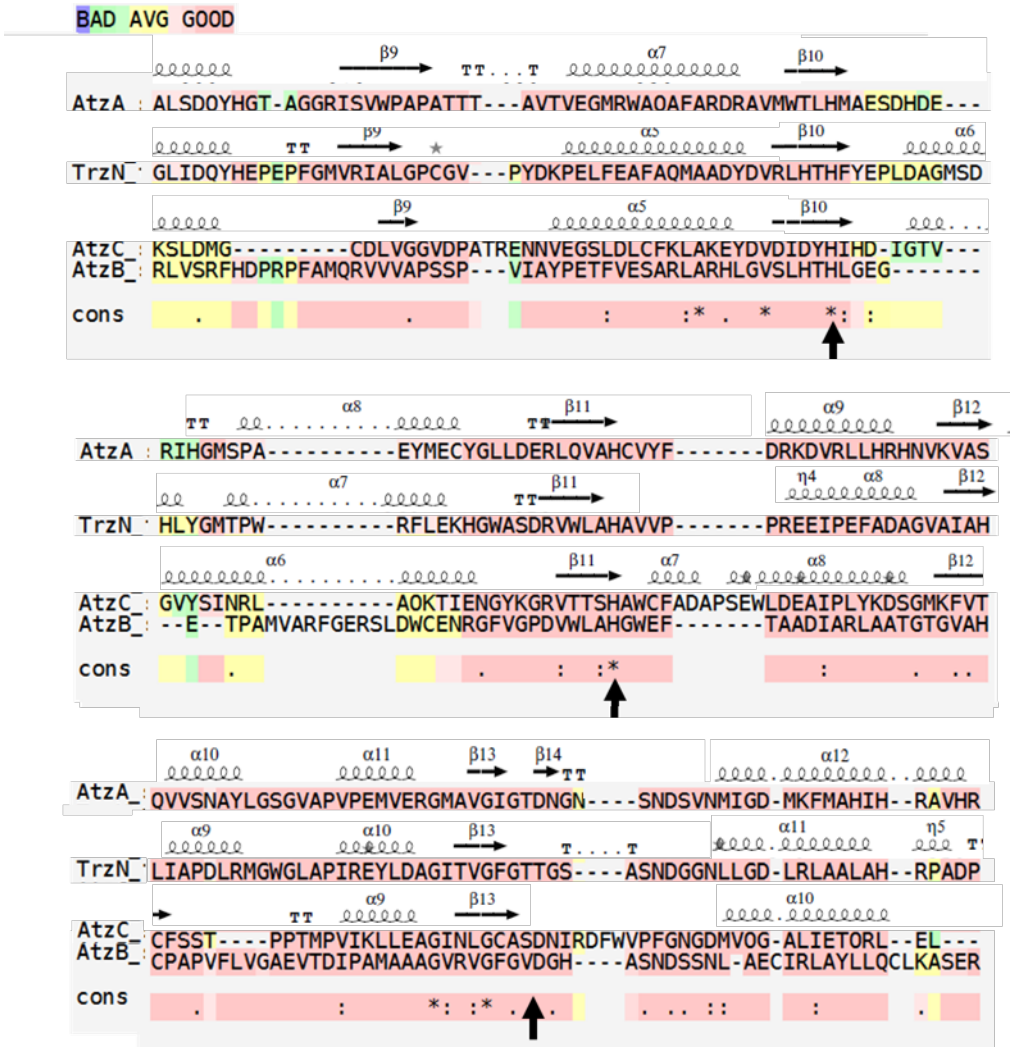


Figure 1-6 Alignment of AtzA, TrzN, AtzC and AtzB amino acid sequences with T-COFFEE (164-168). *The black arrows point towards the four conserved Histidine residues and the aspartate (or threonine in the case of TrzN) involved in the metal binding centre of the enzyme's active sites. The colour reflects the quality of the alignment and the legend is shown on the top left corner. For reason of readability the C-terminal sequences are truncated at amino acid for 353 AtzA, 351 for TrzN, 330 for AtzC and 359 for AtzB. The Beta The secondary sequences of AtzA, TrzN and AtzC have been added using ESPript (169); The η symbol refers to a 3_{10}-helix, α-helices, 3_{10}-helices and π-helices are displayed as medium, small and large squiggles, respectively. β-strands are rendered as arrows, strict β-turns as **TT** letters and strict α-turns as **TTT**. residues in an alternate conformation are highlighted by a grey star.*

1.4.2 Structure-function of the amidohydrolases of the lower pathway

Unlike the enzymes of the upper pathway, the enzymes of the lower pathway do not belong to the amidohydrolase superfamily, instead they belong to the amidase and cyclic amide hydrolase families. Their characteristics can be found in Table 1-5.

AtzD (E.C. 3.5.2.15) produces 1-carboxybiuret by opening the cyanuric acid ring, its k_{cat} and K_M values are 17 sec^{-1} and 350 µM, respectively (Table 1-5). It forms a compact tetramer. Each monomer has been found to contain a metal centre coordinating a Mg^{2+} metal ion; however, it is involved in structure stabilisation and is not catalytic (145) (Fig. 1-7). Phylogenetic analysis of the sequences of AtzD (114) and TrzD (59) with homologous sequences, has revealed that these enzymes belong to a unique family of enzymes (113, 145, 170), which also includes barbituric acid hydrolase (BAH), an enzyme involved in the oxidative pathway of pyrimidines (171). The structural characterisation of AtzD from *Pseudomonas* (145) and cyanuric acid hydrolases (CAH) from *Frankia* sp. strain EuI1c and *Azorhizobium caulinodans* ORS 571 (144, 149) confirmed the novelty of the protein family. AtzD was found to possess a novel fold, named the "Toblerone fold" in reference to the three-fold pseudo-symmetry of the AtzD monomer (Fig. 1-7) (145). The catalytic mechanism of AtzD, probed by mutation and inhibition studies, depends upon a Ser-Lys catalytic dyad (144, 145).

Table 1-5 Characteristics of the amidohydrolases of the lower pathway. *NA: not available.*

E.C. No.	Enzyme	Physiological Function	Kinetic Values[1]	Approx. Size	PFam Family	Accession[2]	Mechanism	Refs
3.5.2.15	AtzD	Cyanuric acid amidohydrolase	K_M = 350 k_{cat} = 17 k_{cat}/K_M = 4.8 x 10^4	Tetramer (4 x 19,000 Da)	Toblerone	**UniProt:** P58329 **PDB:** 4bvq, 4bvr, 4bvs, 4bvt, 4nq3 5t13	Ser-Lys hydrolase	(113, 144, 145, 148, 149)
3.5.1.84	AtzE	Biuret amidohydrolase	NA	NA	Amidase	**UniProt:** Q936X3 **PDB:**	Ser-*cis*Ser-Lys hydrolase	(115, 118)
3.5.1.84	BiuH	Biuret amidohydrolase	K_M = 80 k_{cat} = 12 k_{cat}/K_M = 1.5 x 10^5	Tetramer (4 x 27,000 Da)	Isochorismatase	**UniProt**: Q1M7F4 **PDB:**	Cysteine hydrolase	(118)
3.5.1.54	AtzF	Allophantate amidohydrolase	K_M = 1500 k_{cat} = 16 k_{cat}/K_M = 1.1 x 10^4	Tetramer (4 x 36,000 Da)	Amidase	**UniProt:** Q936X2 **PDB:** 4cp8	Ser-*cis*Ser-Lys hydrolase	(115, 116, 138, 142)

[1]Values are given for the physiological substrate only. Units for kinetic values are as follows: K_M, μM; k_{cat}, s^{-1}; k_{cat}/K_M, $s^{-1}.M^{-1}$

[2]Accession numbers are given for the 'archetypal' enzymes only

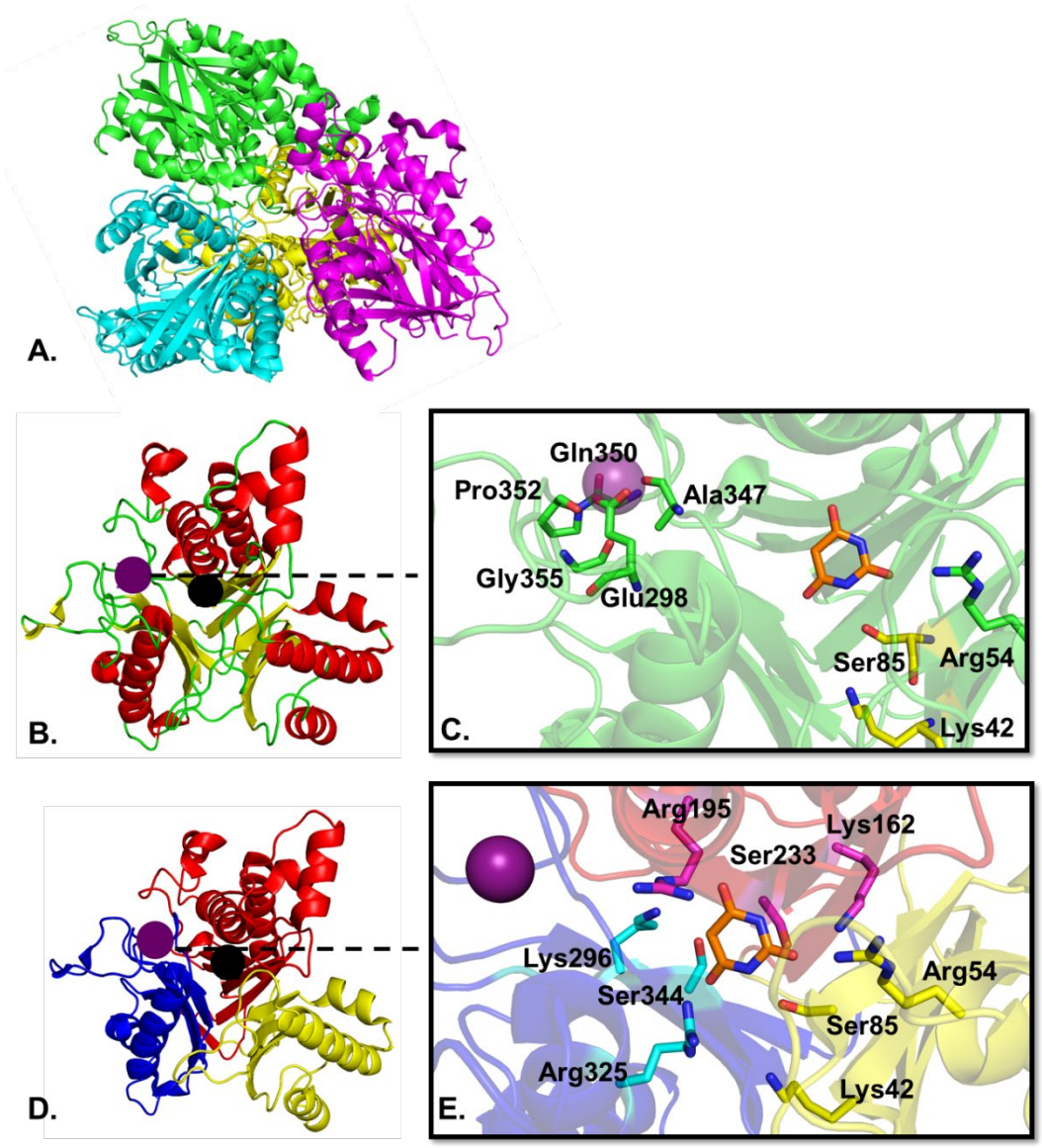


Figure 1-7 Structure of AtzD from* Pseudomonas *sp. ADP (PDB 5hxu). A. *Cartoon representation of the AtzD tetramer;* ***B.*** *Monomer of AtzD showing α-helixes in green, the loops in red and the β-sheets in yellow;* ***C.*** *Zoom in the active site showing the amino acids coordinating the metal ion in green sticks, the catalytic amino acid of the dyad ser85-Lys42 in yellow stick and the Arg54 in green stick around the inhibitor barbituric acid (145);* ***D.*** *Trimeric monomer of AtzD, each repeat unit (RU) is shown in a different colour;* ***E.*** *Zoom in the active site showing the three sets of potential catalytic amino acids around the barbituric acid inhibitor: in cyan sticks for the blue RU, in pink stick for the red RU, in yellow stick for the yellow RU. The black dot represents the location of the active site, the purple dot the location of the metal ion, the Mg^{2+} ion is shown as a purple sphere and barbituric acid in orange stick..*

Due to the three-fold symmetry of the active site, the AtzD monomer contains three potential Ser-Lys dyads: Lys42 and Ser85, Lys162 and Ser233, or Lys296 and Ser344 (Fig. 1-7B,C). It is uncertain which of the Ser-Lys dyads is catalytic; but X-ray data from the *Pseudomonas* AtzD structure suggests that Ser85-Lys42 may form the catalytic dyad (145), while phylogenetic studies suggest that Ser226-Lys156 may fulfil this function (144). The active site of AtzD is largely composed of polar amino acids that stabilise cyanuric acid substrate through an extensive H-bonding network. The reaction mechanism resembles that of other serine hydrolases (Fig. 1-8); the catalytic lysine acts as a general base and deprotonates the serine that in turn attacks cyanuric acid, leading to the formation of a covalent acyl intermediate. Finally, a water molecule attacks the acyl:enzyme complex regenerating the active site serine and liberating 1-carboxybiuret (145).

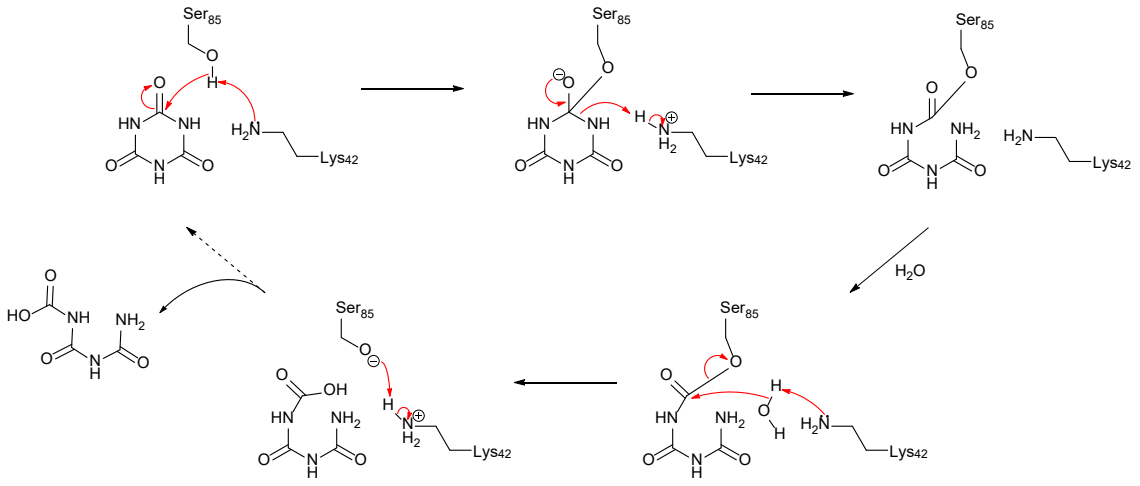


Figure 1-8 Example of a serine hydrolase mechanism: the AtzD mechanism*. catalytic lysine acts as a general base and deprotonates the serine that in turn attacks cyanuric acid, leading to the formation of a covalent acyl intermediate. Finally, a water molecule attacks the acyl:enzyme complex regenerating the active site serine and liberating 1-carboxybiuret (145).*

AtzE and AtzF both belong to the amidase enzyme signature family, characterised by a glycine serine rich sequence (Fig. 1-9) (115, 172). Up until now all attempts to produce AtzE in *E. coli* had failed, and consequently AtzE had not been characterised prior to the beginning of the investigation reported in Chapter 3. However, AtzF has already been biochemically characterised. It is an allophanate amidohydrolase (AH) from the amidase family and possesses the signature amidase catalytic triad (Lys65, *cis*Ser165 and Ser189), which was confirmed *via* mutagenesis and through inhibition studies (115). AtzF is a homotetramer and its k_{cat} and K_M values with allophanate are 16 sec^{-1} and 1500 µM, respectively (Table 1-5) (115).

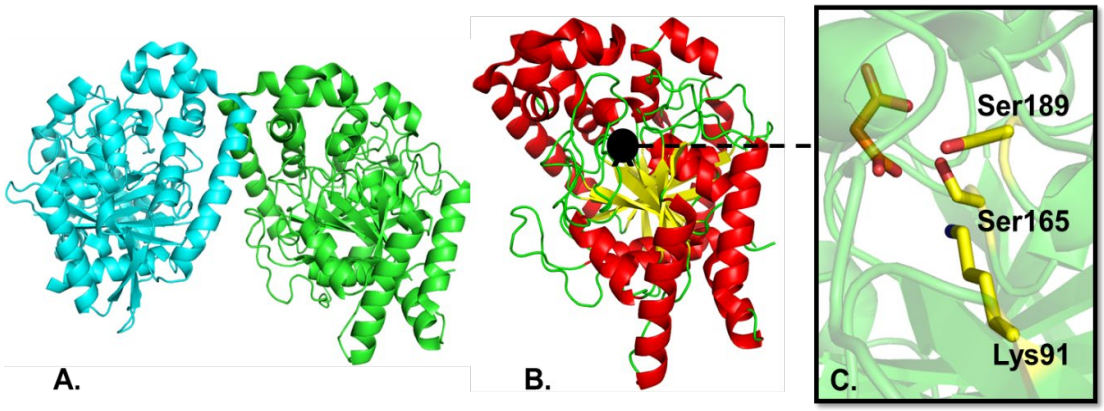


Figure 1-9 Structures of the amidase domain of AtzF from* Pseudomonas *sp. ADP (PDB 4cp8). A. *Cartoon representation of the AtzF dimer;* ***B.*** *AtzF monomer, where the black dot indicates the location of the active site;* ***C.*** *Zoom onto the active site with the amino acids of the catalytic triad are shown in yellow sticks around the malonate inhibitor shown in orange stick.*

AtzF is larger than most amidase family enzymes, as it possesses a *ca.* 15 kDa C-terminal extension (173). The X-ray structure of N-terminal amidase domain of AtzF was obtained. This fragment of AtzF was catalytically competent, with identical K_M and k_{cat} values to the full length protein. Indeed, the only function so far ascribed to the C-terminal extension is to co-ordinate the quaternary structure of the protein, as the truncated amidase domain is dimeric, not tetrameric (173).

The structure of the AtzF amidase domain, truncated from its C-terminal domain, confirmed that the enzyme adopts the α-β-α fold common to this enzyme family (141, 142) (Fig. 1-9). Amidase signature enzymes are found across domains of life and they are involved in a broad range of reactions (172). The alignment of five amidases, sharing the same catalytic mechanism as AtzF, are shown in Fig. 1-10: the fatty acid amide hydrolase from the rat (FAAH) (174), the malonamidase E2 from *Bradyrhizobium japonicum* (175), the peptide amidase from *Stenotrophomonas maltophilia* (176), the AtzF and AtzE from *Pseudomonas* sp. strain ADP (110). This illustrates the conservation of the AS signature and of the catalytic triad Lys Ser Ser (164-168).

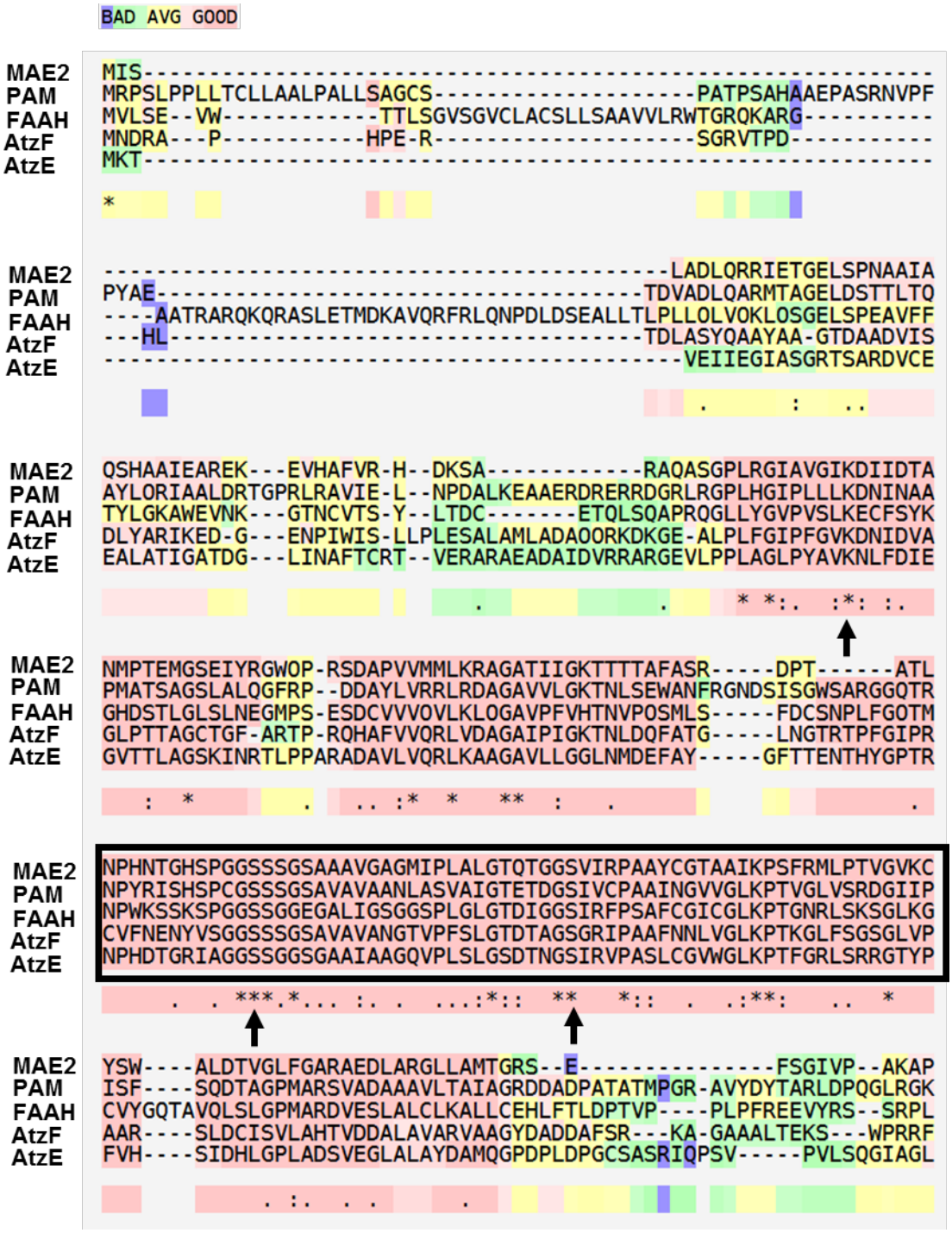


Figure 1-10 Alignment of five AS signature amidase enzymes with T-COFFEE (162-166). *The arrows points towards the catalytic triad Lys* cis*Ser Ser, the black box contains the the Glycine Serine rich sequence called the AS signature, characteristic of this enzyme family.*

The AtzF mechanism is essentially identical to the mechanisms found in other amidases such as the FAAH amidase (176). Ser189 is activated by the Lys-*cis*Ser proton shuttle, and then forms the acyl:enzyme intermediate via the nucleophilic attack on the substrate (liberating ammonia). Finally, a water molecule hydrolyses the acyl-enzyme intermediate releasing dicarboxyamonia and regenerating Ser189 (176, 177).

In addition to its role in cyanuric acid catabolism, allophanate amidohydrolase is also found as part of a multiprotein complex (urea carboxylase) involved in urea degradation in some bacteria, algae, fungi and yeast (178). Urea carboxylase contains two enzymes, urea carboxylase and allophanate amidohydrolase. Urea carboxylase performs an ATP-dependent carboxylation of urea to form allophanate and allophanate amidohydrolase deaminates allophanate, forming ammonia and carbon dioxide (141, 179). Interestingly, Balotra *et al.* suggest that like urea carboxylase, AtzD, AtzE and AtzF may also form a multiprotein complex (141).

As the *Pseudomonas* sp. strain ADP biuret hydrolase (AtzE) was difficult to produce in *E. coli* (110, 116, 118), Cameron *et al.* identified and characterised the biuret hydrolase (BiuH) from *Rhizobium leguminosarum* bv. *viciae* 3841 BiuH. Unrelated to the AtzE amidase, BiuH is a member of the isochorismatase family with homology to cysteine hydrolases. BiuH hydrolyzes biuret to form allophanate (k_{cat}/K_M = 1.5 x 10^5), but does not accept 1-carboxybiurte as a substrate (118). Homologs of BiuH have been described in other bacteria, but are infrequently associated with an AtzD homologue, which may suggest a broader role for biuret catabolism in bacteria (180).

1.5 Overview of the PhD

The work reported in this thesis focuses on the variation in the composition of the cyanuric acid degradation pathways from *Rhizobium leguminasorum* bv. *viciae* 3841 and of *Pseudomonas* sp. strain ADP. AtzE was identified more than twenty years ago (110), and has been identified in the genomes of many *s*-triazine degrading strains (Table 1-1). However, AtzE has remained uncharacterised biochemically or structurally due to its recalcitrance to heterologous expression. Its physiological counterpart in *Rhizobium leguminosarum* bv. *viciae* 3841, BiuH, had been partially biochemically characterised (118), but lacked characterisation at an atomic level.

The goal of my PhD was to obtain and study the X-ray structures of both AtzE and BiuH, to elucidate the structure-function relationships for these two enzymes. Previous studies of this type with xenobiotic-degrading enzymes have been informative about their recent evolution, and it was hoped that this study would provide insights into the convergent evolution of cyanuric acid degradation pathways in these two bacteria.

In Chapter two, BiuH is characterised *Rhizobium leguminasorum* bv. *viciae* 3841, with a particular focus on developing a deeper understanding of the structure-function relationship of this previously under-characterised amidohydrolase. BiuH was purified to apparent homogeneity, its structure obtained and structure function analysis performed by mutagenesis and molecular dynamics modelling.

The third chapter of this PhD focuses on gaining a detailed understanding of the biochemistry and structure of AtzE from *Pseudomonas* sp. strain ADP, which had

previously been identified as a functional analogue of BiuH. Careful and detailed charactisation of AtzE purified from *Pseudomonas* led to the revision of the cyanuric acid catabolism pathway and to the identification of two small and previously overlooked proteins, the function of one of these new proteins (AtzG) is also explored in this chapter.

In Chapter four, the function and structure of the second newly identified small protein in the cyanuric acid catabolism (AtzH) is explored. AtzH is purified to apparent homogeneity from E. coli, its structure obtained and its potential function analysed biochemically, through substrate docking and by mutagenesis. From these data it is inferred that AtzH is like an enzyme acting on a metabolite of the recently revised cyanuric acid catabolic pathway (1,3-dicarboxyurea).

2 Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841

2.1 Overview

The exploration of cyanuric acid degradation pathways in different bacteria showed that substantial differences exist in the CA degradation pathways between microorganisms (153). In *Rhizobium leguminasorum* bv. *viciae* 3841, for example, a biuret hydrolase (BiuH) belonging to the isochorismatase family performs the deamination of biuret to produce allophanate (118); whereas, in the model bacterium *Pseudomonas* sp. strain ADP, it is an amidase that is thought to perform that step. In this chapter, the structure-function study of the biuret hydrolase BiuH is presented. The atomic structure of BiuH was solved and site-directed mutagenesis was used to gain a better understanding of the BiuH catalytic mechanism. Additionally, molecular dynamics simulations highlighted the presence of three channels from the active site to the enzyme surface forming potential substrate channel and a co-product (ammonia) channel and a co-substrate (water) channel.

2.2 Statement of contribution

I performed the cloning, mutagenesis, protein expression and enzyme kinetics. I also analysed and interpreted the data. The crystallography and structural determination were performed by Dr Tom Peat and Dr Janet Newman. All the molecular dynamics work was performed by Dr Del Lucent.

Publication status: Published 01/2018.

Esquirol L, Peat TS, Wilding M, Lucent D, French NG, Hartley CJ, Newman J, Scott C. 2018. Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. viciae 3841. PLoS One **13**.

Australian National University

Statement of Contribution

This thesis is submitted as a Thesis by Compilation in accordance with https://policies.anu.edu.au/ppl/document/ANUP_003405

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

Title and authors:
Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841. Esquirol L*, Peat TS*, Wilding M, Lucent D, French NG, Hartley CJ, Newman J, Scott C.

Current status of paper:
Published 01/2018

Contribution to paper:
I performed the cloning, mutagenesis, protein expression and enzyme kinetics. I also analysed and interpreted the data. The crystallography and structural determination were obtained by Dr Tom Peat and Dr Janet Newman. All the molecular dynamics work was performed by Dr Del Lucent.

Senior author's endorsement:
Lygie's assessment of her contribution to this work is accurate.

Dr. Colin Scott, senior author

LYGIE ESQUIROL
Candidate – Print Name

Signature

25/05/2018
Date

Endorsed

Prof. C. J. Easton
Chair of Supervisory Panel – Print Name

Signature

29 May 2018
Date

G. Otting
Delegated Authority – Print Name

Signature

30 May 2018
Date

2.3 Publication: Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841

PLOS ONE

RESEARCH ARTICLE

Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841

Lygie Esquirol[1,2☯], Thomas S. Peat[3☯], Matthew Wilding[2,3], Del Lucent[4], Nigel G. French[1], Carol J. Hartley[1], Janet Newman[3], Colin Scott[1]*

1 CSIRO Biocatalysis and Synthetic Biology, Canberra, Australian Capital Territory, Australia, **2** Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory, Australia, **3** CSIRO Biomedical Manufacturing, Parkville, Melbourne, Victoria, Australia, **4** Department of Electrical Engineering and Physics, Wilkes University, Wilkes-Barre, Pennsylvania, United States of America

☯ These authors contributed equally to this work.
* colin.scott@csiro.au

Citation: Esquirol L, Peat TS, Wilding M, Lucent D, French NG, Hartley CJ, et al. (2018) Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841. PLoS ONE 13(2): e0192736. https://doi.org/10.1371/journal.pone.0192736

Editor: Renwick Dobson, University of Canterbury, NEW ZEALAND

Received: November 20, 2017

Accepted: January 29, 2018

Published: February 9, 2018

Data Availability Statement: Data are available from the Protein Data Bank (accession numbers: 6AZO, 6AZN, 6AZQ, 6AZS, 5BK6).

Funding: The author(s) received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Biuret deamination is an essential step in cyanuric acid mineralization. In the well-studied atrazine degrading bacterium *Pseudomonas* sp. strain ADP, the amidase AtzE catalyzes this step. However, *Rhizobium leguminosarum* bv. *viciae* 3841 uses an unrelated cysteine hydrolase, BiuH, instead. Herein, structures of BiuH, BiuH with bound inhibitor and variants of BiuH are reported. The substrate is bound in the active site by a hydrogen bonding network that imparts high substrate specificity. The structure of the inactive Cys175Ser BiuH variant with substrate bound in the active site revealed that an active site cysteine (Cys175), aspartic acid (Asp36) and lysine (Lys142) form a catalytic triad, which is consistent with biochemical studies of BiuH variants. Finally, molecular dynamics simulations highlighted the presence of three channels from the active site to the enzyme surface: a persistent tunnel gated by residues Val218 and Gln215 forming a potential substrate channel and two smaller channels formed by Val28 and a mobile loop (including residues Phe41, Tyr47 and Met51) that may serve as channels for co-product (ammonia) or co-substrate (water).

Introduction

The mineralization of cyanuric acid by bacteria is thought to be an ancient metabolic pathway [1]. It is thought that this pathway has been recently 'co-opted' into pathways for the degradation of highly functionalized *s*-triazines as they have become environmentally abundant through human activities since the mid-twentieth century [1–3]. The s-*triazine* mineralization pathways, including the cyanuric acid catabolism pathway, are thought to have evolved in response to an increase in the abundance of *s*-triazines in the environment as a result of human activities [2,4,5]. Although most incidentally exposed bacteria are not sensitive to the *s*-triazines, these anthropogenic compounds are an excellent nitrogen source and bacteria that

can access this nitrogen may have a growth advantage compared with those that cannot: i.e., the ability to use *s*-triazines as a nitrogen source confers a selective advantage [6].

The canonical cyanuric acid catabolism pathway was first described from the atrazine-degrading *Pseudomonas* sp. strain ADP, which converts herbicidal chloro-*s*-triazines to cyanuric acid *via* three sequential hydrolyses that first dechlorinate and then dealkylate the herbicides. These steps are catalyzed by AtzA [7], AtzB [8] and AtzC [9], all of which are metalloenzymes with an amidohydrolase fold (PFAM PF01979). Cyanuric acid is then mineralized by three hydrolases (Fig 1): AtzD [10–12], AtzE [13,14] and AtzF [15–18]. AtzE and AtzF are Ser-*cis*Ser-Lys amidohydrolases and both possess an amidase-fold (PFAM PF01425); the structure of AtzF has been determined experimentally [15,19], while that of AtzE has been inferred from its sequence identity with other amidase proteins [14]. AtzD is also a serine hydrolase, but it belongs to a recently described structural family (the Toblerone fold; PFAM PF09663) that is unrelated to Ser-*cis*Ser-Lys hydrolases [10,11,20].

Other *s*-triazine catabolizing enzymes and pathways have also evolved in response to the increased abundance of *s*-triazines in the environment. For example, AtzA is substituted for a physiologically isofunctional, but non-homologous amidohydrolase TrzN in some bacterial species [21–23]. TriA, a melamine aminohydrolase that is 98% identical to, but biochemically distinct from AtzA, allows the use of the triamino triazine melamine as a nitrogen source in *Pseudomonas* sp. strain NRRL B-12227 [24–26]. *Rhizobium leguminosarum* bv. *viciae* 3841 has an unusual cyanuric acid catabolic pathway that uses an *atzD* homolog to ring-open cyanuric acid, but lacks an *atzE* homolog to catalyze the subsequent step [27]. Instead, *R. leguminosarum* bv. *viciae* 3841 employs a cysteine hydrolase that fulfils the role of a biuret amidohydrolase (BiuH). The genes encoding the cyanuric acid amidohydrolase and biuret hydrolase are co-located on one of six large plasmids (the ~0.5 Mbp 'symbiosis' plasmid, pRL10 [27]), albeit this location may not have a specific functional consequence as most *s*-triazine-degradation genes are associated with mobile genetic elements [2,5].

Potential applications for many of the genes and enzymes of the *s*-triazine degrading pathway have been identified, including environmental bioremediation and the detection of adulterants in food and animal feed [28–32]. Biuret is sometimes used as an inexpensive adulterant to increase the total nitrogen content of food and feed, thereby increasing its value [33]. Biuret is also a side product in the production of urea fertilizer and can be phytotoxic [34,35]. BiuH may have utility in detecting, quantifying and degrading biuret in food, feed and fertilizer.

Herein, we describe and characterize the structure and catalytic mechanism of BiuH. Through a combination of structural studies, mutagenesis and molecular modelling we are able to propose a plausible catalytic mechanism for BiuH, which closely resembles that of other cysteine hydrolases. We also identify three potential channels from the active site to the bulk solvent, for the ingress of substrates (biuret and water) to the active site and the egress of products (ammonia and allophanate).

Materials and methods

Cloning

A synthetic version of the biuret hydrolase gene (*biuH*) from *Rhizobium leguminasorum* bv. *viciae* 3841, codon optimized for expression in *E. coli*, encoding for a protein identical to Q1M7F4 (Uniprot), EMBL database accession no. AM236084.1 was ordered from GenScript (Piscataway, NJ, USA; S1 Fig) and provided as an insert in pUC57 (Genscript) with *Nde*I and *Bam*HI (New England Biolabs) restriction sites engineered 3' and 5', respectively, of the structural gene. The gene was subcloned into the *Nde*I and *Bam*HI sites of pETcc2, described

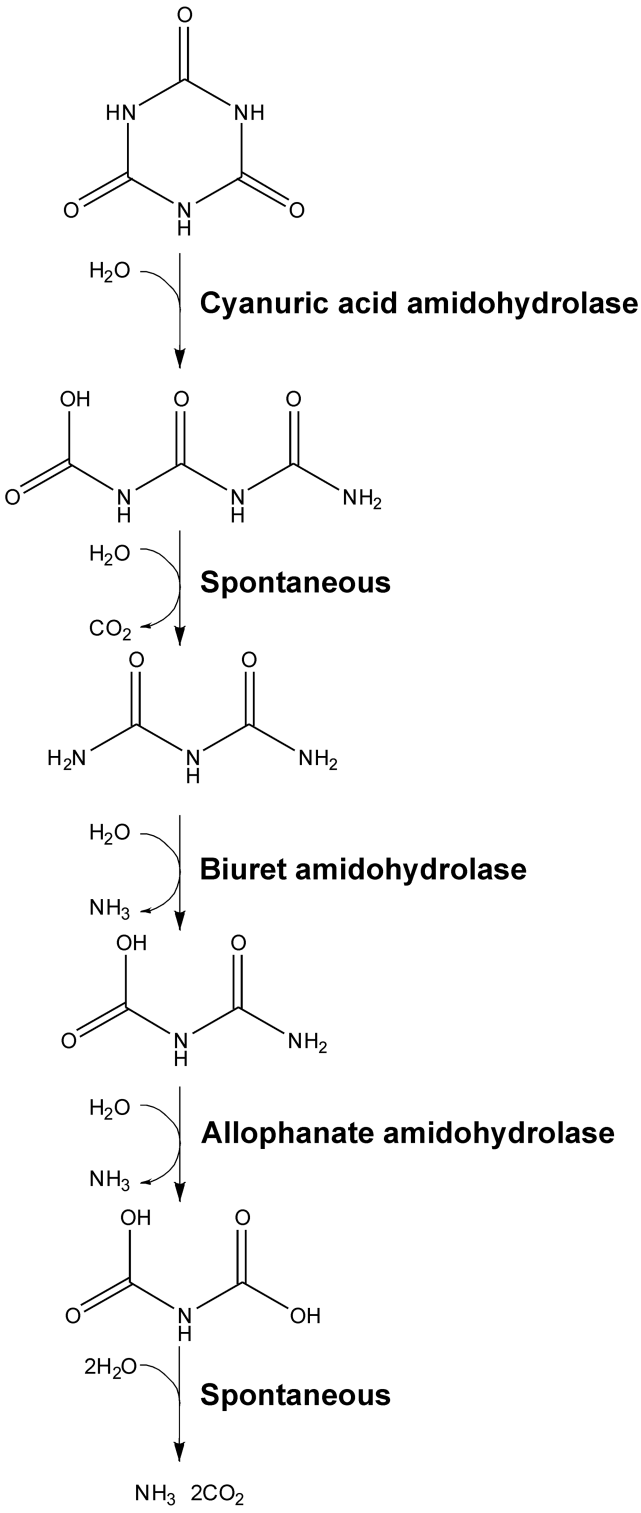


Fig 1. Cyanuric acid mineralization by bacteria. Cyanuric acid is mineralized to CO_2 and NH_3 by cyanuric acid mineralizing bacteria by the enzymes cyanuric acid amidohydrolase, biuret hydrolase and allophanate hydrolase. The product of cyanuric acid amiidohydrolase (1-carboxybiuret) is unstable under physiological conditions and decarboxylates to form biuret and CO_2. The product of allophanate amidohydrolase (dicarboxyammonia) is also unstable under physiological conditions and decomposes to CO_2 and NH_3.

https://doi.org/10.1371/journal.pone.0192736.g001

in Peat *et al*., 2013. A sequence coding for a 6xHis-tag before a thrombin cleavage site (MGSSHHHHHHSSGLVPRGSH; S1 Fig) was introduced by the subcloning to facilitate protein purification.

Genes encoding single amino acid substitution variants of BiuH (Asp36Ala, Asp36Asn, Asp36Gln, Asp36Glu, Phe41Ala, Phe41Leu, Phe41Tyr, Phe41Trp, Lys142Ala, Lys142His and Lys142Arg, Lys145Ala Lys145His, Lys145Arg, Cys175Ala, Cys175Ser, Gln215Ala, Gln215Asn, Gln215Asp and Gln215GluGln215Glu) were produced by overlap extension PCR as described in Ho *et al* [36]. The *biuH* gene was used as template and mutagenic primers were obtained from Integrated DNA Technologies (IDT, Singapore) and their sequences are detailed in S1 Table. Mutant *biuH* genes were cloned into the pETcc2 expression vector using *Nde*I and *Bam*HI.

Protein expression and purification

pETcc2 derivatives encoding the wild-type biuret hydrolase and twenty variants were used to transform *Escherichia coli* BL21 (λDE3) cells (New England Biolabs). Bacterial cultures were grown on Luria-Bertani (LB) medium, supplemented with 100 μg/mL ampicillin where required. Cells were grown with shaking at 200 rpm at 37 ˚C for the wild-type biuret hydrolase and 28 ˚C for the variants. Protein expression was induced at an OD_{600} of 0.8 by addition of 1 mM isopropyl β–D-1-thiogalactopyranoside (IPTG).

A seleno-L-methionine (SeMet)-BiuH was obtained by growth in minimal medium containing 60 mg/L of SeMet as the only source of methionine, as described in Doublié 1997 [37]. Purification was performed as described below, except that 0.2 mM EDTA and 5 mM dithiothreitol (DTT) were added to all the buffers in order to retard oxidation of the SeMet-BiuH during purification.

Cells were harvested 24 hours after induction by centrifugation at 5,000 x *g* for 15 minutes, resuspended in lysis buffer (5 mM imidazole, 25 mM potassium phosphate pH 7.5) and lysed by passage through a Microfluidics homogenizer M-110P (Massachusetts, USA) five times at 15,000 PSI. The lysis was followed by centrifugation at 18,000 x *g* for 45 minutes, using an Aventi J-E centrifuge to pellet the cells debris and the soluble fraction was used for further purification.

The soluble fraction was syringe filtered through a 0.22 μm filter. The filtrate was applied to a 5 mL Ni-NTA Superflow cartridge (GE Healthcare) and protein eluted with a gradient from 5 mM to 500 mM imidazole in ten column volumes (CV). SDS-PAGE analysis was performed to assess the purity of the fractions, using NuPAGE® Novex 4–12% acrylamide gradient Bis-Tris gels (Invitrogen). Fractions that eluted between 150–350 mM imidazole were found to contain a protein with a mass corresponding to that of biuret hydrolase.

After pooling, the protein containing fractions were concentrated to 12 mL using an Amicon Ultra-15 centrifugal filter unit and further purified by size exclusion chromatography using a 130 mL column packed with Superdex 200 preparation grade resin (GE Healthcare Life Sciences), equilibrated initially with 50 mM HEPES pH 7.5, 100 mM NaCl. After a preliminary results from the differential scanning fluorimetry (DSF) the gel filtration buffer was swapped to one containing 50 mM Tris pH 7.5, 100 mM NaCl. All chromatography steps were performed using an ÄKTA purifier UPC 10 (GE Healthcare Life Sciences).

Protein concentration was estimated using NanoDrop spectrophotometer (Thermo Pierce) by reading the absorbance at λ = 280 nm. The molar extinction coefficients used for BiuH and its variants was 34,295 $M^{-1}.cm^{-1}$, except for BiuH Phe41Tyr (25,785 $M^{-1}.cm^{-1}$) and BiuH Phe41Trp (39,795 $M^{-1}.cm^{-1}$). Molar extinction coefficients were calculated using ProtParam on the ExPasy server (https://web.expasy.org/protparam/).

Differential scanning fluorimetry (DSF)

DSF was used to determine an appropriate formulation for crystallization trials, using a standard, published protocol [38]. Briefly, 0.3 uL of protein and 0.3 uL of a 1:20 dilution of Sypro dye (Sigma S5692) was diluted into a final volume of 20 uL, and heated in steps of 0.5 ˚C from 20–90 ˚C in an RT-PCR machine (BioRad CXF 96). The protein at 3.6 mg/mL in a preliminary size exclusion formulation (50 mM HEPES pH 7.5, 100 mM NaCl) was tested in triplicate against an array of 13 different buffers at pH values ranging from 5 to 9, at two concentrations of NaCl ('Buffer Screen 9',). The T_m in the HEPES/NaCl formulation was 47.8+/-0.1 ˚C (all T_m estimations were extracted using the program Meltdown [39]). The protein was slightly more stable in 50 mM Tris chloride pH 8, 50 mM NaCl (51.5+/- 0.2 ˚C). The protein was simultaneously treated with thrombin to remove the N-terminal His-tag (100 uL protein at 3.6 mg/mL was added to a 0.2 mL tube containing 10 units of lyophilised thrombin, and enough $CaCl_2$ to give a final concentration of 3 mM). This mix was dialyzed into 50 mM Tris pH 8, 50 mM NaCl (3.5 kDa cutoff membrane) and this sample was assayed without further purification against the same buffer screen by DSF. This increased the T_m to 55.1 ˚C. A Meltdown report for the Buffer Screen 9 analysis of the thrombin/tris treated protein is included as supplementary information (S2 Fig).

To confirm the proper folding of the BiuH variants, thermal melt analyses were performed on each variant, using the N-terminal His tagged protein. Protein concentration was 5 mg/mL, in 50 mM Tris chloride pH 7.5, 100 mM NaCl and 0.2 μL of 1:20 diluted Sypro Orange dye was added to each 20 μL experiment. The samples were run in 3–8 fold replication. The wild-type protein shows a T_m of around 56 ˚C, similar to that found from buffer screen 9, the variants ranged from 41 ˚C to 68 ˚C, but all showed a clear melting transition. See S2 Table and S3 Fig.

Protein crystallization and structure solution

Modestly diffracting (≈3Å) wild type BiuH crystals suitable for X-ray analysis were eventually grown using a combination of seeding, *in situ* proteolysis, formulation variation and additive screening. The best native crystal tested was a thin, stacked plate grown from protein at 3 mg/mL in Tris/NaCl buffer, with an *in situ* chymotrypsin treatment, where 100 uL protein solution was added to 10 μg of freeze-dried chymotrypsin and this mix was set up with no further purification. The reservoir consisteded of 0.18 M lithium chloride, 0.3 M NDSB 195 (non-detergent sulfobetaine 195), 18% polyethylene glycol 6000, 0.09 M sodium MES pH 6. The crystals grew in droplets of 200 nL protein + 200 nL reservoir, and were set up in SwissSci SD2 sitting drop plates (Molecular Dimensions, UK) at 20 ˚C. As there was no obvious molecular replacement model available, SeMet protein in 50 mM Tris pH 8, 50 mM NaCl, 5 mM DTT at 10 mg/mL was set up against an optimization screen based around the successful wild-type BiuH condition, with the same *in situ* chymotrypsin treatment. A SeMet crystal from this optimization screen was harvested, and used to collect single wavelength anomalous data at 0.97919 wavelength to 2.46 Å resolution at the Australian Synchrotron MX2 beamline. The data showed a strong anomalous signal (CCanom > 0.15) to about 2.90 Å and the structure was solved using Crank2 [40] which automatically built 884 residues in 19 chains in the C2 spacegroup. The structure was manually rebuilt to give four independent chains in the asymmetric unit which formed a tight tetramer. Several higher resolution data sets (Table 1) of mutants and soaks were available and this model was used to solve these structures using Phaser [41] (in two new space groups, $P2_12_12_1$ and $P22_12_1$). The $P22_12_1$ spacegroup (the two K142 mutant structures) also has a single tetramer in the asymmetric unit whereas the $P2_12_12_1$ spacegroup (the C175S structure with and without biuret) has two tetramers in the asymmetric unit. The structures

Table 1. Data collection and refinement statistics.

Data Collection					
PDB code	**6AZO**	**6AZN**	**6AZQ**	**6AZS**	**5BK6**
Crystal	SeMet	C175S	C175S + biuret	K142A	K142H
Spacegroup	C2	$P2_12_12_1$	$P2_12_12_1$	$P22_12_1$	$P22_12_1$
Cell (a x b x c)	135.7 x 101.0 x 65.6	74.2 x 86.9 x 343.0	73.7 x 87.3 x 341.7	62.1 x 122.2 x 136.1	62.1 x 122.7 x 135.7
Cell (α x β x γ)	90 x 91.8 x 90	90 x 90 x 90	90 x 90 x 90	90 x 90 x 90	90 x 90 x 90
Resolution (Å)	2.46	1.75	2.22	1.59	1.59
Completeness (%)	99.7 (98.0)	100 (100)	99.9 (98.6)	100 (100)	100 (100)
Rmerge %	0.279 (0.713)	0.090 (0.759)	0.159 (0.726)	0.183 (1.122)	0.071 (0.726)
Rpim %	0.091 (0.425)	0.056 (0.546)	0.093 (0.434)	0.074 (0.530)	0.045 (0.588)
Mean I/sigI	11.4 (2.7)	9.0 (1.1)	7.7 (2.5)	8.2 (1.6)	11.7 (1.6)
# unique reflections	32,094	223,984	110,156	139,549	139,814
Multiplicity	19.8 (7.4)	6.5 (5.2)	7.4 (7.1)	13.5 (10.4)	6.1 (4.3)
CC1/2	0.991 (0.817)	0.998 (0.847)	0.990 (0.652)	0.996 (0.665)	0.999 (0.627)
Anomalous completeness	99.5 (95.6)				
Anomalous multiplicity	10.0 (3.8)				
ΔAnom correlation between half sets	0.233 (inner = 0.795)				
# Se	24				
Wavelength (Å)	0.97919				
Refinement					
Resolution (Å)	45.0–2.46	44.5–1.75	42.9–2.22	50.0–1.59	50.0–1.59
No. Reflections	30,444	210,888	104,573	132,605	132,830
Rwork %	22.5	17.4	24.9	15.4	15.0
Rfree %	27.6	20.4	27.7	17.4	17.4
# atoms (total)	7,462	15,574	14,376	8,241	8,285
# waters	413	1445	409	879	959
# buffer/biuret/inhibitor atoms	0	0	49	112 / 22	56
Mean B value overall ($Å^2$)	17.3	29.1	32.8	16.6	22.1
Mean B value protein ($Å^2$)	17.9	29.3	33.5	16.3	21.8
Mean B value water ($Å^2$)	10.7	36.8	26.3	26.8	32.4
Mean B value buffer/biuret/inhibitor ($Å^2$)	NA	NA	24.1	20.6 / 30.2	31.3
r.m.s.d. bond lengths ($Å^2$)	0.012	0.018	0.010	0.015	0.018
r.m.s.d. bond angles (°)	1.634	1.760	1.433	1.787	1.893
Ramachandran analysis (%) preferred/ outliers	97.8 / 0.1	97.9 / 0.1	98.2 / 0	97.9 / 0	97.9 / 0.1

https://doi.org/10.1371/journal.pone.0192736.t001

were manually rebuilt using Coot [42] and refined with Refmac [43]. The biuret and inhibitor constraints were generated with the eLBOW function in Phenix [44]. All the crystal trials were set up in the SD-2 sitting drop plates which were used for the native protein; details of the crystallization conditions can be found in S3 Table. The structures obtained in this study have been lodged in the Protein Data Bank: SeMet BiuH wild-type (PDB: 6AZO), BiuH Cys175Ser (PDB: 6AZN), BiuH Cys175Ser with biuret (PDB: 6AZQ), BiuH Lys142Ala with *N*-carbamoyl-D,L-aspartic acid (PDB: 6AZS) and BiuH Lys142His (PDB: 5BK6).

Molecular dynamics

A 500 ns molecular dynamics simulation was performed on the wild-type enzyme to investigate its substrate-active site interactions as well as the enzyme's conformational plasticity. Beginning from the crystal structure (with ligand coordinates taken from the C175S mutant

structure) the AmberTools package [45] was used to setup molecular dynamics simulations with the Amber2014SB force field for the protein degrees of freedom [46], the GAFF force field for the ligand degrees of freedom [47], and the OBC2 generalized-Born implicit solvent model [48,49]. Simulations were performed using the OpenMM simulation library [50] with a non-bonded cut-off of 1 nanometer, a salt concentration of 150 mM, and hydrogen bond lengths constrained. Equations of motion were integrated with a Langevin integrator (time step of 2 femtoseconds, a temperature of 300 Kelvin, and a collision frequency of 91/picosecond). After energy minimization, the system was equilibrated for 1 nanosecond, followed by 500 nanoseconds of simulation (collecting the positions of all atoms every 200 picoseconds).

The trajectory was subjected to further analysis *via* conformational clustering using the k-means algorithm with k = 20. These states were then assembled into a Markovian state model with a lag time of 40 nanoseconds using the PyEMMA python library [51]. The Caver algorithm [52,53] was used to identify a number of transiently forming tunnels linking the active sites of each monomer to the surface as well as the central cavity of the tetramer.

An additional simulation was performed using the same protocol to characterize the Cys175Ser mutant (which was observed to bind biuret but was non-catalytic)

Enzyme assays

A glutamate dehydrogenase (GDH) coupled reaction was used to measure ammonia release in the biuret hydrolase (BiuH) dependent reactions. GDH catalyzes the NADH-dependent amination of α-ketoglutarate (S4 Fig). Ammonia production by BiuH was followed using the decrease of absorbance at 340 nm by UV spectrophotometry, which was due to the oxidation of NADH by GDH. 1.25 U of GDH was used per 250 μL reaction, the final concentrations of α-ketoglutarate and NADH were 3.5 mM and 0.2 mM, respectively.

Biuret hydrolase specific activity was obtained by using 22 nM of biuret hydrolase wild type or 0.22 μM of the variants and 5 mU/μL of GDH in presence of 1.2 mM of biuret in 25 mM potassium phosphate buffer pH 8.5, at 28 ˚C. Biuret hydrolase kinetic data were measured for the wild type and all the variants having a residual specific activity above 1% of the wild type enzymes, by using 22 nM of biuret hydrolase enzyme and either 2.9 μM or 0.9 μM of the variants, depending on their performance in presence of various concentrations of biuret ranging from 0–4 mM, using the GDH-coupled assay. All the kinetics constants were calculated using GraphPad Prism (GraphPad Software, San Diego, USA) fitting the rate data to the Michaelis-Menten equation:

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_M + [S]}$$

Inhibition study

Inhibition of the BiuH's activity was measured in the presence of *N*-carbamoyl-D,L-aspartic acid, using 22 nM of biuret hydrolase enzyme. A 100 mM stock solution of *N*-carbamoyl-D,L-aspartic acid was prepared in 50 mM HEPES buffer, pH 8.5. The IC_{50} was determined by measuring the catalytic rate of BiuH against 0.2 mM of biuret at 28 ˚C in 50 mM HEPES buffer, pH 8.5 in presence of increasing amount of inhibitor ranging from 0 to 20 mM (S5 Fig). *N*-carbamoyl-D,L-aspartic acid was shown to not limit GDH activity at 0–20 mM.

Results and discussion

Structure of biuret hydrolase

Although BiuH had been partially characterized previously [54], the structure had not been determined and the molecular detail of its biochemical activity had not been investigated. We obtained a number of X-ray structures of BiuH (Table 1), which allowed a more complete analysis of the enzyme. The native protein was purified (20 mg from 1 litre of culture, concentrated to 10 mg/mL) and crystallized after removal of the His-tag *in situ* using chymotrypsin, in neutral conditions containing medium or high molecular weight polyethylene glycols as precipitants. Proteolysis did not impact the activity of the purified enzyme (S6 Fig).

The protein is a tetramer (consistent with SEC; S7 Fig); each monomer adopts a five-stranded parallel β-sheet which is surrounded by α-helices (Fig 2). Helices α2, α4 and α5 (residues 95–102, 178–186, and 202–213, respectively) make symmetric interactions to another protomer in the tetramer, forming a dimer, as part of the basis of the quaternary structure. Each of these dimer interfaces (A-D, B-C) cover an area of over 1800 $Å^2$ whereas the interface between protomers not in the dimer (A-C, B-D) is significantly smaller (*ca*. 700 $Å^2$). BiuH is similar to several other structures in the PDB ID: 3irv and 3uao (NicF; maleamate amidohydrolase; Uniprot: Q88FY5) [55–57] as examples (rmsd of Cα atoms of 1.3 to 1.6 Å over about 180 residues, sequence identities ranging from 26 to 29% identical). Excess densityassociated with Cys175, Cys114, Cys190 and Cys196 in the SeMet structure suggests that these residues have tendency to be oxidized. Depending on the structure and the chain, the loop containing residues 44–53 is at least somewhat disordered and has a higher B-factor than most of the rest of the protein. The average B factor for the chains where the loop is mobile (in more than one conformation) is about double that of the average B factor for both the whole of the protein chain and double what is seen for the same loop (44–53) in other chains where there is a single conformation.

BiuH active site and catalytic mechanism

BiuH belongs to the same family of proteins as RutB (ureidoacrylate peracid amidohydrolase; Uniprot: P75897) [58,59], PncA (nicotinamidase; Uniprot: P21369) [60–62] and NicF [55–57]. These enzymes are involved in the catabolism of heterocyclic compounds; PncA deaminates nicotinamide as part of the salvage pathway, RutB is required for the aerobic catabolism of pyrimidines and NicF is essential for nicotinic acid catabolism (Fig 2). PncA, NicF and RutB are all amidohydrolases; however, they catalyze slightly different reactions, with NicF and PncA perform analogous reactions producing ammonia and maleic acid (NicF) or nicotinic acid (PncA), while RutB produces aminoacrylate and carbamate. BiuH could potentially catalyze either reaction to produce either ammonia and allophanate or carbamate and urea. However, NMR studies with ^{13}C labelled biuret provide strong evidence for the production of allophanate, rather than urea (Fig 3) [54].

Like RutB and NicF, BiuH contains no metal in the active site [58,59,63]. In this regard it differs from PncA, which contains a zinc that is con-ordinated by two histidine residues, an aspartate residue and two water atoms [64]. The bipyrimidal co-ordination of the active site zinc is completed by the nitrogen heteroatom of the nicotinamide ring, which positions the substrate for hydrolysis by the active site nucleophile (Cys159) [64]. This suggests that biuret hydrolase is more distantly related to PncA that it is to NicF or RutB.

A comparison of the sequences of biuret hydrolase with seven homologs (PDB ID: 3irv, 3kl2, 1nba, 3uao, 2wta, 3hu5, and 3r2j) indicated that nine amino acids are highly conserved in the active site of this enzyme family (Asp36, Gln38, Phe41, Lys142, Phe148, Thr151, Gly169,

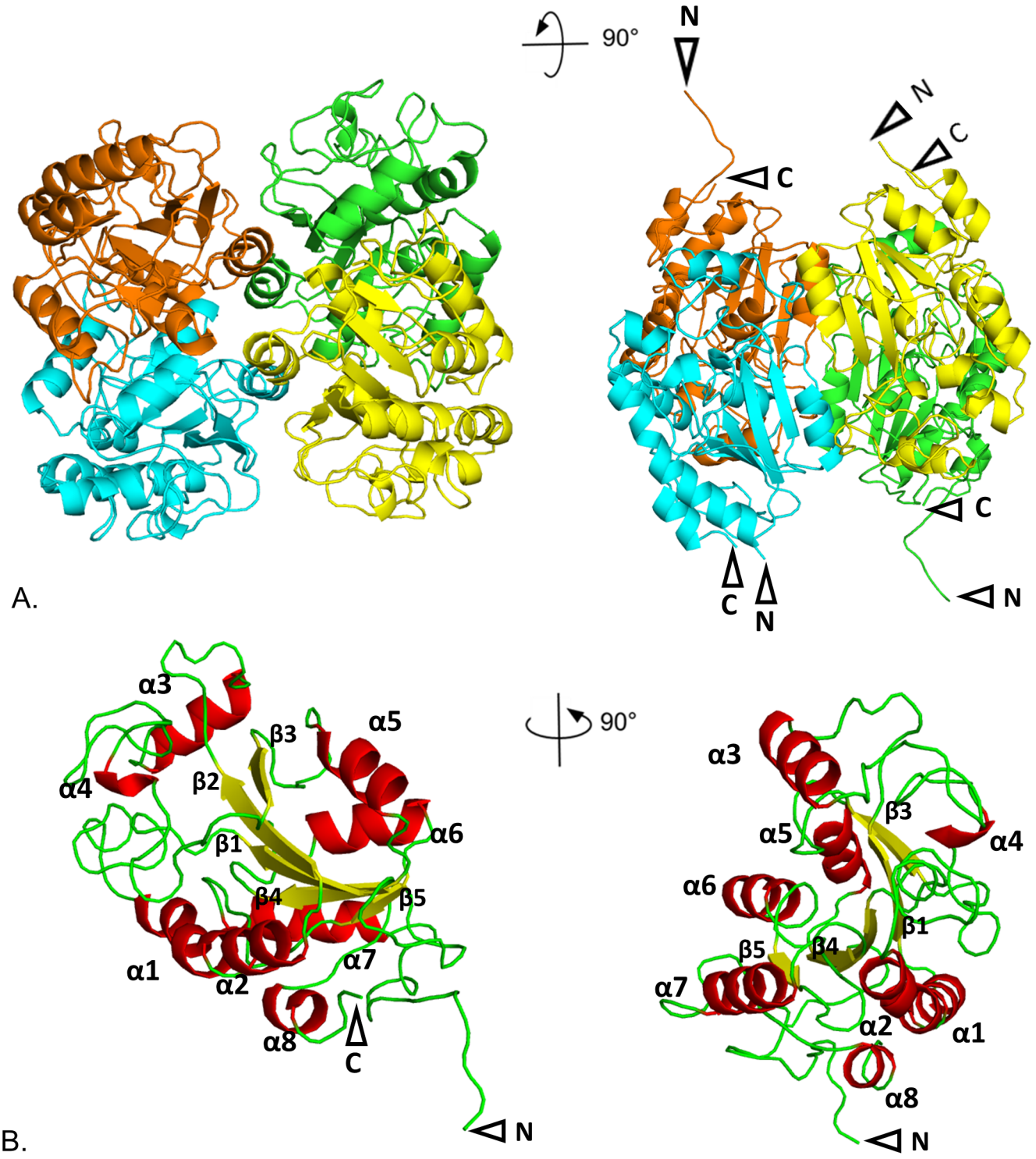


Fig 2. Structure of the biuret amidohydrolase enzyme. Cartoon representation of: A. BiuH tetrameric structure, each color represents a subunit; B. monomeric subunit of BiuH with alpha helices shown in red (numbered α1- α8), beta strands in yellow (numbered β1- β5) and loops in green. In every cases the N and/or C termini were visible, they are indicated by a black arrow. Figures 1, 3 and 6 were generated with PyMol [65].

https://doi.org/10.1371/journal.pone.0192736.g002

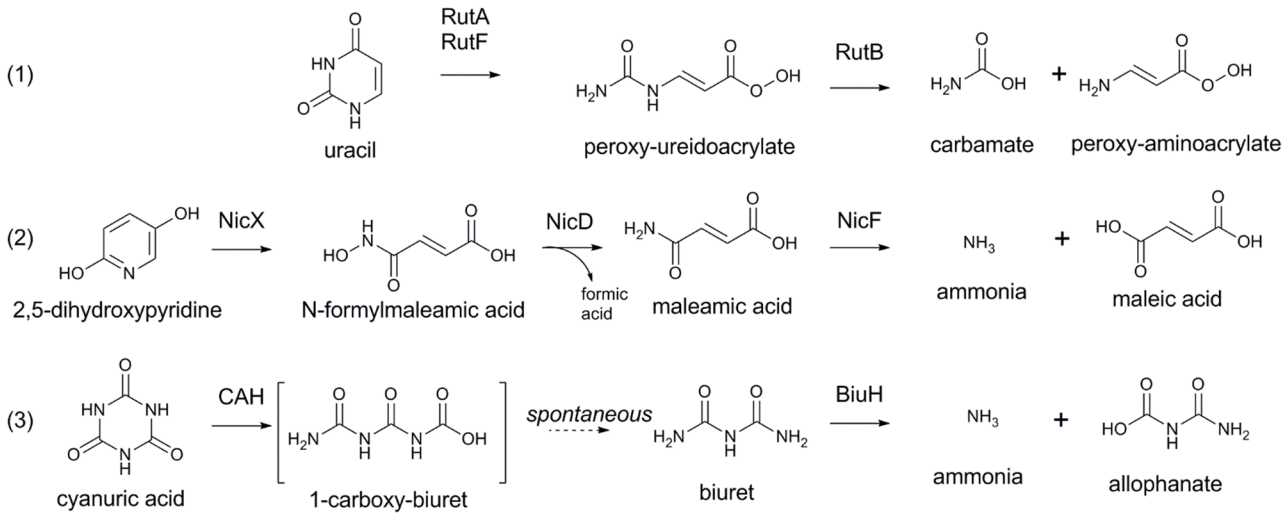


Fig 3. Reactions catalyzed by biuret amidohydrolase and homologs involved in heterocycle catabolism. 1) Ureidoacrylate peracid amidohydrolase (RutB) produces carbamate and peroxy-amino acrylate from peroxy-ureidoacrylate, which is produced by ring opening of uracil by RutA and RutF; 2) maleamate amidohydrolase (NicF) produces ammonia and maleic acid from maleamaic acid, produced by NicX and NicD from 2,5-dihydroxypyridine during nicotinic acid catabolism; and, 3) biuret amidohydrolase (BiuH) produces ammonia and allophanate from biuret during cyanuric acid catabolism.

https://doi.org/10.1371/journal.pone.0192736.g003

Cys175, Thr179). As Cys175 is the only conserved cysteine, it was probable that it was the active site nucleophile; indeed, substitution of Cys175 for isosteric serine abolished catalytic activity (without impacting protein folding; S2 Table). A structure of the inactive Cys175Ser variant with biuret in the active site was obtained (PDB ID: 6AZQ). Given the relatively minor change to the active site relative to that of the wild-type, this structure is a close approximation of the Michaelis complex before the nucleophilic attack on the substrate by active site nucleophile.

In the Cys175Ser variant, biuret has an extensive hydrogen bond network with both side-chains and the protein backbone: Asp36, Lys145, Thr171, Ser175 and Gln215 have sidechain interactions; Ile170 and Thr171 contribute hydrogen bonds from backbone atoms. The active site Gln215 is provided by from the adjacent monomer of the dimer, which protrudes into the active site of its neighbouring monomer and binds the terminal amide that is distal from the nucleophile (Fig 4). The pocket is further constrained by Phe41, Tyr47 and Val174, with side-chain atoms of these residues between 3.2 and 3.9 Å away from the biuret molecule (Fig 4). The center of the Phe41 ring is about 3.0 Å from the N6 nitrogen, which is adjacent to the carbon presumed to be under attack by Cys175 in the wild-type enzyme. In this structure, Ser175, Asp36 and Lys142 are within hydrogen-bonding distance of one another; suggesting that, in addition to Cys175, Asp36 and Lys142 comprise the catalytic triad for BiuH. The amide oxygen closest to the nucleophile, which forms the oxyanion during catalysis, occupies a pocket that is formed by the main-chain nitrogens of Cys175 and Thr171. In NicF the oxyanion hole is formed by main-chain nitrogens of the nucleophilic cysteine (Cys150) and a threonine residue (Thr146) [63], and in PncA it is formed by main-chain nitrogens of a *cis*Ala (*cis*Ala155) and Phe (Phe158) [64]

The catalytic mechanism of BiuH is therefore likely to be identical to that of other cysteine amidohydrolases (Fig 5) in which the active site nucleophile (Cys175) is deprotonated by a general base (Asp36). The function of Lys142 is likely to be the same as that of Lys117 in NicF, which serves to increase the acidity of Asp29 (the equivalent of BiuH Asp36) [55]. The

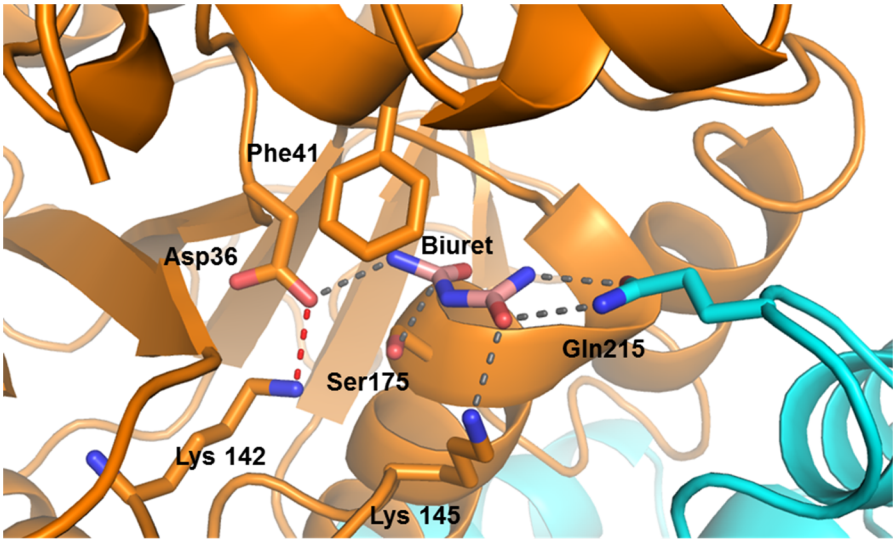


Fig 4. Cys175Ser BiuH variant showing biuret in the active site. BiuH is represented in cartoon style, with the exception of the active site amino acids; biuret is shown in pink, the hydrogen bonds between the residues and biuret are shown in grey, the hydrogen bonds between Asp36 and Lys142 in red; Gln215 is shown in cyan as it belongs to another enzyme subunit. The difference density map (Fo –Fc) of biuret in the active site is shown in S11 Fig.

https://doi.org/10.1371/journal.pone.0192736.g004

nucleophile forms a tetrahedral covalent intermediate with the substrate, with the developing oxyanion stabilised by the main-chain nitrogens of Cys175 and Thr171. Thereafter ammonia is released and an acyl intermediate is formed between the substrate and enzyme. This intermediate is subsequently hydrolyzed to release the product (allophanate) and regenerate the active site for further catalysis (Fig 5). Mutagenesis of Asp36 and Lys142 support this

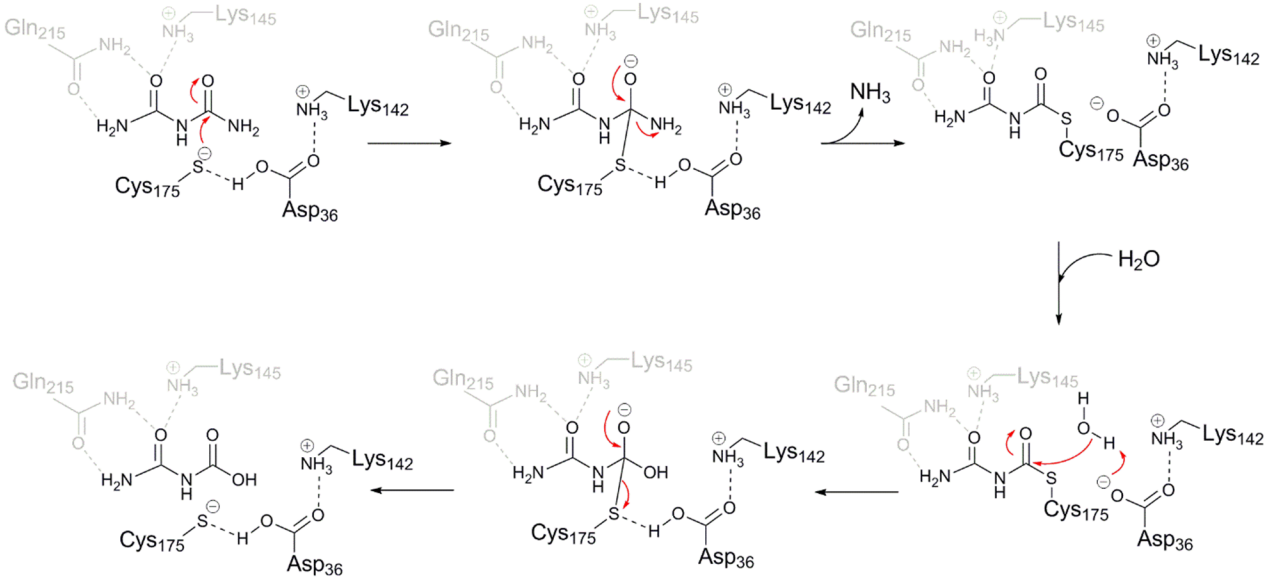


Fig 5. Suggested mechanism of the BiuH. Lys142 stabilizes Asp36 that will act as a general base and deprotonate Cys175, allowing Cys175 to perform a nucleophilic attack on the carbonyl end of biuret. Cys175 then binds to biuret forming a tetrahedral intermediate. Asp36 then acts as a general acid, leading to the collapse of the intermediate and the production of an ammonia and a thioester intermediate. Following the addition of a water molecule, Asp36 deprotonates the molecule of water leading to the hydrolysis of the thioester intermediate, forming a new tetrahedral intermediate. Finally, the enzyme is restored to its original state, releasing the allophanate product.

https://doi.org/10.1371/journal.pone.0192736.g005

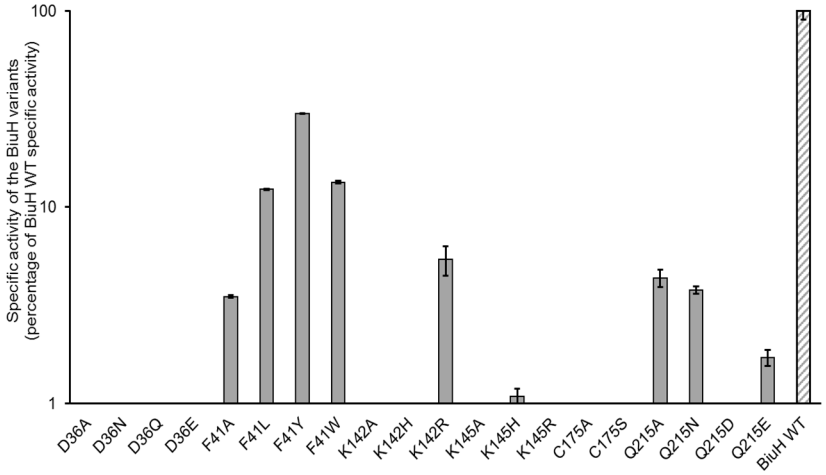


Fig 6. Specific activity of variant enzymes compared to the BiuH wild type enzyme. The specific activity is shown for each variant as percentage of the wild type BiuH specific activity, using 1.2 mM biuret as substrate (n = 3).

https://doi.org/10.1371/journal.pone.0192736.g006

mechanism (Fig 6, S2 Table): Asp36 cannot be substituted for other amino acids and Lys142 can only be replaced by an arginine residue (with a ~25-fold reduction in activity relative to the wild-type enzyme).

BiuH and its variants were treated with *N*-carbamoyl aspartic acid, which is a substrate analog previously demonstrated to inhibit BiuH activity [54]. The structure of the Lys142Ala variant was obtained after treatment with the inhibitor (PDB ID: 6AZS), and an acyl (thiocarbamate) covalent complex was captured between the deaminated inhibitor and Cys175 (Fig 7). This unequivocally demonstrates that Cys175 is the active site nucleophile, and also suggests a role for Lys142 in regenerating the active site after the formation of the covalent intermediate.

Variants were made that disrupted the hydrogen bonding interactions between the active site through hydrogen bonds with Gln215, Asp36, Lys145 and substrate. The purified variants were well folded, as shown by DSF, with T_m values from 41 ˚C (BiuH Asp36Glu) to 68 ˚C (BiuH Lys142His; PDB: 5BK6). The wild-type BiuH had an intermediate melting temperature of 56 ˚C. The Ser175 hydroxyl is less than 3 Å away from one of the two carbons in biuret, looking poised to attack. It should be noted that all heteroatoms (oxygen, nitrogen) in the biuret molecule hydrogen bond to the protein in an exquisite arrangement that makes the protein specific to its substrate. Steady state kinetics of the variants were obtained (Table 2). Substitution of Gln215 with Ala or Glu increased the K_M of BiuH for biuret by more than 25-fold (Table 2), consistent with the role of Gln215 in binding biuret. Interestingly, replacement of Gln215 with Asn slightly decreased the enzyme's K_M for biuret, but reduced its k_{cat} by ~8-fold, presumably by modifying the orientation of the substrate in the active site relative to the nucleophilic thiolate of Cys175. Variants of Lys145 in which Lys was substituted for Ala, Arg or His were effectively inactive (albeit DSF showed them to be folded correctly; Table 2), suggesting that Lys145 may be essential for substrate binding (Lys145 has a 3.1 Å hydrogen bond directly to biuret). Phe41 is found in a mobile loop that appears to occlude the active site when not occupied by the substrate. Replacement of Phe41 with Ala, Leu, Tyr or Trp increased K_M and reduced k_{cat}, suggesting a role in substrate binding and positioning, or in product egress.

Molecular dynamics

The binding of biuret to the wild type enzyme (PDB ID: 6AZO) was further investigated using molecular dynamics. A 500 ns trajectory indicated that biuret was held relatively tightly in the

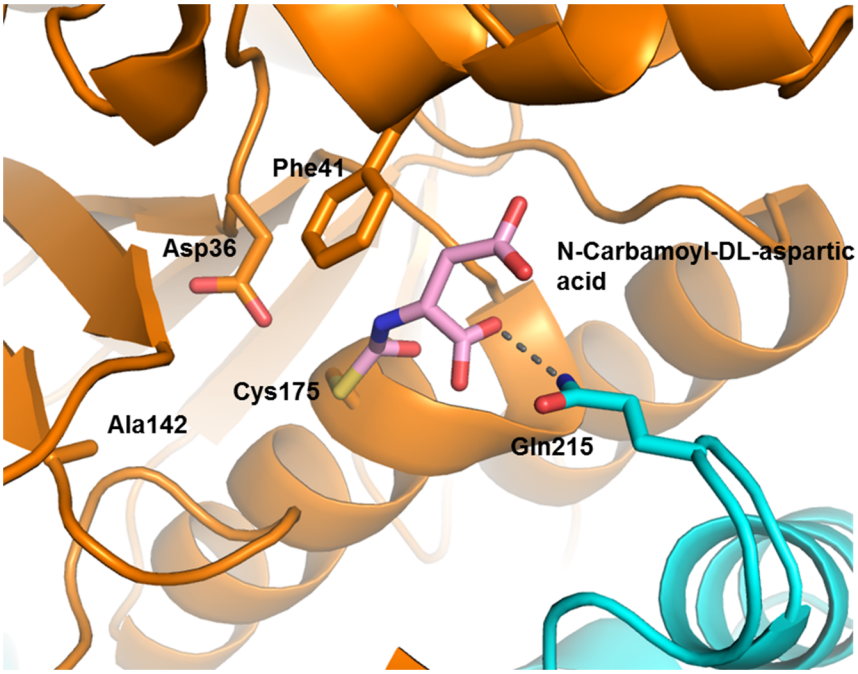


Fig 7. The active site of the Lys142Ala variant of BiuH showing Cys175 bound to the inhibitor *N*-carbamoyl-D,L-aspartic acid. BiuH is represented in cartoon style, with the exception of the active site amino acids; *N*-carbamoyl-D,L-aspartic acid is shown in pink, the hydrogen bond between the Gln215 and *N*-carbamoyl-D,L-aspartic acid are shown in grey, the hydrogen bonds between Asp36 and Ala142 in red; Gln215 is shown in cyan as it belongs to another enzyme subunit. The inhibition profile of BiuH with *N*-carbamoyl-D,L-aspartic acid is shown in S5 Fig and the difference density map (Fo −Fc) of the covalently bound inhibitor is shown in S11 Fig.

https://doi.org/10.1371/journal.pone.0192736.g007

active site moving only slightly from the crystallographic conformation. Additionally, the catalytic residues Asp36 and Cys175 were rigid with Lys142 being somewhat more mobile. Examination of the conformations of the other active site residues revealed that a number of residues distal from the catalytic residues had either relatively high RMSD from the crystallographic conformation, or were shown to adopt two or three distinct rotomeric states (S8 and S9 Figs). From the Markovian state model there are four highly populated states,three more states with moderately low populations, and the remaining states had very low stationary probabilities (S9 Fig). Using these states as a guide, three putative tunnels were identified that extended from the surface of the protein into the active site, and were gated by the residues distal from the catalytic residues (Fig 8). The most persistent tunnel was gated by interactions with Val218 and

Table 2. Steady state kinetic parameters of BiuH and its variants. (n = 5).

	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.M^{-1}$)
BiuH WT	79 ± 7	11.90 ± 0.300	149742
Phe41Ala	505 ± 23	0.14 ± 0.002	277
Phe41Leu	333 ± 15	0.07 ± 0.001	210
Phe41Tyr	749 ± 67	0.81 ± 0.026	1081
Phe41Trp	127 ± 8	0.26 ± 0.003	2047
Gln215Ala	2073 ± 100	0.74 ± 0.016	357
Lys142Arg	139 ± 8	0.07 ± 0.001	504
Gln215Asn	48 ± 2	0.19 ± 0.002	3958
Gln215Glu	2528 ± 197	0.29 ± 0.012	115

https://doi.org/10.1371/journal.pone.0192736.t002

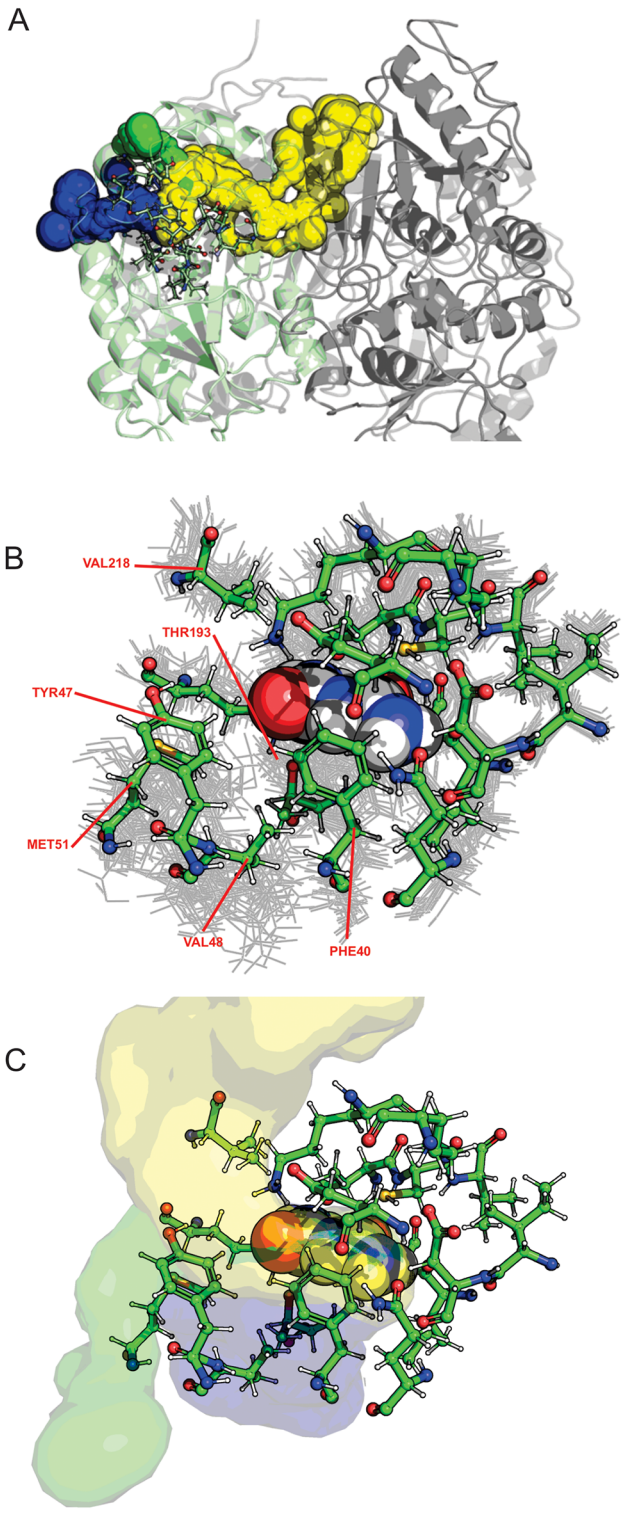


Fig 8. Solvent accessible channels in BiuH. Three solvent accessible channels emerging from the active site are shown in blue, green, and yellow. A) The blue and green tunnels quickly reach the surface of the monomer, while the yellow tunnel extends into the subunit interface and central cavity of the tetramer. The amino acids constituting the active site are shown as sticks, and the subunit that contained biuret is shown in pale green. B) The active site residues are shown as green sticks (for a representative structure of cluster 0) and as grey lines (for representative structures of other 19 clusters) and biuret is shown as Van der Waals spheres. The residues that are responsible for gating these tunnels are labeled in red C) The three dominant tunnels are shown on the right (green, yellow, and blue). Figure generated with Caver [52].

https://doi.org/10.1371/journal.pone.0192736.g008

Gln215 in the adjacent subunit while the smaller tunnels were gated by movement of Phe41, Tyr47, Val48, and Met51. Although further simulation is needed to verify this, it would seem that these secondary tunnels might serve to allow egress of ammonia and admission of water to enable the hydrolysis of the covalent intermediate.

Comparing simulations of the Cys175Ser mutant and the wild type revealed that Ser175 adopted a number of conformations that placed the nucleophilic oxygen an Ångstrom further from the appropriate atom in the substrate when compared to the wild-type (S10 Fig). This was exacerbated by the absence of productive hydrogen bonds with Glu84 and presence of non-productive hydrogen bonds with Thr171. Combined with the relatively higher pK_a of serine, these results help explain why the mutant tightly binds biuret, but does not catalyze its hydrolysis.

Conclusion

Herein, we have described and characterized the structure and catalytic mechanism of BiuH, a cysteine hydrolase that hydrolytically deaminates biuret during cyanuric acid catabolism by *R. leguminosarum* bv. *viciae* 3841. BiuH is structurally and mechanistically different from the 'canonical' AtzE from *Pseudomonas* sp. strain ADP, which fulfils the same physiological role; i.e., it liberates a nitrogen atom from the triazine ring and allows further degradation to make all three nitrogen atoms bioavailable. Curiously, there is no equivalent of *atzF* in close proximity to the *atzD* and *BiuH* genes in *R. leguminosarum* bv. *viciae* 3841. Further study will be required to determine the metabolic fate of the allophanate produced by BiuH.

BiuH has an analogous function to enzymes in both pyridine and pyrimidine catabolism with which it shares sequence, structure and mechanistic similarity (i.e., NicF and RutB). This suggests that BiuH may have evolved from an ancestor with ureidoacrylate peracid amidohydrolase or maleamate amidohydrolase activity. Interestingly, although the gene cluster containing the cyanuric acid and biuret hydrolase in *Rhizobium leguminasorum* bv. *viciae* 3841 is the only such cluster characterized to date, similar (as yet uncharacterized) clusters are present in other bacteria: e.g., *Gordonia* sp. KTR9 (GenBank: AFR50980.1 and AFR50981.1), *Rhizobium leguminosarum* strain Vaf10 (GenBank:ANP91478.1 and ANP91508.1), *Rhizobium leguminosarum* bv. *trifolii* WSM1325 (GenBank:ACS61202.1 and ACS61203.1), *Rhizobium tropici* CIAT 899 (GenBank:AGB75492.1 and AGB75491.1) and *Agrobacterium vitis* S4 (GenBank:ACM38740.1 and ACM38741.1). This may suggest that BiuH homologs play a more important role than anticipated in the environmental fate of *s*-triazine compounds.

Supporting information

S1 Fig. Protein sequence of the biuret hydrolase protein from *Rhizobium leguminosarum* bv. *viciae* 3841. EMBL database accession no. AM236084.1, containing the 6xhis-tag and the thrombin cleavage sites added through the cloning (in red).
(PDF)

S2 Fig. Meltdown output for BiuH.
(PDF)

S3 Fig. Melting temperature of BiuH and its variants. The melting temperature (Tm) was measured by differential scanning fluorimetry in ˚C (n = 3–16 depending on the variants).
(PDF)

S4 Fig. Biuret hydrolase (BiuH) activity assays. A glutamate dehydrogenase (GDH) coupled reaction was used to measure ammonium release in the assays of the biuret hydrolase WT and its variants.
(PDF)

S5 Fig. Percentage of BiuH activity in function of the *N*-Carbamoyl-D,L-aspartic acid inhibitor concentration (mM). The activity of BiuH was measured with 0.2 mM biuret and the GDH-coupled assay in the presence of increasing amounts of inhibitor (n = 6); *N*-Carbamoyl-DL-aspartic acid structure is shown on the top right corner.
(PDF)

S6 Fig. Steady state kinetic parameters of BiuH with/without his-tag and comparison to published data.
(PDF)

S7 Fig. Size exclusion chromatographies from BiuH or its variants. Number 8 represents the elution time for a protein having the size of an octamer of BiuH (red), 4 represents the size of a tetramer of BiuH (blue), 2 the elution time for a protein having the size of a dimer of BiuH (black) and 1 the elution time expected for a protein of the size of the BiuH monomer. The size exclusion from the K145A/H/R variants seem to present a slight shift towards the dimeric size when compared to the other variants.
(PDF)

S8 Fig. Root mean square distance for BiuH active site and biuret during molecular dynamics simulation. Top: Root mean square distance from the crystallographic conformation for active site residues and biuret over the course of a 500 ns molecular dynamics simulation. Bottom: Distributions of root mean square distances from the crystallographic conformation for active side residues during a 500 ns molecular dynamics simulation.
(PDF)

S9 Fig. Conformational clustering of active site residues throughout a 500ns molecular dynamics simulation. States were assigned by clustering the Cartesian coordinates of the active site residues and biuret using the k-means algorithm (k = 20). Top: Discrete trajectory showing the conformational state of the active site throughout the course of the simulation. Bottom: Network diagram showing Markovian state model (lag time = 40 ns) of the conformational transitions between clustered states. The area of each node is proportional to the equilibrium probability of the state while the thickness of the arrows is proportional to the transition probability.
(PDF)

S10 Fig. Comparison of the biuret–enzyme interactions observed during molecular dynamics simulations for the WT enzyme (blue) and the C175S mutant (orange). The top row of plots shows the time trace of the number of hydrogen bonds between the active site residues and biuret, a histogram of these hydrogen bonds, and a bar chart highlighting differences in hydrogen bonding between WT and mutant. On the bottom, the distance between the nucleophile and the electrophilic carbon of biuret is shown as a function of time on the left, and distributions on the right.
(PDF)

S11 Fig. Difference density map (Fo −Fc) shown around biuret and the N-carbamoyl-D,L-aspartic acid inhibitor. A. Active site of the Cys175Ser BiuH variant shown in cartoon, with the active site amino acids and biuret shown in stick and the difference density map (Fo −Fc)

in green mesh. B. Active site of the Lys142Ala variant of BiuH showing Cys175 bound to the inhibitor N-carbamoyl-D,L-aspartic acid and the difference density map (Fo –Fc), with the active site amino acids and the inhibitor shown in stick and the difference density map (Fo –Fc) in green mesh.
(PDF)

S1 Table. Mutagenic primers used for introducing point mutation in BiuH's sequence by overlapping PCR. In bold are the base pairs causing the mutation.
(PDF)

S2 Table. Specific activity of BiuH and its variants. The specific activity was measured in presence of 1.2 mM of biuret (n = 3) in μmoles.sec-1.mg enzyme-1, Tm: melting temperature measured by differential scanning fluorimetry in ˚C (n = 3–16 depending on the variants).
(PDF)

S3 Table. Crystallisation conditions.
(PDF)

Acknowledgments

We thank the beamline scientists of the MX beamlines of the Australian Synchrotron for help during data collection, and acknowledge the contribution of SIEF funds to enable access to the synchrotron. Crystallization experiments were performed in the CSIRO Collaborative Crystallisation Centre. We would also like to thank Drs Sarah Rottet and Andrew Warden for their constructive comments during the preparation of this manuscript.

Author Contributions

Conceptualization: Matthew Wilding, Carol J. Hartley, Colin Scott.

Formal analysis: Thomas S. Peat, Del Lucent, Janet Newman.

Funding acquisition: Colin Scott.

Investigation: Lygie Esquirol, Nigel G. French, Janet Newman.

Methodology: Carol J. Hartley, Janet Newman.

Supervision: Matthew Wilding, Carol J. Hartley, Colin Scott.

Writing – original draft: Lygie Esquirol, Thomas S. Peat, Janet Newman, Colin Scott.

Writing – review & editing: Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Del Lucent, Nigel G. French, Carol J. Hartley, Janet Newman, Colin Scott.

References

1. Wackett LP (2009) Questioning our perceptions about evolution of biodegradative enzymes. Current Opinion in Microbiology 12: 244–251. https://doi.org/10.1016/j.mib.2009.05.001 PMID: 19477677
2. Udikovic-Kolic N, Scott C, Martin-Laurent F (2012) Evolution of atrazine-degrading capabilities in the environment. Applied Microbiology and Biotechnology 96: 1175–1189. https://doi.org/10.1007/s00253-012-4495-0 PMID: 23076592
3. Wackett LP (2004) Evolution of new enzymes and pathways: Soil microbes adapt to s-triazine herbicides. In: Gan JJ, Zhu PC, Aust SD, Lemley AT, editors. Pesticide Decontamination and Detoxification. Washington: Amer Chemical Soc. pp. 37–48.
4. Wackett LP (2005) Biodegradation of *s*-triazine herbicides. Abstracts of Papers of the American Chemical Society 230: U2153–U2153.

5. Wackett LP (2001) Evolution and global distribution on triazine-catabolic enzymes. Biochemistry 40: 8636–8636.
6. Russell RJ, Scott C, Jackson CJ, Pandey R, Pandey G, Taylor MC, et al. (2011) The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. Evolutionary Applications 4: 225–248. https://doi.org/10.1111/j.1752-4571.2010.00175.x PMID: 25567970
7. Peat TS, Newman J, Balotra S, Lucent D, Warden AC, Scott C (2015) The structure of the hexameric atrazine chlorohydrolase AtzA. Acta Crystallographica Section D-Biological Crystallography 71: 710–720.
8. Boundy-Mills KL, de Souza ML, Mandelbaum RT, Wackett LP, Sadowsky MJ (1997) The *atzB* gene of *Pseudomonas* sp strain ADP encodes the second enzyme of a novel atrazine degradation pathway. Applied and Environmental Microbiology 63: 916–923. PMID: 9055410
9. Balotra S, Warden AC, Newman J, Briggs LJ, Scott C, Peat TS (2015) X-Ray structure and mutagenesis studies of the N-isopropylammelide isopropylaminohydrolase, AtzC. PLoS One 10.
10. Peat TS, Balotra S, Wilding M, French NG, Briggs LJ, Panjikar S, et al. (2013) Cyanuric acid hydrolase: evolutionary innovation by structural concatenation. Molecular Microbiology 88: 1149–1163. https://doi.org/10.1111/mmi.12249 PMID: 23651355
11. Seffernick JL, Erickson JS, Cameron SM, Cho S, Dodge AG, Richman JE, et al. (2012) Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. Journal of Bacteriology 194: 4579–4588. https://doi.org/10.1128/JB.00791-12 PMID: 22730121
12. Fruchey I, Shapir N, Sadowsky MJ, Wackett LP (2003) On the origins of cyanuric acid hydrolase: Purification, substrates, and prevalence of AtzD from *Pseudomonas* sp strain ADP. Applied and Environmental Microbiology 69: 3653–3657. https://doi.org/10.1128/AEM.69.6.3653-3657.2003 PMID: 12788776
13. Cameron SM, Sadowsky MJ, Wackett LP (2011) Novel biuret hydrolase reveals presence of food toxicant, cyanuric acid. Abstracts of Papers of the American Chemical Society 241: 1.
14. Martinez B, Tomkins J, Wackett LP, Wing R, Sadowsky MJ (2001) Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. Journal of Bacteriology 183: 5684–5697. https://doi.org/10.1128/JB.183.19.5684-5697.2001 PMID: 11544232
15. Balotra S, Newman J, Cowieson NP, French NG, Campbell PM, Briggs LJ, et al. (2015) X-Ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. Applied and Environmental Microbiology 81: 470–480. https://doi.org/10.1128/AEM.02783-14 PMID: 25362066
16. Shapir N, Cheng G, Sadowsky MJ, Wackett LP (2006) Purification and characterization of TrzF: Biuret hydrolysis by allophanate hydrolase supports growth. Applied and Environmental Microbiology 72: 2491–2495. https://doi.org/10.1128/AEM.72.4.2491-2495.2006 PMID: 16597948
17. Cheng G, Shapir N, Sadowsky MJ, Wackett LP (2005) Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism. Applied and Environmental Microbiology 71: 4437–4445. https://doi.org/10.1128/AEM.71.8.4437-4445.2005 PMID: 16085834
18. Shapir N, Sadowsky MJ, Wackett LP (2005) Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp strain ADP. Journal of Bacteriology 187: 3731–3738. https://doi.org/10.1128/JB.187.11.3731-3738.2005 PMID: 15901697
19. Balotra S, Warden AC, Newman J, Briggs LJ, Scott C, Peat TS (2015) X-Ray Structure and Mutagenesis Studies of the N-Isopropylammelide Isopropylaminohydrolase, AtzC. Plos One 10: 15.
20. Peat TS, Balotra S, Wilding M, Hartley CJ, Newman J, Scott C (2017) High-resolution X-ray structures of two functionally distinct members of the cyclic amide hydrolase family of Toblerone fold enzymes. Applied and Environmental Microbiology 83: 13.
21. Shapir N, Pedersen C, Gil O, Strong L, Seffernick J, Sadowsky MJ, et al. (2006) TrzN from *Arthrobacter aurescens* TC1 is a zinc amidohydrolase. Journal of Bacteriology 188: 5859–5864. https://doi.org/10.1128/JB.00517-06 PMID: 16885454
22. Shapir N, Rosendahl C, Johnson G, Andreina M, Sadowsky MJ, Wackett LP (2005) Substrate specificity and colorimetric assay for recombinant TrzN derived from *Arthrobacter aurescens* TC1. Applied and Environmental Microbiology 71: 2214–2220. https://doi.org/10.1128/AEM.71.5.2214-2220.2005 PMID: 15870302
23. Sugrue E, Carr PD, Scott C, Jackson CJ (2016) Active site desolvation and thermostability trade-offs in the evolution of catalytically diverse triazine hydrolases. Biochemistry 55: 6304–6313. https://doi.org/10.1021/acs.biochem.6b00731 PMID: 27768291
24. Jutzi K, Cook AM, Hutter R (1981) The degradative pathway of the *s*-triazine melamine. Experientia 37: 1231–1232.

25. Cook AM, Hutter R (1981) *s*-Triazines as nitrogen sources for bacteria. Journal of Agricultural and Food Chemistry 29: 1135–1143.

26. Seffernick JL, de Souza ML, Sadowsky MJ, Wackett LP (2001) Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. Journal of Bacteriology 183: 2405–2410. https://doi.org/10.1128/JB.183.8.2405-2410.2001 PMID: 11274097

27. Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, Hull KH, et al. (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biology 7.

28. Jackson CJ, Coppin CW, Carr PD, Aleksandrov A, Wilding M, Sugrue E, et al. (2014) 300-Fold increase in production of the Zn^{2+}-dependent dechlorinase TrzN in soluble form *via* apoenzyme stabilization. Applied and Environmental Microbiology 80: 4003–4011. https://doi.org/10.1128/AEM.00916-14 PMID: 24771025

29. Scott C, Jackson CJ, Coppin CW, Mourant RG, Hilton ME, Sutherland TD, et al. (2009) Catalytic improvement and evolution of atrazine chlorohydrolase. Applied and Environmental Microbiology 75: 2184–2191. https://doi.org/10.1128/AEM.02634-08 PMID: 19201959

30. Scott C, Lewis SE, Milla R, Taylor MC, Rodgers AJW, Dumsday G, et al. (2010) A free-enzyme catalyst for the bioremediation of environmental atrazine contamination. Journal of Environmental Management 91: 2075–2078. https://doi.org/10.1016/j.jenvman.2010.05.007 PMID: 20570036

31. Radian A, Aukema KG, Aksan A, Wackett LP (2015) Silica gel for enhanced activity and hypochlorite protection of cyanuric acid hydrolase in recombinant *Escherichia coli*. Mbio 6: 11.

32. Yeom S, Mutlu BR, Aksan A, Wackett LP (2015) Bacterial cyanuric acid hydrolase for water treatment. Applied and Environmental Microbiology 81: 6660–6668. https://doi.org/10.1128/AEM.02175-15 PMID: 26187963

33. Mac Mahon S, Begley TH, Diachenko GW, Stromgren SA (2012) A liquid chromatography-tandem mass spectrometry method for the detection of economically motivated adulteration in protein-containing foods. Journal of Chromatography A 1220: 101–107. https://doi.org/10.1016/j.chroma.2011.11.066 PMID: 22197251

34. Achor DS, Albrigo LG (2005) Biuret toxicity symptoms in citrus leaves mimics cell senescence rather than nutritional deficiency chlorosis. Journal of the American Society for Horticultural Science 130: 667–673.

35. Shen RC (1959) Fertilizer contaminants -rate of biuret formation from urea. Journal of Agricultural and Food Chemistry 7: 762–763.

36. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension by the polymerase chain-reaction. Gene 77: 51–59. PMID: 2744487

37. Doublie S (1997) Preparation of selenomethionyl proteins for phase determination. Macromolecular Crystallography, Pt A 276: 523–530.

38. Seabrook SA, Newman J (2013) High-throughput thermal scanning for protein stability: Making a good technique more robust. Acs Combinatorial Science 15: 387–392. https://doi.org/10.1021/co400013v PMID: 23710551

39. Rosa N, Ristic M, Seabrook SA, Lovell D, Lucent D, Newman J (2015) Meltdown: A tool to help in the interpretation of thermal melt curves acquired by differential scanning fluorimetry. Journal of Biomolecular Screening 20: 898–905. https://doi.org/10.1177/1087057115584059 PMID: 25918038

40. Skubak P, Pannu NS (2013) Automatic protein structure solution from weak X-ray data. Nature Communications 4: 6.

41. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallographica Section D-Biological Crystallography 63: 32–41.

42. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. Acta Crystallographica Section D-Biological Crystallography 60: 2126–2132.

43. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallographica Section D-Biological Crystallography 67: 355–367.

44. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallographica Section D-Biological Crystallography 66: 213–221.

45. D.A. Case DA, Cerutti DS, Cheatham III TE, Darden TA, Duke RE, Giese TJ, et al. (2017) AMBER 2017. San Francisco: University of California.

46. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C (2015) ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. Journal of Chemical Theory and Computation 11: 3696–3713. https://doi.org/10.1021/acs.jctc.5b00255 PMID: 26574453

47. Wang JM, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general AMBER force field. Journal of Computational Chemistry 25: 1157–1174. https://doi.org/10.1002/jcc.20035 PMID: 15116359

48. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins-Structure Function and Bioinformatics 55: 383–394.

49. Nguyen H, Roe DR, Simmerling C (2013) Improved generalized Born solvent model parameters for protein simulations. Journal of Chemical Theory and Computation 9: 2020–2034. https://doi.org/10.1021/ct3010485 PMID: 25788871

50. Eastman P, Friedrichs MS, Chodera JD, Radmer RJ, Bruns CM, Ku JP, et al. (2013) OpenMM 4: A reusable, extensible, hardware independent library for high performance molecular simulation. Journal of Chemical Theory and Computation 9: 461–469. https://doi.org/10.1021/ct300857j PMID: 23316124

51. Scherer MK, Trendelkamp-Schroer B, Paul F, Perez-Hernandez G, Hoffmann M, Plattner N, et al. (2015) PyEMMA 2: A software package for estimation, validation, and analysis of Markov models. Journal of Chemical Theory and Computation 11: 5525–5542. https://doi.org/10.1021/acs.jctc.5b00743 PMID: 26574340

52. Pavelka A, Sebestova E, Kozlikova B, Brezovsky J, Sochor J, Damborsky J (2016) CAVER: Algorithms for analyzing dynamics of tunnels in macromolecules. IEEE-ACM Transactions on Computational Biology and Bioinformatics 13: 505–517. https://doi.org/10.1109/TCBB.2015.2459680 PMID: 27295634

53. Pavelka A, Chovancova E, Damborsky J (2009) HotSpot Wizard: a web server for identification of hot spots in protein engineering. Nucleic Acids Research 37: W376–W383. https://doi.org/10.1093/nar/gkp410 PMID: 19465397

54. Cameron SM, Durchschein K, Richman JE, Sadowsky MJ, Wackett LP (2011) New family of biuret hydrolases involved in *s*-triazine ring metabolism. ACS Catalysis 1: 1075–1082.

55. Kincaid VA, Sullivan ED, Klein RD, Noel JW, Rowlett RS, Snidert MJ (2012) Structure and catalytic mechanism of nicotinate (vitamin B3) degradative enzyme maleamate amidohydrolase from *Bordetella bronchiseptica* RB50. Biochemistry 51: 545–554. https://doi.org/10.1021/bi201347n PMID: 22214383

56. Jimenez JI, Juarez JF, Garcia JL, Diaz E (2011) A finely tuned regulatory circuit of the nicotinic acid degradation pathway in *Pseudomonas putida*. Environmental Microbiology 13: 1718–1732. https://doi.org/10.1111/j.1462-2920.2011.02471.x PMID: 21450002

57. Jimenez JI, Canales A, Jimenez-Barbero J, Ginalski K, Rychlewski L, Garcia JL, et al. (2008) Deciphering the genetic determinants for aerobic nicotinic acid degradation: The *nic* cluster from *Pseudomonas putida* KT2440. Proceedings of the National Academy of Sciences of the United States of America 105: 11329–11334. https://doi.org/10.1073/pnas.0802273105 PMID: 18678916

58. Parales RE, Ingraham JL (2010) The surprising Rut pathway: an unexpected way to derive nitrogen from pyrimidines. Journal of Bacteriology 192: 4086–4088. https://doi.org/10.1128/JB.00573-10 PMID: 20562306

59. Kim KS, Pelton JG, Inwood WB, Andersen U, Kustu S, Wemmer DE (2010) The Rut pathway for pyrimidine degradation: Novel chemistry and toxicity problems. Journal of Bacteriology 192: 4089–4102. https://doi.org/10.1128/JB.00201-10 PMID: 20400551

60. Gazzaniga F, Stebbins R, Chang SZ, McPeek MA, Brenner C (2009) Microbial NAD metabolism: Lessons from comparative genomics. Microbiology and Molecular Biology Reviews 73: 529-+. https://doi.org/10.1128/MMBR.00042-08 PMID: 19721089

61. French JB, Cen Y, Sauve AA, Ealick SE (2010) High-resolution crystal structures of *Streptococcus pneumoniae* nicotinamidase with trapped intermediates provide insights into the catalytic mechanism and inhibition by aldehydes. Biochemistry 49: 8803–8812. https://doi.org/10.1021/bi1012436 PMID: 20853856

62. Frothingham R, MeekerOconnell WA, Talbot EAS, George JW, Kreuzer KN (1996) Identification, cloning, and expression of the *Escherichia coli* pyrazinamidase and nicotinamidase gene, *pncA*. Antimicrobial Agents and Chemotherapy 40: 1426–1431. PMID: 8726014

63. Kincaid VA, Sullivan ED, Klein RD, Noel JW, Rowlett RS, Snidert MJ (2012) Structure and Catalytic Mechanism of Nicotinate (Vitamin B-3) Degradative Enzyme Maleamate Amidohydrolase from Bordetella bronchiseptica RB50. Biochemistry 51: 545–554. https://doi.org/10.1021/bi201347n PMID: 22214383

64. Fyfe PK, Rao VA, Zemla A, Cameron S, Hunter WN (2009) Specificity and mechanism of *Acinetobacter baumanii* nicotinamidase: implications for activation of the front-Line tuberculosis drug pyrazinamide. Angewandte Chemie-International Edition 48: 9176–9179. https://doi.org/10.1002/anie.200903407 PMID: 19859929

65. Schrodinger, LLC (2015) The PyMOL Molecular Graphics System, Version 1.8.

S1 Figure: Protein sequence of the biuret hydrolase protein from *Rhizobium leguminosarum* bv. *viciae* 3841. EMBL database accession no. AM236084.1, containing the 6xhis-tag and the thrombin cleavage sites added through the cloning (in red).

MGSSHHHHHHSSGLVPRGSHMPWMDAMVETNRHFIDADPYPWPYNGALRPDNTALIIIDMQT

DFCGKGGYVDHMGYDLSLVQAPIEPIKRVLAAMRAKGYHIIHTREGHRPDLADLPANKRWRS

QRIGAGIGDPGPCGRILTRGEPGWDIIPELYPIEGETIIDKPGKGSFCATDLELVLNQKRIE

NIILTGITTDVCVSTTMREANDRGYECLLLEDCCGATDYGNHLAAIKMVKMQGGVFGSVSNS

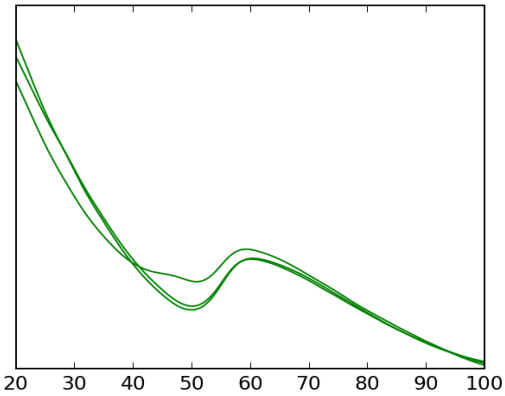
AALVEALP

S1 Table: Mutagenic primers used for introducing point mutation in BiuH's sequence by overlapping PCR. In bold are the base pairs causing the mutation.

Amino acid mutated	Mutagenic primer 5' → 3'
Asp36Ala Fwd	gccctcatcatcatc**gcc**atgcagacggatttc
Asp36Ala Rev	gaaatccgtctgcat**ggc**gatgatgatgagggc
Asp36Asn Fwd	ccgccctcatcatcatc**aac**atgcagacggatttc
Asp36Asn Rev	gaaatccgtctgcatg**ttg**atgatgatgagggcgg
Asp36Gln Fwd	ccgccctcatcatcatc**cag**atgcagacggatttctg
Asp36Gln Rev	cagaaatccgtctgcat**ctg**gatgatgatgagggcgg
Asp36Glu Fwd	gccctcatcatcatc**gag**atgcagacggatttctg
Asp36Glu Rev	cagaaatccgtctgcat**ctc**gatgatgatgagggc
Lys142Ala Fwd	gagacgatcatcgac**gcg**cccggcaagggttc
Lys142Ala Rev	gaacccttgccggg**cgc**gtcgatgatcgtctc
Lys142His Fwd	gagacgatcatcgac**cat**cccggcaagggttcg
Lys142His Rev	cgaacccttgccggg**atg**gtcgatgatcgtctc
Lys142Arg Fwd	ggcgagacgatcatcgac**agg**cccggcaag
Lys142Arg Rev	cttgccggg**cct**gtcgatgatcgtctcgcc
Cys175Ala Fwd	gatcaccaccgatgt**cgc**cgtctcgacgacgatg
Cys175Ala rev	catcgtcgtcgagacg**gcg**acatcggtggtgatc
Cys175Ser Fwd	caccaccgatgtc**agc**gtctcgacgac
Cys175Ser rev	gtcgtcgagac**gct**gacatcggtggtg
Lys145Ala Fwd	tcgacaagcccgg**cgc**gggttcgttctgcg
Lys145Ala Rev	cgcagaacgaaccc**gcg**ccgggcttgtcga
Lys145His Fwd	tcgacaagcccggc**cat**ggttcgttctgcgc
Lys145His Rev	gcgcagaacgaacc**atg**gccgggcttgtcga
Lys145Arg Fwd	atcgacaagcccggc**agg**ggttcgttctg
Lys145Arg Rev	cagaacgaacc**cct**gccgggcttgtcgat
Gln215Ala Fwd	atcaagatggtgaagatg**gcg**ggcggcgtcttcg
Gln215Ala Rev	cgaagacgccgcc**cgc**catcttcaccatcttgat
Gln215Asn Fwd	caagatggtgaagatg**aat**ggcggcgtcttcggc
Gln215Asn Rev	gccgaagacgccgcc**att**catcttcaccatcttg
Gln215Asp Fwd	caagatggtgaagatg**gat**ggcggcgtcttcggc
Gln215Asp Rev	gccgaagacgccgcc**atc**catcttcaccatcttg
Gln215Glu Fwd	tcaagatggtgaagatg**gag**ggcggcgtct
Gln215Glu Rev	agacgccgcc**ctc**catcttcaccatcttga
Phe41Ala Fwd	catcatcgacatgcagacggat**gcc**tgcggcaaggg
Phe41Ala Rev	cccttgccgca**ggc**atccgtctgcatgtcgatgatg
Phe41Leu Fwd	catcgacatgcagacggat**tta**tgcggcaaggg
Phe41Leu Rev	cccttgccgca**taa**atccgtctgcatgtcgatg
Phe41Tyr Fwd	catcatcgacatgcagacggat**tat**tgcggcaagggc
Phe41Tyr Rev	gcccttgccgca**ata**atccgtctgcatgtcgatgatg
Phe41Trp Fwd	tcatcgacatgcagacggat**tgg**tgcggcaaggg
Phe41Trp Rev	cccttgccgca**cca**atccgtctgcatgtcgatga

MELTDOWN v2.0.0 Melt Curve Analysis

cysteinehydrolase-tris-thrombin_Buffer-sc...



Full interpretation of the results requires you to look at the individual melt curves.

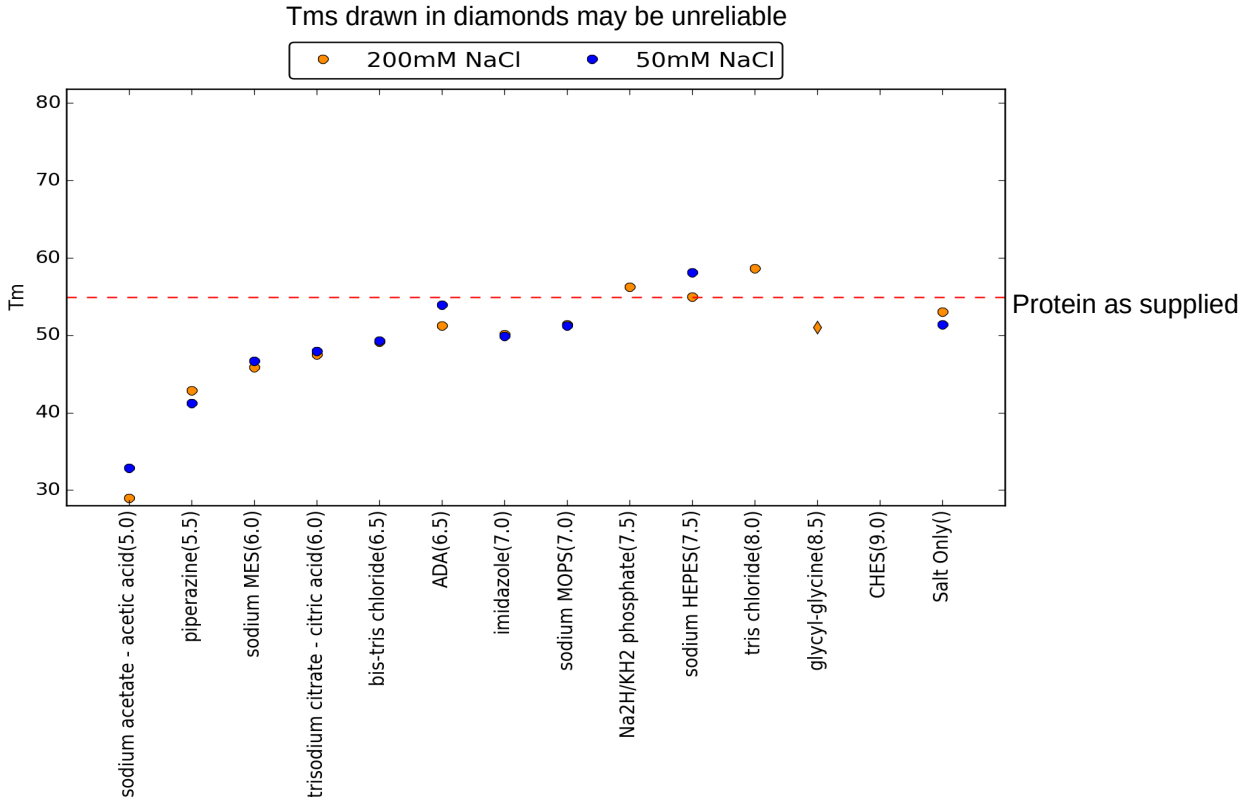
77% of curves were used in Tm estimations

Average estimation of error is **0.3 C**

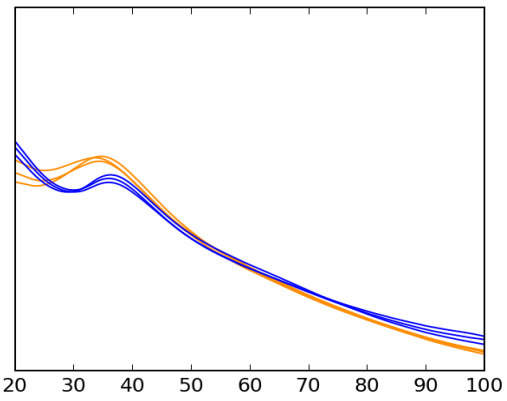
Protein as supplied is **well behaved**

Protein as supplied: Tm = 54.92(+/-0.09)

Lysozyme Control: Passed
No Dye Control: Passed
No Protein Control: Passed

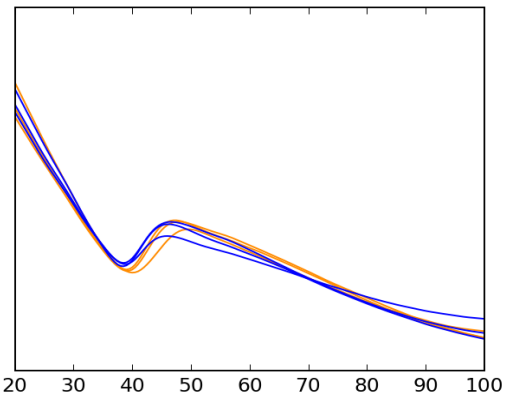


Highest Tm = 58.68 +/- 0.25
(tris chloride / 200mM NaCl)



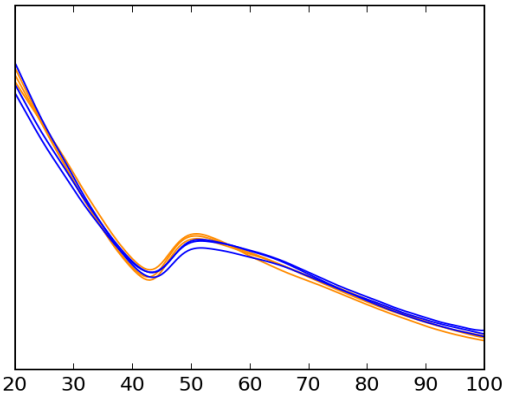
sodium acetate - acetic acid (5.0)

Grouped by	Tm
200mM NaCl	29.0 (+/-0.1)
50mM NaCl	32.83 (+/-0.07)



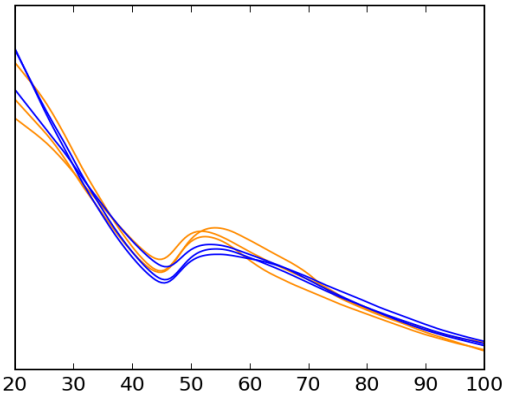
piperazine (5.5)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	42.87 (+/-0.85)	5.08
50mM NaCl	41.27 (+/-0.09)	5.11



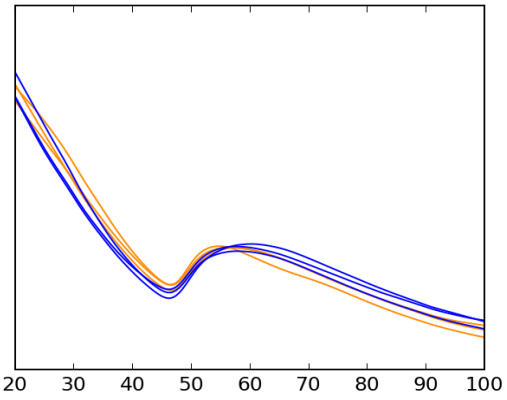
sodium MES (6.0)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	45.85 (+/-0.13)	5.8
50mM NaCl	46.68 (+/-0.07)	5.79



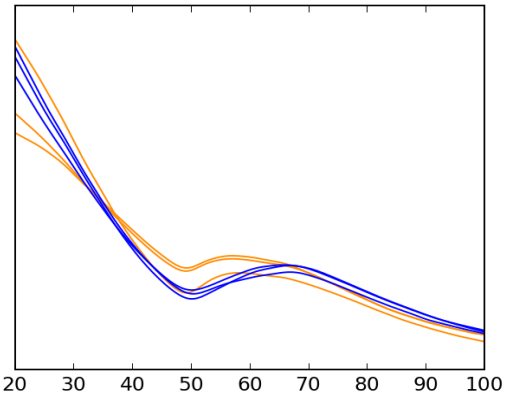
trisodium citrate - citric acid (6.0)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	47.55 (+/-0.21)	6.11
50mM NaCl	47.98 (+/-0.13)	6.11



bis-tris chloride (6.5)

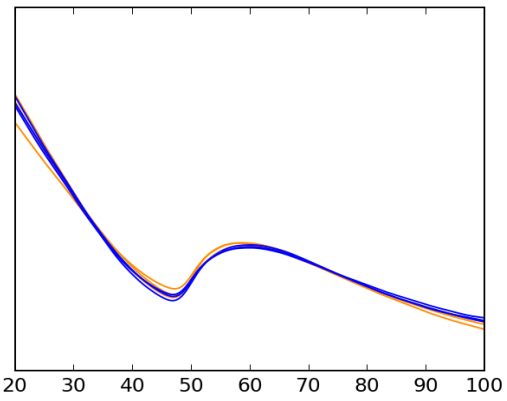
Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	49.17 (+/-0.06)	6.12
50mM NaCl	49.34 (+/-0.06)	6.12



ADA (6.5)

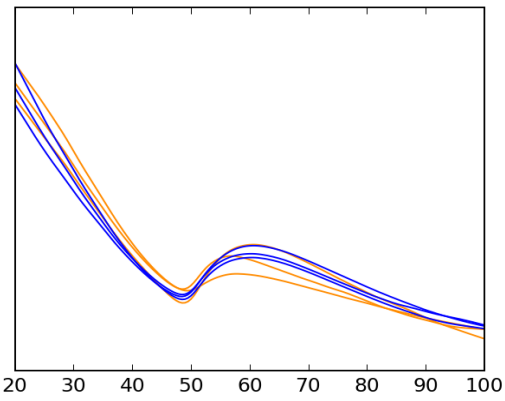
Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	51.27 (+/-0.12)	6.39
50mM NaCl	53.98 (+/-0.56)	6.38

Monotonic, saturated, in the noise, and outlier curves are dotted, and excluded from Tm calculations
Curves drawn with dashed lines have unreliable Tm estimates



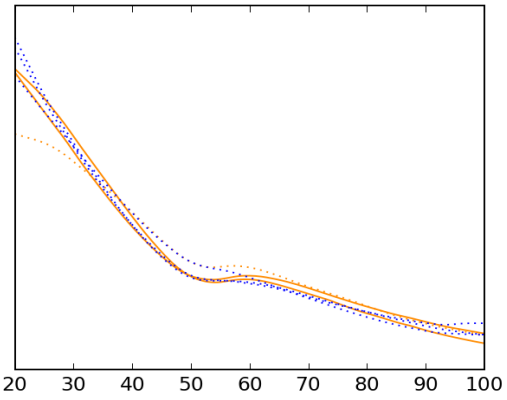
imidazole (7.0)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	50.12 (+/-0.08)	6.49
50mM NaCl	49.94 (+/-0.03)	6.49



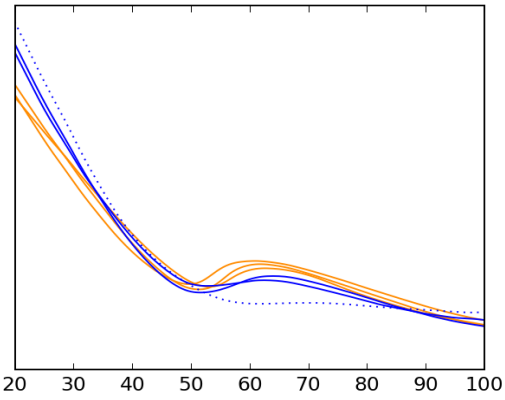
sodium MOPS (7.0)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	51.43 (+/-0.26)	6.74
50mM NaCl	51.25 (+/-0.04)	6.74



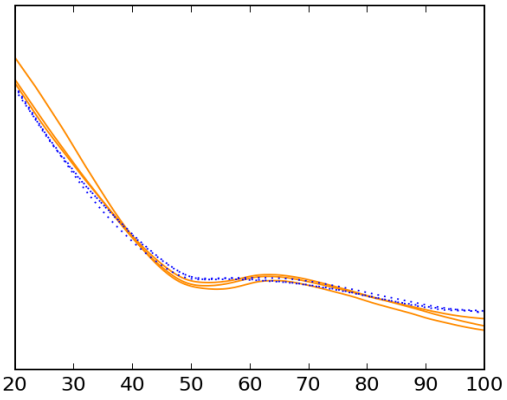
Na2H/KH2 phosphate (7.5)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	56.27 (+/-0.01)	7.7
50mM NaCl	None	



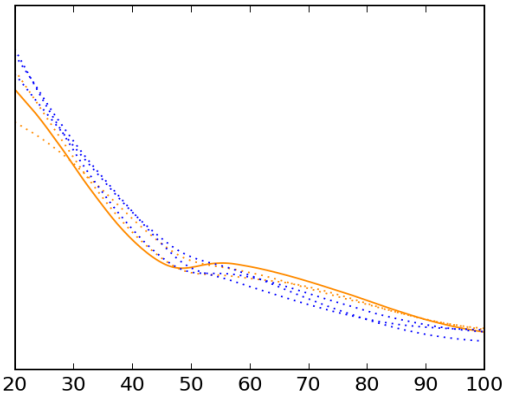
sodium HEPES (7.5)

Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	55.01 (+/-1.24)	7.29
50mM NaCl	58.1 (+/-0.35)	7.27



tris chloride (8.0)

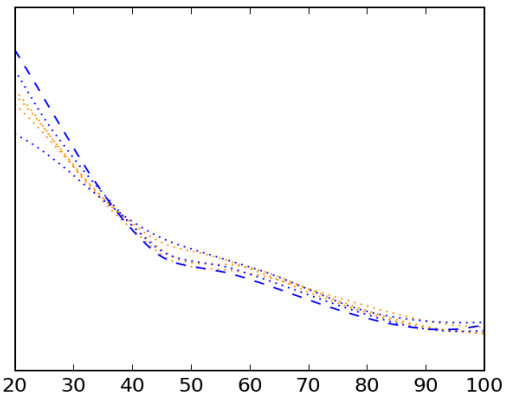
Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	58.68 (+/-0.25)	7.22
50mM NaCl	None	



glycyl-glycine (8.5)

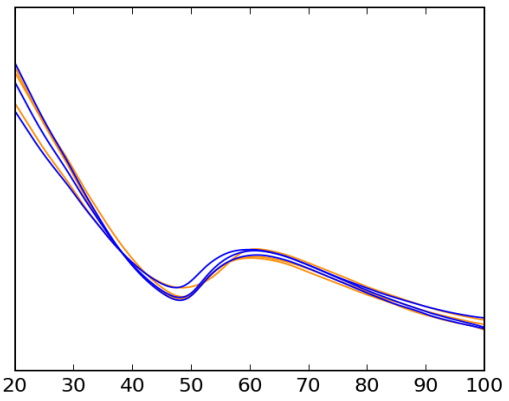
Grouped by	Tm	Adjusted pH at Tm
200mM NaCl	51.04	7.98
50mM NaCl	None	

Monotonic, saturated, in the noise, and outlier curves are dotted, and excluded from Tm calculations
Curves drawn with dashed lines have unreliable Tm estimates



CHES (9.0)

Grouped by	Tm
200mM NaCl	None
50mM NaCl	None



Salt Only ()

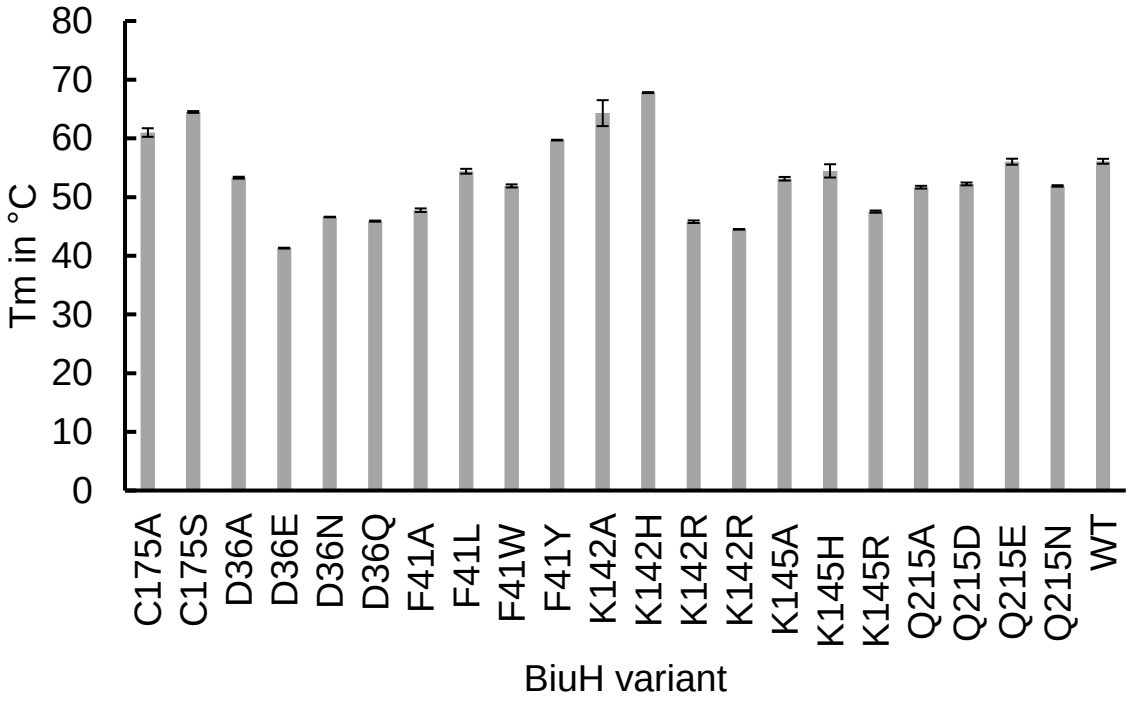
Grouped by	Tm
200mM NaCl	53.06 (+/-1.76)
50mM NaCl	51.4 (+/-0.29)

Monotonic, saturated, in the noise, and outlier curves are dotted, and excluded from Tm calculations
Curves drawn with dashed lines have unreliable Tm estimates

S2 Table: Specific activity of BiuH and its variants. The specific activity was measured in presence of 1.2 mM of biuret (n=3) in µmoles.sec^{-1}.mg enzyme^{-1}, Tm: melting temperature measured by differential scanning fluorimetry in °C (n=3-16 depending on the variants).

	Specific activity	**Tm (°C)**
BiuH WT	51.02 ± 9.670	58.0 ± 0.45
Asp36Ala	0.05 ± 0.012	53.3 ± 0.16
Asp36Asn	0.03 ± 0.006	46.6 ± 0.05
Asp36Gln	0.05 ± 0.009	45.9 ± 0.09
Asp36Glu	0.04 ± 0.087	41.3 ± 0.07
Phe41Ala	1.79 ± 0.033	47.8 ± 0.31
Phe41Leu	6.28 ± 0.057	54.4 ± 0.41
Phe41Tyr	15.33 ± 0.068	59.7 ± 0.07
Phe41Trp	6.84 ± 0.082	51.9 ± 0.24
Lys142Ala	0.04 ± 0.015	64.3 ± 2.21
Lys142His	0.05 ± 0.009	67.8 ± 0.08
Lys142Arg	2.75 ± 0.466	45.8 ± 0.22
Lys145Ala	0.23 ± 0.004	53.1 ± 0.31
Lys145His	0.55 ± 0.074	54.5± 1.13
Lys145Arg	0.34 ± 0.04	47.5 ± 0.19
Cys175Ala	0.04 ± 0.032	61.0 ± 0.73
Cys175Ser	0.05 ± 0.011	64.5 ± 0.15
Gln215Ala	2.22 ± 0.22	51.7 ± 0.24
Gln215Asn	1.92 ± 0.08	51.9 ± 0.13
Gln215Asp	0.22 ± 0.007	52.2 ± 0.24
Gln215Glu	0.87 ± 0.081	56.0 ± 0.51

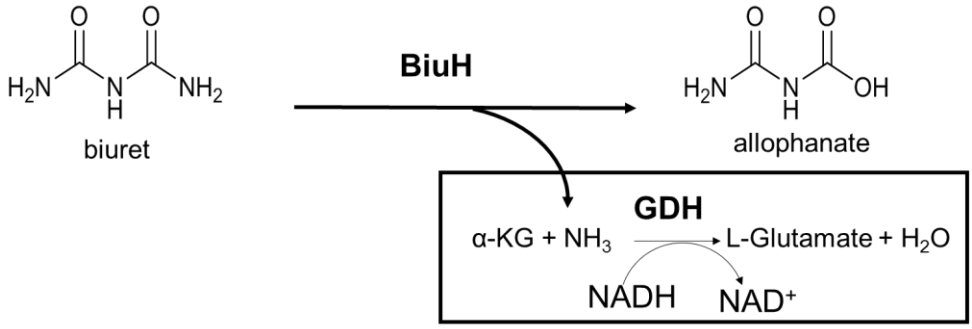
S3 Fig.: Melting temperature of BiuH and its variants. The melting temperature (T_m) was measured by differential scanning fluorimetry in °C (n=3-16 depending on the variants).



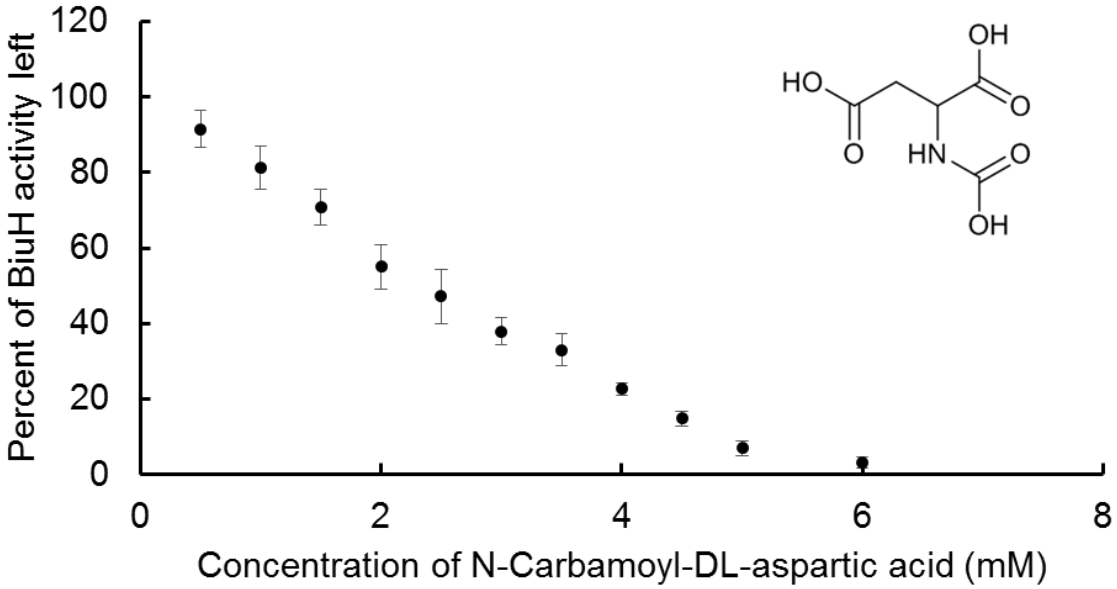
S3 Table: Crystallisation conditions

PDB code	Protein variant	Protein details	Treatment	Additions	Crystallisation condition	Temperature
6AZO	SeMet native	4mg/mL in 50 mM tris pH 8 50 mM NaCl 5 mM DTT	chymotrypsin		0.1 M bis-tris chloride pH 6.14 0.089 M lithium sulfate 20.7 %w/v polyethylene glycol 4000	8 °C
6AZN	C175S	13 mg/mL in 50 mM tris pH 7.5 100 mM NaCl	thrombin	5 mM biuret (from 500 mM stock in DMSO)	0.1 M citrate buffer pH 5.5 0.02 M calcium chloride 12 %w/v polyethylene glycol 8000	20 °C
6AZQ	C175S	13 mg/mL in 50 mM tris pH 7.5 100 mM NaCl	thrombin	Saturated biuret in 50 mM tris pH 7.5, 100 mM NaCl	0.1 M citrate buffer pH 5.46 0.027 M calcium chloride 11.9 %w/v polyethylene glycol 8000	20 °C
6AZS	K142A	7 mg/mL in 50 mM tris pH 7.5 100 mM NaCl	thrombin	1 mM N-carbamoyl DL aspartic acid (from 100 mM stock in DMSO)	0.1 M bis tris chloride pH 6.09 23.6 %w/v polyethylene glycol 3350 0.161 M sodium thiocyanate	8 °C
5BK6	K142H	100 mg/mL in 50 mM tris pH 7.5 100 mM NaCl			0.1 M bis tris chloride pH 5.5 17 %w/v polyethylene glycol 10000 0.1 M sodium acetate	20 C

S4 Fig.: Biuret hydrolase (BiuH) activity assays. A glutamate dehydrogenase (GDH) coupled reaction was used to measure ammonium release in the assays of the biuret hydrolase WT and its variants.



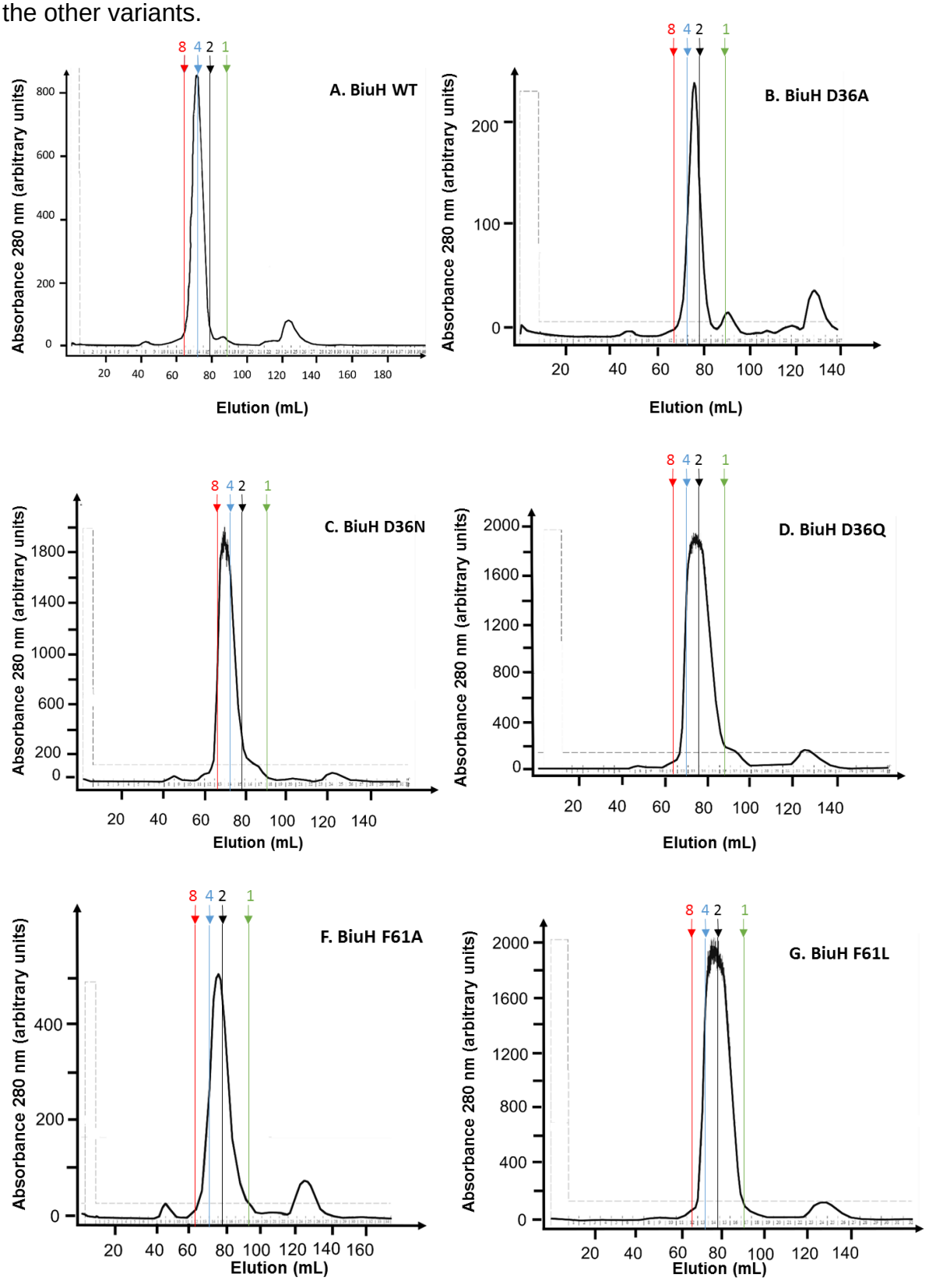
S5 Fig.: Percentage of BiuH activity in function of the *N*-Carbamoyl-D,L-aspartic acid inhibitor concentration (mM). The activity of BiuH was measured with 0.2 mM biuret and the GDH-coupled assay in the presence of increasing amounts of inhibitor (n=6); *N*-Carbamoyl-DL-aspartic acid structure is shown on the top right corner.

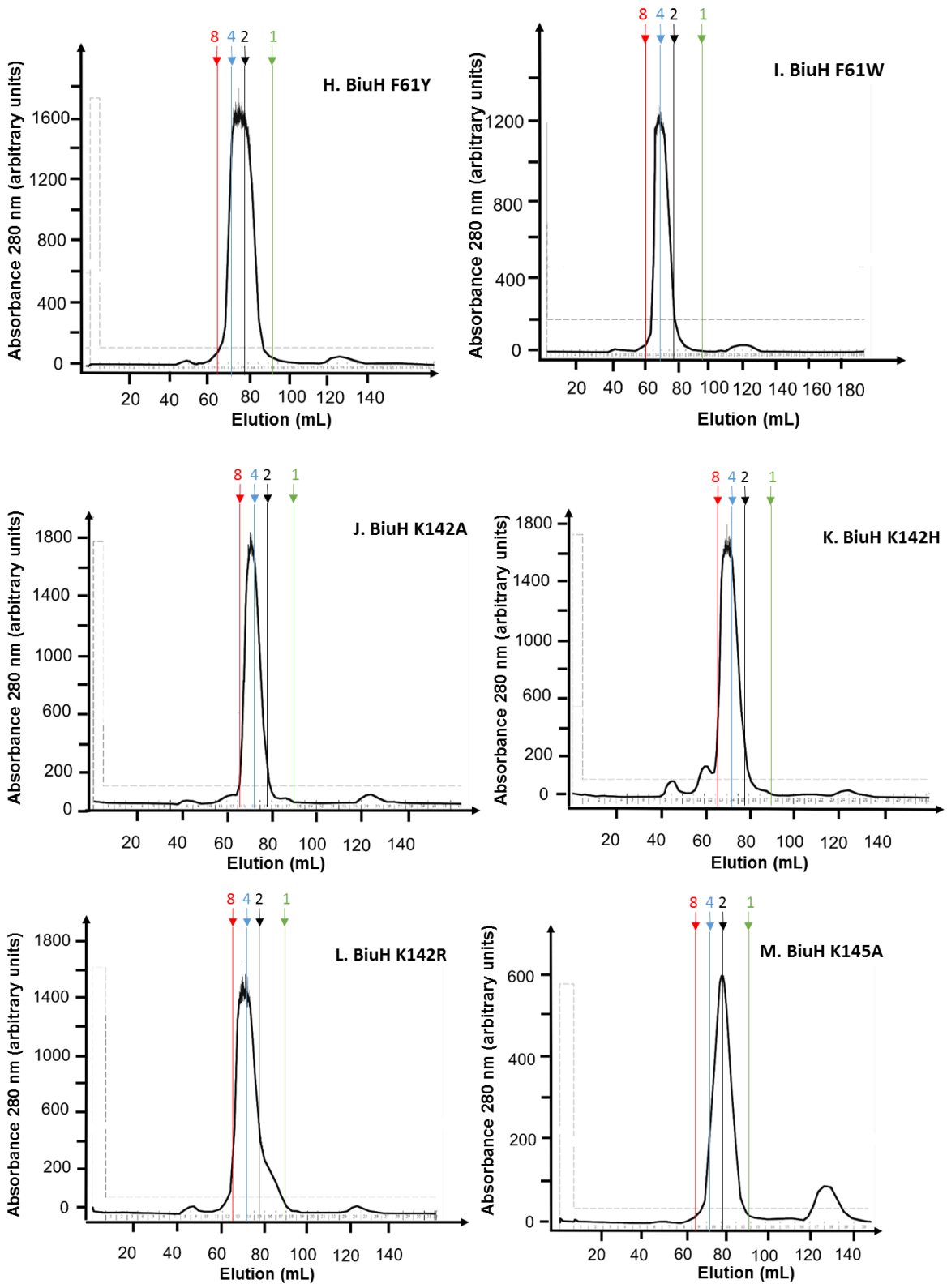


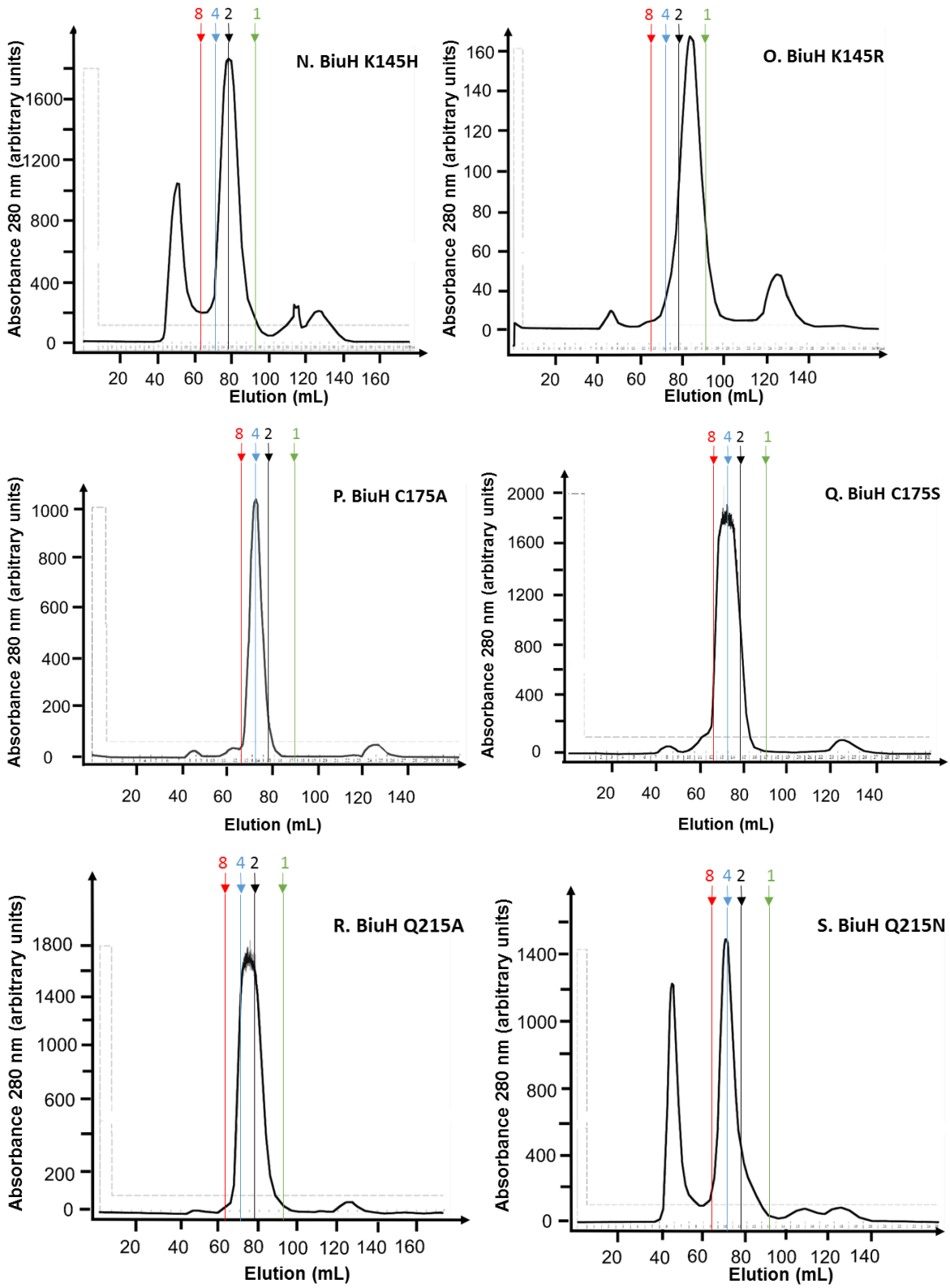
S6 Fig.: Steady state kinetic parameters of BiuH with/without his-tag and comparison to published data.

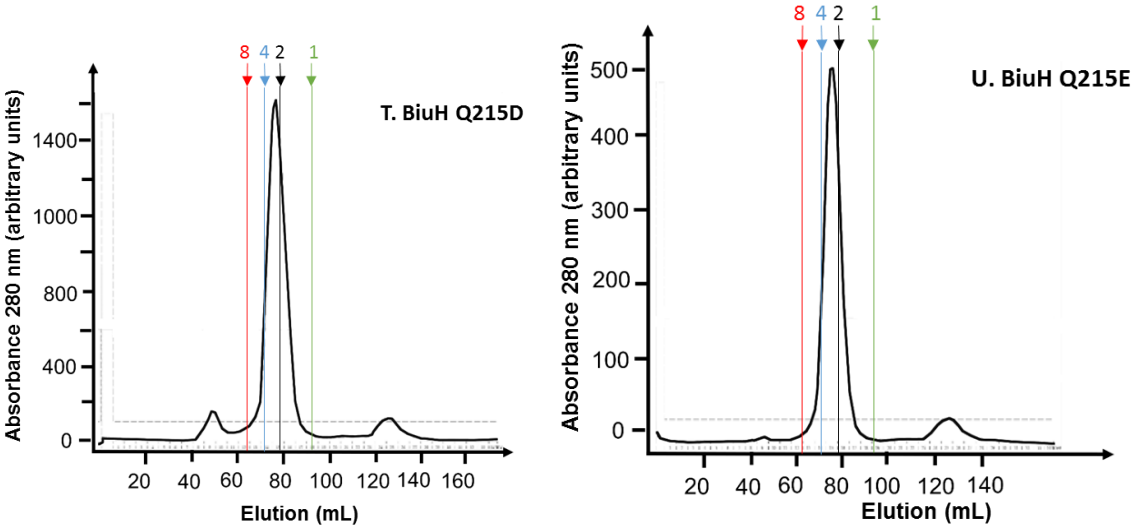
	K_M (µM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.M^{-1}$)
BiuH WT with his-tag	79 ± 7	11.9 ± 0.3	1.5×10^5
BiuH WT without his-tag	74 ± 2	8.8 ± 0.1	1.2×10^5
BiuH WT, as published by Cameron *et al* 2011	23 ± 4	4.0 ± 0.2	1.7×10^5

S7 Fig.: Size Exclusion chromatographies from BiuH or its variants. Number 8 represents the elution time for a protein having the size of an octamer of BiuH (red), 4 represents the size of a tetramer of BiuH (blue), 2 the elution time for a protein having the size of a dimer of BiuH (black) and 1 the elution time expected for a protein of the size of the BiuH monomer. The size exclusion from the K145A/H/R variants seem to present a slight shift towards the dimeric size when compared to the other variants.

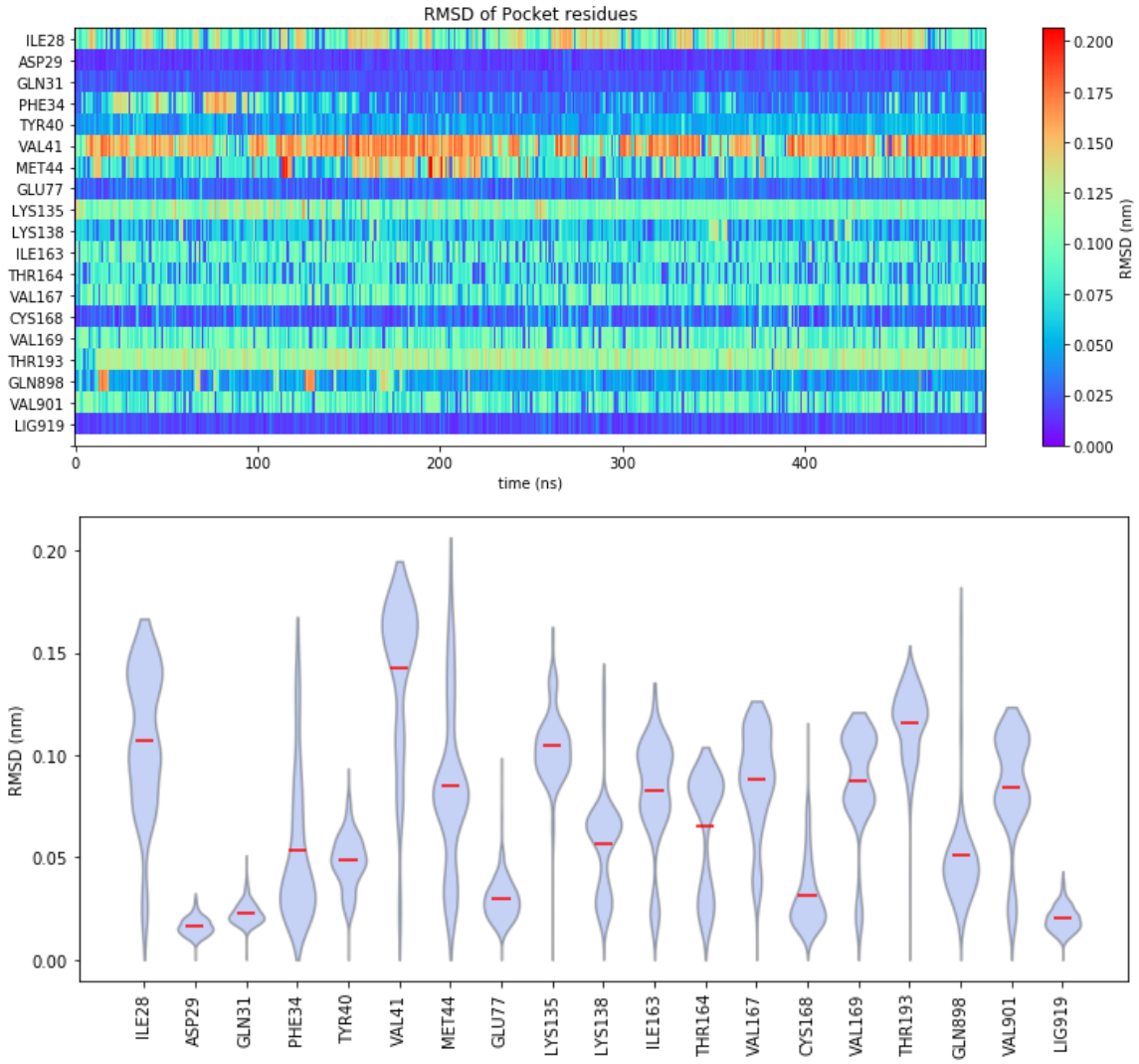



H. BiuH F61Y
I. BiuH F61W
J. BiuH K142A
K. BiuH K142H
L. BiuH K142R
M. BiuH K145A
8 4 2 1
Absorbance 280 nm (arbitrary units)
Elution (mL)

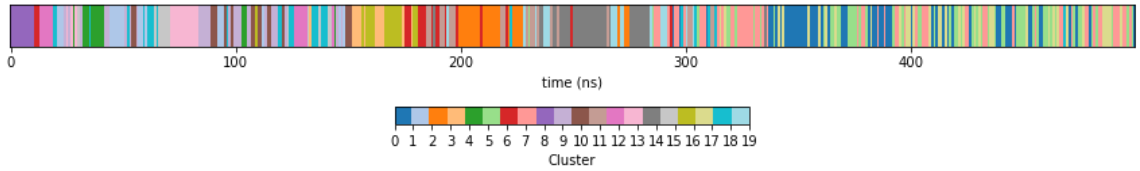

N. BiuH K145H
O. BiuH K145R
P. BiuH C175A
Q. BiuH C175S
R. BiuH Q215A
S. BiuH Q215N
8 4 2 1
Absorbance 280 nm (arbitrary units)
Elution (mL)

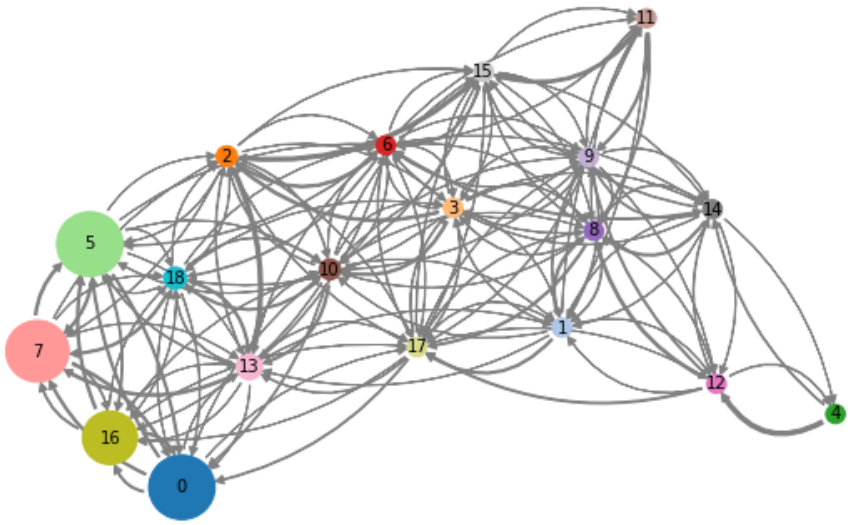

T. BiuH Q215D
U. BiuH Q215E
Absorbance 280 nm (arbitrary units)
Elution (mL)
8
4
2
1
0
400
800
1000
1400
100
200
300
500
20
40
60
80
100
120
140
160

S8 Fig.: Root mean square distance for BiuH active site and biuret during molecular dynamics simulation. Top: Root mean square distance from the crystallographic conformation for active site residues and biuret over the course of a 500 ns molecular dynamics simulation. Bottom: Distributions of root mean square distances from the crystallographic conformation for active side residues during a 500 ns molecular dynamics simulation.

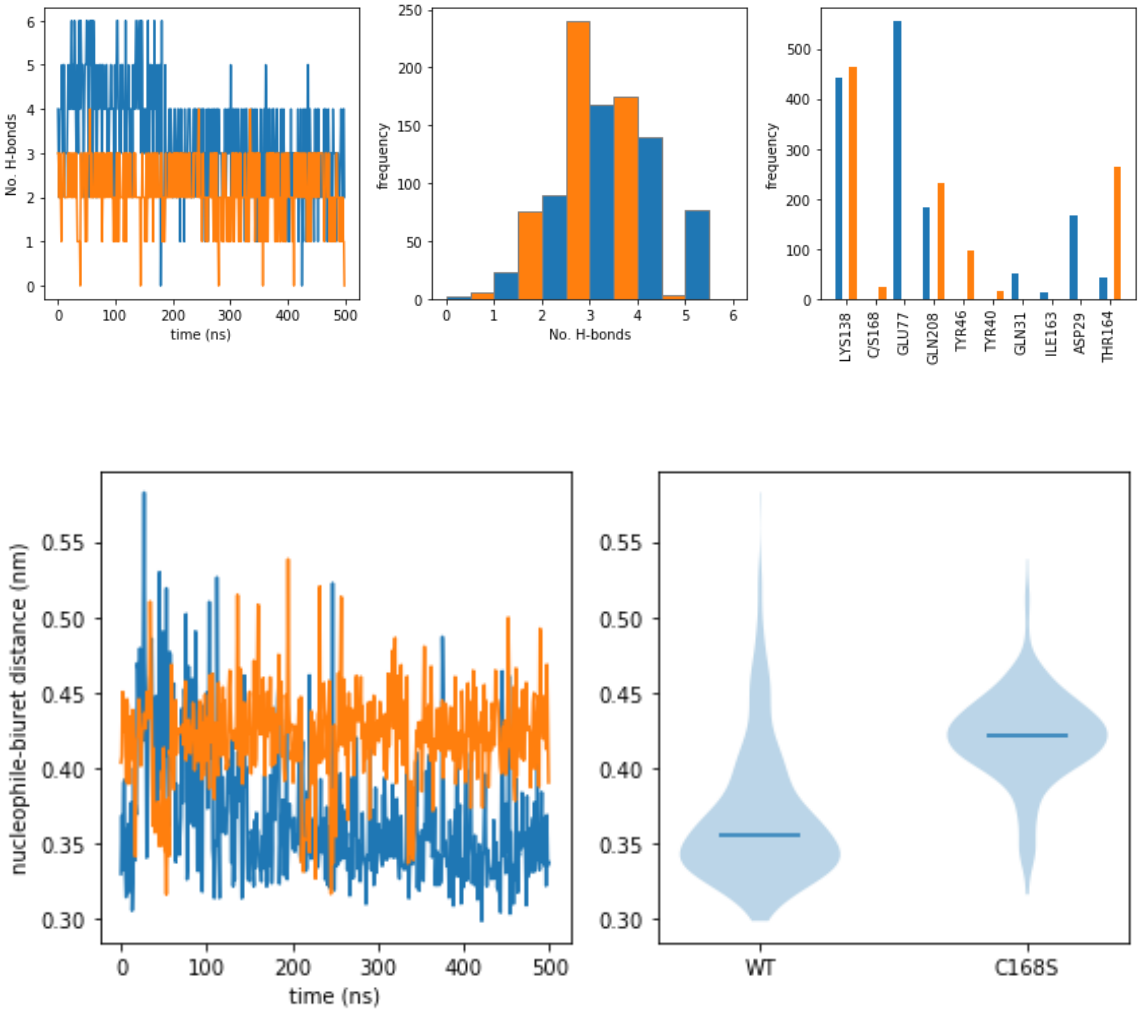


S9 Fig.: Conformational clustering of active site residues throughout a 500ns molecular dynamics simulation. States were assigned by clustering the Cartesian coordinates of the active site residues and biuret using the k-means algorithm (k = 20). Top: Discrete trajectory showing the conformational state of the active site throughout the course of the simulation. Bottom: Network diagram showing Markovian state model (lag time = 40 ns) of the conformational transitions between clustered states. The area of each node is proportional to the equilibrium probability of the state while the thickness of the arrows is proportional to the transition probability.

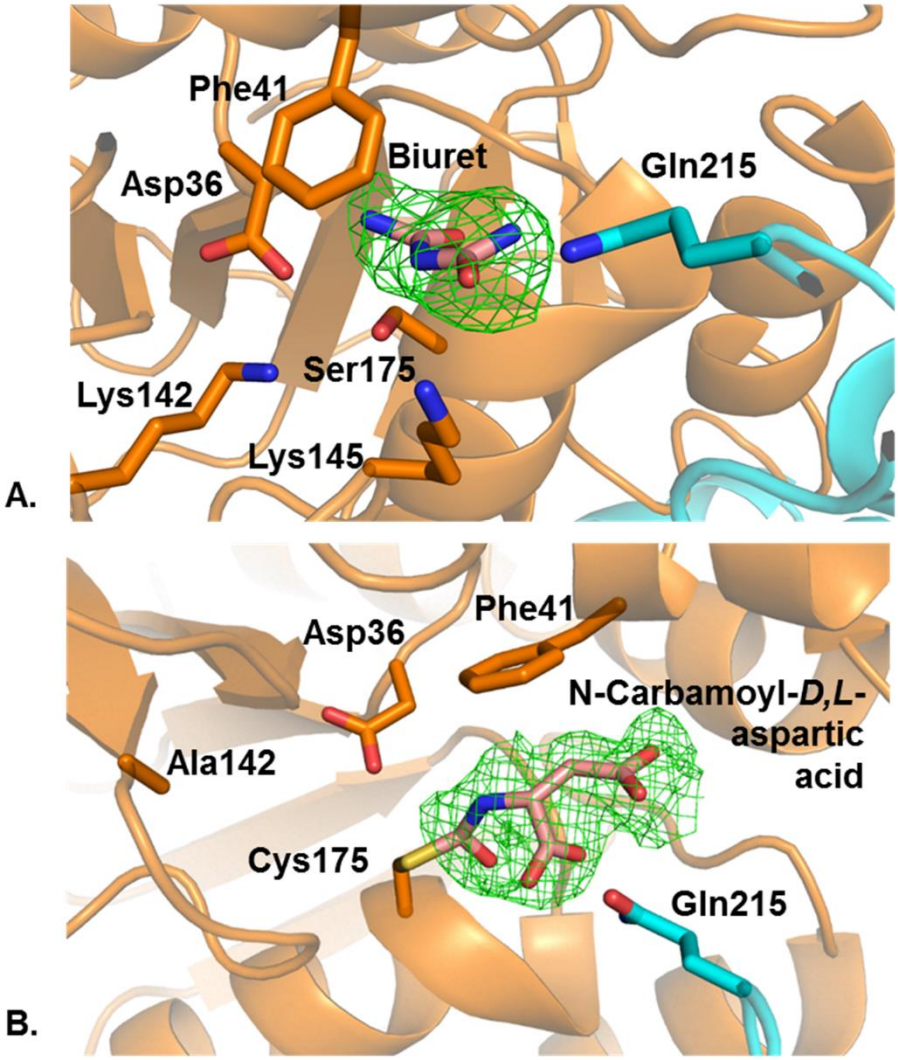




S10 Fig.: Comparison of the biuret–enzyme interactions observed during molecular dynamics simulations for the WT enzyme (blue) and the C175S mutant (orange). The top row of plots shows the time trace of the number of hydrogen bonds between the active site residues and biuret, a histogram of these hydrogen bonds, and a bar chart highlighting differences in hydrogen bonding between WT and mutant. On the bottom, the distance between the nucleophile and the electrophilic carbon of biuret is shown as a function of time on the left, and distributions on the right.



S11 Fig.: Difference density map (Fo –Fc) shown around Biuret and the N-Carbamoyl-D,L-Aspartic acid inhibitor. A. Active site of the Cys175Ser BiuH variant shown in cartoon, with the active site amino acids and biuret shown in stick and the difference density map (Fo –Fc) in green mesh. B. Active site of the Lys142Ala variant of BiuH showing Cys175 bound to the inhibitor N-carbamoyl-D,L-aspartic acid and the difference density map (F o –F c), with the active site amino acids and the inhibitor shown in stick and the difference density map (Fo –Fc) in green mesh.



3 An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme

3.1 Overview

Although the atrazine degradation pathway of *Pseudomonas* sp. strain ADP has been studied for over twenty years and is generally well understood at a molecular level, no one had previously purified and characterised AtzE, the penultimate enzyme of the atrazine degradation pathway, thought to hydrolyse biuret to form allophanate (116, 118). This paper reports the purification of the native AtzE from *Pseudomonas* sp. strain ADP, allowing its structural characterisation. The structure revealed the presence of a small, essential protein (AtzG), with which AtzE forms a heterotetramer. Biochemical characterisation and molecular dynamics experiments revealed that AtzE acts as a 1-carboxybiuret hydrolase, not as a biuret hydrolase as previously thought. Finally, this work suggests that the AtzE might have evolved from the GatCAB transamidosome complex.

This work was selected as an 'Editor's Pick' by the editorial staff at the Journal of Biological Chemistry. Manuscripts selected as Editor's Picks fall within the top 2% of all papers accepted by the Journal of Biological Chemistry.

3.2 Statement of contribution

I developed the purification method of native AtzE from *Pseudomonas* sp. strain ADP and the LC-MS methods with the technical support of Nigel French and Dr Hideki Onagi, respectively. I performed enzyme purification, cloning, enzyme kinetics assays and data analysis. Molecular dynamics simulations were performed by Dr Matthew Wilding. The crystallography and structural determination were performed by Dr Janet Newman and Dr Tom Peat, respectively. Proteomic analysis was performed with Dr Jian-Wei Liu.

Publication status: Published 05/2018.

Lygie Esquirol*, Thomas S. Peat, Matthew Wilding, Jian-Wei Liu, Nigel G. French, Carol J. Hartley, Hideki Onagi, Thomas Nebl, Christopher J. Easton, Janet Newman, and Colin Scott (2018). An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme. J. Biol. Chem. 2018 293: 7880-7891. Published 05/2018. doi:10.1074/jbc.RA118.001996 http://www.jbc.org/content/293/20/7880

Australian National University

Statement of Contribution

This thesis is submitted as a Thesis by Compilation in accordance with https://policies.anu.edu.au/ppl/document/ANUP_003405

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

Title and authors:
An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme. Esquirol L*, Peat TS*, Wilding M, Liu JW, French NG, Hartley CJ, Onagi H, Nebl T, Easton CJ, Newman J, Scott C.

Current status of paper:
Published 03/2018

Contribution to paper:
I developed the purification method of native AtzE from Pseudomonas sp. strain ADP and the LC-MS methods with the technical support of Nigel French and Dr Hideki Onagi, respectively. I performed enzyme purification, cloning, enzyme kinetics assays and data analysis. Molecular dynamics simulations were performed by Dr Matthew Wilding. The crystallography and structural determination were done by Dr Janet Newman and Dr Tom Peat, respectively. Proteomic analysis was performed with Dr Jian-Wei Liu.

Senior author's endorsement:

Lygie's assessment of her contribution to this work is accurate.

Dr. Colin Scott, senior author

LYGIE ESQUIROL
Candidate – Print Name

Signature

25/05/2018
Date

Endorsed

Prof. C. J. Easton
Chair of Supervisory Panel – Print Name

Signature

29 May 2018
Date

G. Otting
Delegated Authority – Print Name

Signature

30 May 2018
Date

3.3 Publication: An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme

Erratum to this publication:

Page 7888, Section Cloning, Line 7:

…we reamplified *atzE* flanked with *Nco*I and *Sac*I (not *Avr*II as stated)…

Page 7888, Section Cloning, Line 19:

…before being subcloned into the second cloning site of the pACYCDuet-1 vector using *Xho*I and *Avr*II (not *Nde*I and *Xho*I as stated)…

Full details are given in the appendix on page 180.

An unexpected vestigial protein complex reveals the evolutionary origins of an *s*-triazine catabolic enzyme

Received for publication, January 24, 2018, and in revised form, February 28, 2018 Published, Papers in Press, March 9, 2018, DOI 10.1074/jbc.RA118.001996

Lygie Esquirol[‡§1], Thomas S. Peat[¶1], Matthew Wilding[§¶], Jian-Wei Liu[‡], Nigel G. French[‡], Carol J. Hartley[‡], Hideki Onagi[§], Thomas Nebl[¶], Christopher J. Easton[§], Janet Newman[¶], and Colin Scott[‡‖2]

From the [‡]Biocatalysis and Synthetic Biology Team and [‖]Synthetic Biology Future Science Platform, CSIRO Land and Water, Canberra, Australian Capital Territory 2601, the [§]Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory 2601, and [¶]CSIRO Biomedical Manufacturing, Parkville, Melbourne, Victoria 3052, Australia

Edited by Karin Musier-Forsyth

Cyanuric acid is a metabolic intermediate of *s*-triazines, such as atrazine (a common herbicide) and melamine (used in resins and plastics). Cyanuric acid is mineralized to ammonia and carbon dioxide by the soil bacterium *Pseudomonas* sp. strain ADP via three hydrolytic enzymes (AtzD, AtzE, and AtzF). Here, we report the purification and biochemical and structural characterization of AtzE. Contrary to previous reports, we found that AtzE is not a biuret amidohydrolase, but instead it catalyzes the hydrolytic deamination of 1-carboxybiuret. X-ray crystal structures of apo AtzE and AtzE bound with the suicide inhibitor phenyl phosphorodiamidate revealed that the AtzE enzyme complex consists of two independent molecules in the asymmetric unit. We also show that AtzE forms an $\alpha 2\beta 2$ heterotetramer with a previously unidentified 68-amino acid-long protein (AtzG) encoded in the cyanuric acid mineralization operon from *Pseudomonas* sp. strain ADP. Moreover, we observed that AtzG is essential for the production of soluble, active AtzE and that this obligate interaction is a vestige of their shared evolutionary origin. We propose that AtzEG was likely recruited into the cyanuric acid–mineralizing pathway from an ancestral glutamine transamidosome that required protein–protein interactions to enforce the exclusion of solvent from the transamidation reaction.

The presence of anthropogenic compounds in the environment provides selection pressures that can lead to the evolution of new metabolic pathways (1). One of the most well-studied examples of an evolutionary response by bacteria to the presence of synthetic xenobiotics is that of the *s*-triazines (2). This family of compounds includes fertilizers (*e.g.* atrazine and ametryn), resins and plastics (*e.g.* melamine), explosives (*e.g.* Royal Detonation Explosive), and disinfectants (*e.g.* cyanuric acid). Interestingly, many of the *s*-triazine catabolic pathways that have evolved share a common structure, in which specific "upper pathways" convert different *s*-triazines to cyanuric acid (1,3,5-triazine-2,4,6-triol) by hydrolysis of the three substituents that decorate the heterocyclic ring (3–5). A common "lower pathway" then mineralizes cyanuric acid to carbon dioxide and ammonia.

The best characterized of the bacterial catabolic pathways that have evolved in response to *s*-triazines in the environment is the atrazine mineralization pathway of *Pseudomonas* sp. strain ADP. This soil bacterium was isolated in the mid-1990s by Wackett and co-workers (6). The atrazine catabolism pathway of *Pseudomonas* sp. strain ADP is composed of six hydrolases encoded by genes on a self-transmissible plasmid (pADP1) (5). The upper pathway (Fig. 1*A*) is composed of three hydrolytic steps: dechlorination, followed by two sequential *N*-alkyl chain hydrolyses. The upper pathway is catalyzed by AtzA (4, 7), AtzB (8), and AtzC (9), respectively; all three enzymes are metalloenzymes of the amidohydrolase superfamily (9, 10).

The lower pathway from *Pseudomonas* sp. strain ADP (Fig. 1*B*) is also composed of three hydrolases: AtzD, AtzE, and AtzF. Unlike the upper pathway, these three enzymes are serine hydrolases; AtzD is an unusual Ser-Lys hydrolase and is the archetype of a recently discovered protein fold (the Toblerone fold (11, 12)), whereas AtzE and AtzF are Ser-*cis*-Ser-Lys hydrolases (1, 13). The lower pathway progresses via the AtzD-mediated ring opening of cyanuric acid to form an unstable product, 1-carboxybiuret, which is known to spontaneously decarboxylate to biuret in an aqueous environment (1). It was therefore thought that the proceeding catabolic steps consisted of the deamination of biuret by AtzE to form allophanate, and allophanate deamination by AtzF to form carbamic acid. Interestingly, these enzymatic products are unstable in an aqueous environment and can undergo spontaneous reactions; allophanate decarboxylates to urea (14), whereas carbamic acid decomposes to ammonia and carbon dioxide in water (13, 15). AtzD and AtzF have been overproduced in heterologous expression systems, which has allowed them to be studied biochemically (14–17) and structurally (11–13, 18). However, expression and purification of AtzE in a heterologous system have been challenging (1), and structural and biochemical data had not been obtained previously.

The lower pathway has a degree of plasticity in its composition. AtzD can be replaced by homologs, such as TrzD from *Pseudomonas* sp. strain NRRLB-12227 (56% identity to AtzD

The authors declare that they have no conflicts of interest with the contents of this article.
This article was selected as one of our Editors' Picks.
This article contains Figs. S1–S3 and Tables S1–S3.
The atomic coordinates and structure factors (codes 6C62 and 6C6G) have been deposited in the Protein Data Bank (http://wwpdb.org/).
[1] Both authors contributed equally to this work.
[2] To whom correspondence should be addressed: CSIRO Land and Water, Foundation Bldg., Black Mountain Science and Innovation Park, Canberra, ACT 2601, Australia. E-mail: colin.scott@csiro.au.

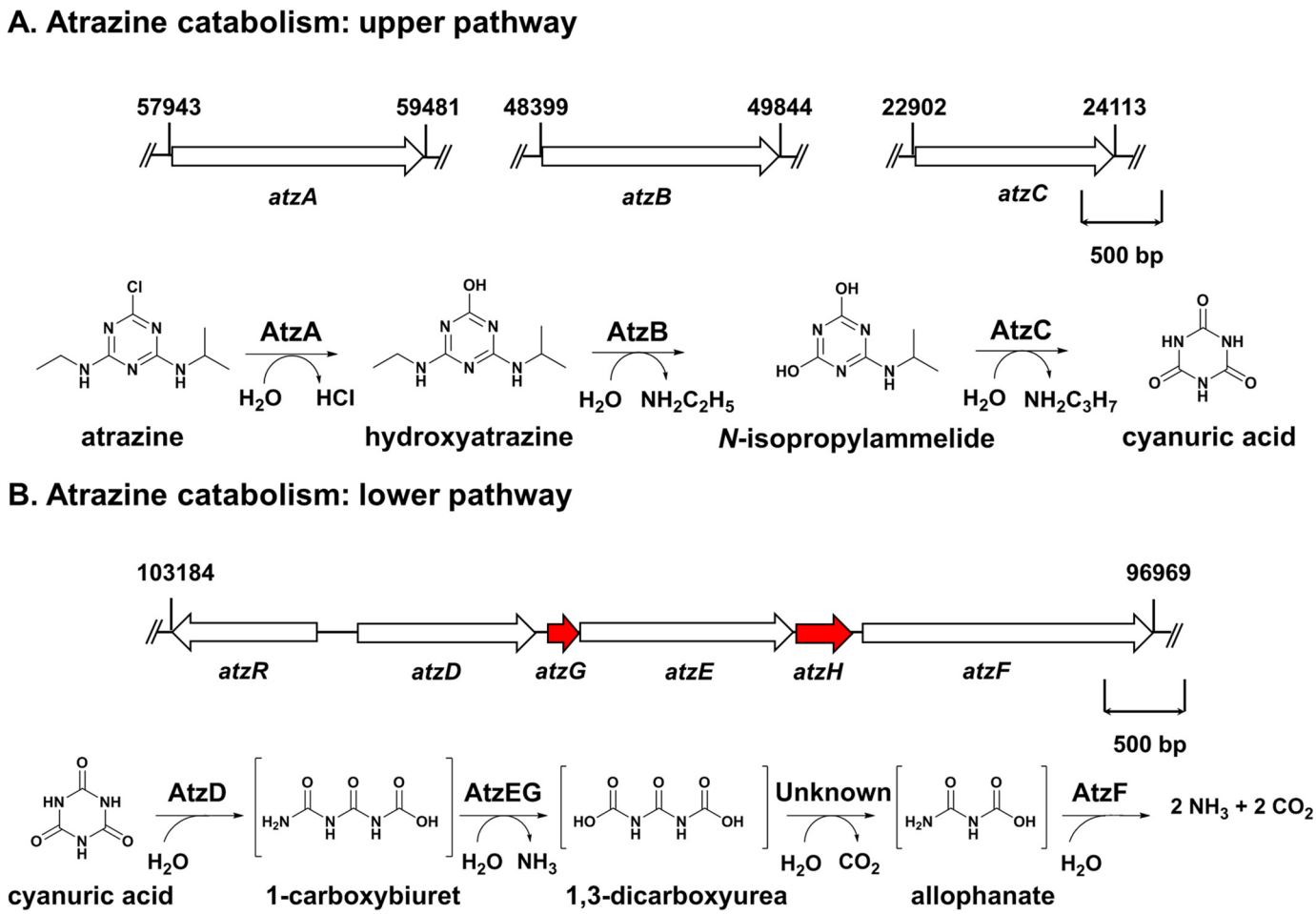


Figure 1. Catabolism of the *s*-triazine herbicide atrazine by *Pseudomonas* sp. strain ADP. The atrazine catabolic pathway is composed of an upper pathway (*A*), which transforms atrazine to cyanuric acid, and a lower pathway (*B*), which catabolizes cyanuric acid to ammonia and carbon dioxide. The entire *Pseudomonas* sp. strain ADP atrazine catabolic pathway is encoded on a single conjugative plasmid (pADP1; GenBank™ NZ_CM003636.1), and the co-ordinates of each gene are shown. The genes of the upper pathway (*atzA*, *atzB*, and *atzC*) encode three hydrolases (AtzA, AtzB, and AtzC). The genes of the lower pathway (*atzD*, *atzG*, *atzE*, *atzH*, and *atzF*) are found as a single operon; *atzG* and *atzH* (shown in *red*) were identified in the work reported here. AtzE had previously been reported as a biuret amidohydrolase (1); however, in the work presented here, we have demonstrated that it is a 1,3-dicarboxyurea forming 1-carboxybiuret amidohydrolase (as shown in the updated lower pathway).

(16, 17)). AtzF is sometimes substituted by the homologous TrzF found in *Enterobacter cloacae* strain 99 (70% identity to AtzF (14)), and AtzE has been replaced with the nonhomologous biuret amidohydrolase (BiuH,[3] a $(\beta/\alpha)_8$ TIM barrel cysteine hydrolase) in *Rhizobium leguminasorum* bv. *viciae* 3841 (Table 1) (1, 19).

Herein, we describe the purification and biochemical and structural characterization of AtzE directly from the model organism *Pseudomonas* sp. strain ADP. We show that AtzE forms a stable $\alpha_2\beta_2$ heterotetramer with a previously unknown small protein, AtzG, that is required for soluble expression of AtzE. Co-expression of AtzE and AtzG in *Escherichia coli* allows, for the first time, the heterologous overproduction of AtzE.

[3] The abbreviations used are: BiuH, biuret hydrolase; BAH, barbituric acid hydrolase; CFE, cell-free extract; bis-tris, 2-[bis(2-hydroxyethyl) amino]-2-(hydroxymethyl)propane-1,3-diol; PPDI, phenyl phosphorodiamidate; PDB, Protein Data Bank; GDH, glutamate dehydrogenase; FDR, false discovery rate; MM-PBSA, Molecular Mechanics Poisson-Boltzmann Surface Area; AS, ammonium sulfate; CV, column volume; MRM, multiple reaction monitoring; MD, molecular dynamics.

Results

AtzE is a 1-carboxybiuret hydrolase

We attempted to produce active AtzE recombinantly in *E. coli* but were unsuccessful (consistent with previous attempts; see Refs. 1, 18). As heterologous expression was unsuccessful, we purified native AtzE from *Pseudomonas* sp. strain ADP. The expression of the *atzDEF* operon is induced under low nitrogen conditions in the presence of cyanuric acid (5), and so we cultured *Pseudomonas* sp. strain ADP on a minimal medium and supplied cyanuric acid as the sole nitrogen source. We were able to detect cyanuric acid amidohydrolase and allophanate amidohydrolase activities in the cell-free extracts (CFE) of these cultures. Cyanuric acid added to the CFE was mineralized (*i.e.* three molecules of ammonia were formed for each molecule of cyanuric acid added). However, we were unable to detect any biuret hydrolysis.

We were able to detect and follow AtzE during CFE fractionation and purification by measuring ammonia production when supplementing fractions with purified AtzD and cyanuric acid. Although the cruder samples during fractionation had

Table 1
Steady-state kinetic parameters for AtzE
The structures of the substrates and nonsubstrates tested (biuret, malonamide, asparagine, glutamine, lysine, citrulline and 2-amino-3-ureidopropionic acid) are shown in Table S2. The standard errors for the K_m and k_{cat} values are given.

Enzyme	Substrate	k_{cat}	K_m	k_{cat}/K_m
		s^{-1}	μM	$s^{-1}\cdot M^{-1}$
AtzD	Cyanuric acid	17.0 ± 1.7	350 ± 42	4.8×10^4
BiuH	Biuret	11.5 ± 0.2	80 ± 7	1.5×10^5
AtzE (*Pseudomonas* sp. ADP)	1-Carboxybiuret	15.5 ± 0.1	63 ± 3	2.5×10^5
AtzE (*E. coli*)	1-Carboxybiuret	14.1 ± 0.2	61 ± 3	2.3×10^5
	1-Nitrobiuret	9.4 ± 3.3	785 ± 217	1.2×10^4
	1-Carboxymalonamide	7.0 ± 0.1	167 ± 17	4.2×10^4
	Succinamic acid	19.0 ± 0.4	523 ± 30	3.6×10^4

high background levels of ammonia, additional ammonia produced from cyanuric acid by AtzD and AtzE was distinguishable from this background. Using ammonia production as a proxy for enzyme activity, we were able to isolate 0.125 mg of pure AtzE per liter of *Pseudomonas* sp. strain ADP culture via ammonium sulfate precipitation and four chromatography steps (Fig. S1*A* and Table S1).

The mass of the isolated protein was 48,119 Da by MS, consistent with the predicted mass of AtzE, and the fragmentation pattern obtained by tryptic digest confirmed that the isolated protein was AtzE. Differential scanning fluorimetry revealed that the melting temperature (T_m) for AtzE isolated from *Pseudomonas* sp. strain ADP was greater than 60 °C. Previous reports had suggested that AtzE could not be produced in *E. coli* because of low stability (1); however, the high T_m of purified AtzE suggested that the poor expression of AtzE in *E. coli* was not due to protein instability.

Purified AtzE had no detectable biuret aminohydrolase activity under any condition tested. However, AtzE-mediated ammonia production was observed when incubated with cyanuric acid and pure AtzD. This suggested that the product of cyanuric acid hydrolysis by AtzD, 1-carboxybiuret, was in fact the natural substrate for AtzE. LC-MS was subsequently used to follow the reactions directly. There was no detectable biuret amidohydrolase activity with AtzE (Fig. 2*A*), but it was detected using the biuret amidohydrolase from a *Rhizobium* species, BiuH (Fig. 2*B*) (19). Moreover, biuret accumulates in reactions in which cyanuric acid is treated with AtzD (Fig. 2*C*), but it does not when AtzE is also included (Fig. 2*D*), suggesting that AtzE prevents biuret's formation via the hydrolytic deamination of 1-carboxybiuret.

Steady-state kinetic data for AtzD, AtzE, and the biuret hydrolase were obtained (Table 1). The product inhibition of AtzD reported previously (11) was alleviated by the addition of AtzE, allowing the determination of rates at concentrations above 0.2 mM cyanuric acid for the first time. The kinetic parameters determined for AtzD were found to be a k_{cat} of 17 s^{-1}, a K_m of 350 μM, and a k_{cat}/K_m of 4.8×10^4 $s^{-1}\cdot M^{-1}$. The steady-state kinetic parameters for AtzE were found to be similar to those of the *Rhizobium* biuret amidohydrolase for its physiological substrate, with a k_{cat} of 16 s^{-1}, a K_m of 63 μM, and a k_{cat}/K_m of 2.5×10^5 $s^{-1}\cdot M^{-1}$ for AtzE, and a k_{cat} of 12 s^{-1}, a K_m of 80 μM, and a k_{cat}/K_m of 1.5×10^5 $s^{-1}\cdot M^{-1}$ for BiuH (Table 1).

We also tested AtzE with a number of potential additional substrates, including the product of barbituric acid hydrolase (1-carboxymalonamide), analogs of AtzE's natural substrate (1-nitrobiuret and succinamic acid), structurally related α-amino acids (citrulline, lysine, glutamine, asparagine, and 2-amino-3-propionic acid) and the biuret analog malonamide (Table 1 and Table S2). None of the α-amino acids were substrates for AtzE nor was the biuret analog malonamide (Table S2). However, 1-carboxymalonamide, 1-nitrobiuret, and succinamic acid were all substrates for AtzE with k_{cat}/K_m values of 4.2×10^4, 1.2×10^4, and 3.6×10^4 $s^{-1}\cdot M^{-1}$, respectively (Table 1; Fig. 3).

LC-MS was used to identify that the product of AtzE-mediated hydrolysis of 1-carboxymalonamide is 1-carboxymalonamic acid through hydrolysis of the terminal amine. Succinamic acid has a terminal amide rather than a ureido group, suggesting that AtzE hydrolyzes the amide from this substrate in an equivalent reaction to that of the hydrolysis of 1-carboxymalonamide. We therefore propose an update to the cyanuric acid catabolic pathway, in which AtzE acts as a 1-carboxybiuret amidohydrolase, producing 1,3-dicaboxyurea and ammonia (Figs. 1*B* and 2*D*).

Structure of AtzE reveals the presence of an unexpected but essential ancillary protein

X-ray crystal structures were obtained for purified AtzE (PDB code 6C62; Table 2), and purified AtzE was treated with the suicide inhibitor phenyl phosphorodiamidate (PPDI; PDB code 6C6G; Table 2). The crystals of AtzE had two independent molecules of AtzE in the asymmetric unit, and the entire chain could be traced for both molecules (Fig. 4). Surprisingly, after positioning the AtzE chains in the structure, additional protein density was observed (Fig. 4). The density was of high enough quality to allow for initial estimation of the unknown sequence, which was identical to the predicted translation product of a short (204 bp) unannotated intergenic region between *atzD* and *atzE* in the *atzDEF* operon on the pADP1 plasmid (Fig. 1*B*).

Examination of purified AtzE by SDS-PAGE indicated a small co-purified protein of approximately the same size as the protein that co-crystallized with AtzE (Fig. S1*B*). Mass spectrometry of the AtzE purified from *Pseudomonas* confirmed the presence of an unidentified protein having a mass of 7422 Da (Fig. S1*C* and Table S3), and analysis of tryptic fragments of the 7.4-kDa protein confirmed that it was encoded by the small unannotated ORF located between the *atzD* and *atzE* genes (Fig. S2), which we have termed *atzG* (Fig. S1). Along with the two molecules of AtzE, there are two molecules of AtzG in the asymmetric unit (Fig. 4*B*).

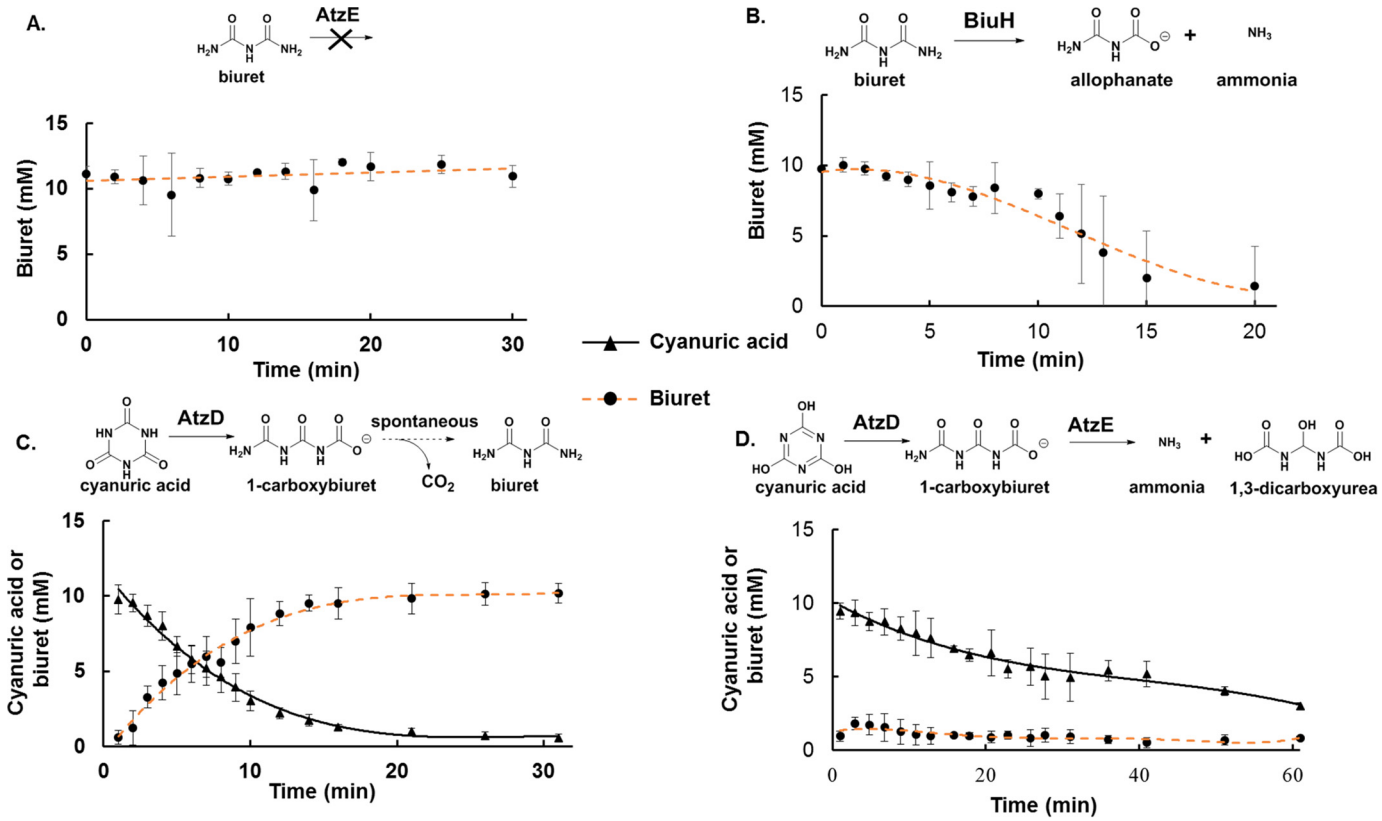


Figure 2. LC-MS analysis of AtzE activity with biuret and 1-carboxybiuret. The concentrations of cyanuric acid (*triangles*) and biuret (*circles*) present in each reaction are shown. Four reaction conditions were tested as follows: *A,* biuret in presence of AtzE; *B,* biuret in presence of the biuret hydrolase BiuH; *C,* cyanuric acid in presence of AtzD; and *D,* cyanuric acid in presence of AtzD and AtzE. Reactions were done in triplicate, and the standard deviations are shown.

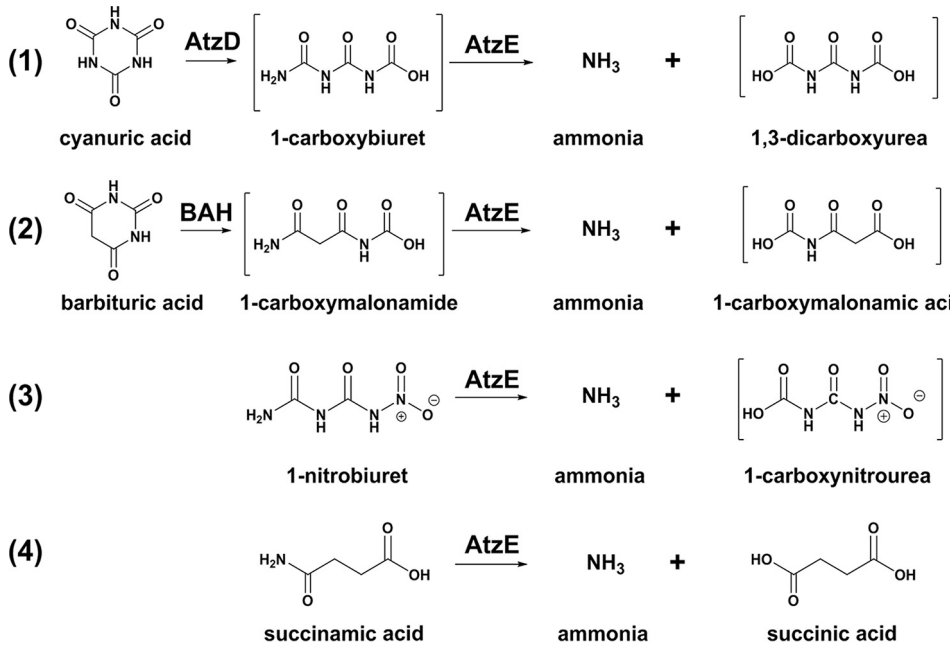


Figure 3. AtzE-dependent deamination reactions. The schemes for AtzE-dependent deamination are shown as follows: *1,* 1-carboxybiuret; *2,* 1-carboxymalonamide; *3,* 1-nitrobiuret and *4,* succinamic acid. 1-Carboxybiuret and 1-carboxymalonamide were produced *in situ* by cyanuric acid amidohydrolase (AtzD) and barbituric acid amidohydrolase (BAH), respectively. *Square brackets* indicate compounds that spontaneously decarboxylate under the reaction conditions used.

Of the protein structures available in the PDB (20), AtzE is most similar to the glutamine deaminase component (chain A; 35% sequence identity) of 4WJ3, a glutamine transamidosome from *Pseudomonas* (21); 428 residues aligned (out of 457) with a 1.5-Å root mean square deviation over the aligned backbone atoms (Fig. 5). The transamidosome is a three-protein complex (GatCAB) that is essential for the production of correctly charged tRNA in bacteria via the transfer of ammonia from

Table 2
Crystal structure parameters
Note: values in parentheses are for the highest resolution shell. NA means not applicable.

	6C62 (Native)	6C6G (PPDI)
Data collection		
Space group	I2	I2
Cell dimensions		
a, *b*, *c* (Å)	79.5, 89.0, 141.7	78.6, 88.9, 141.9
α, β, γ (°)	90, 101.9, 90	90, 101.3, 90
Resolution (Å)	1.95 (1.99–1.95)	2.10 (2.16–2.10)
R_{merge}	0.233 (1.545)	0.326 (1.282)
R_{pim}	0.095 (0.622)	0.126 (0.496)
$I/\sigma I$	6.9 (2.1)	6.7 (2.0)
CC1/2	0.991 (0.804)	0.982 (0.656)
Completeness (%)	100 (100)	100 (100)
Redundancy	6.9 (7.1)	7.6 (7.6)
Refinement	**2 dimers**	**2 dimers**
Resolution (Å)	69.3–1.95	69.6–2.10
Unique reflections	66,810	53,225
R_{work}/R_{free}	15.3/19.1	16.7/21.0
No. of atoms	8659	8523
Protein	7930	7882
Inhibitor	NA	8
Water	726	639
B-Factors (Å^2)	17.5	16.8
Protein	17.4	16.9
Inhibitor	NA	12.4
Water	24.6	21.1
Root mean square deviations		
Bond lengths (Å)	0.017	0.014
Bond angles (°)	1.726	1.643

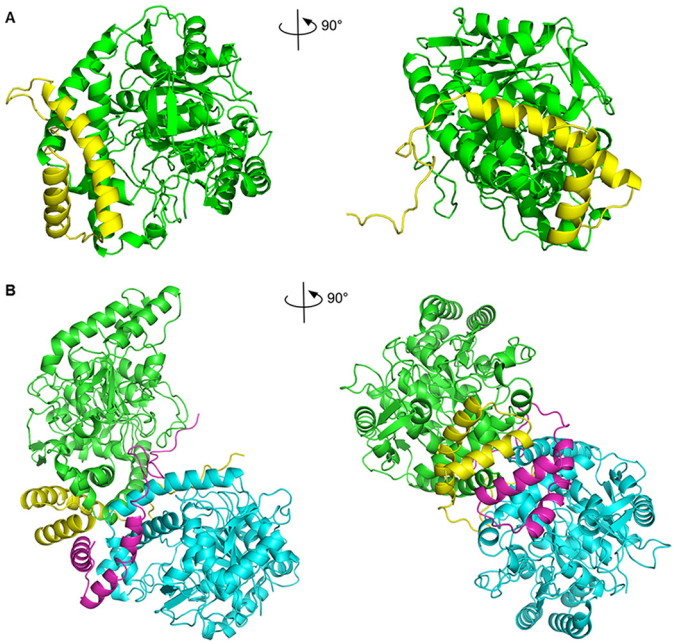


Figure 4. X-ray crystal structure of the unexpected AtzE complex. *A*, AtzE monomer (*green*) has a typical amidase fold. Surprisingly, it is associated with a small (68 amino acids) protein (*yellow*). *B*, $\alpha_2\beta_2$ heterotetramer, containing two molecules of AtzE (*green* and *cyan*) and two molecules of the 68-amino acid protein (*yellow* and *magenta*).

glutamine to acidic amino acid-charged tRNA (21). Deamination occurs in the active site of an amidase (GatA). The ammonia is then channeled, via an "ammonia tunnel," to a phosphate-activated, aspartate-charged tRNA in the second active site (in GatB). Channeling prevents loss of ammonia and excludes water from the active site of GatB (22).

AtzG is homologous to the GatC protein in the GatCAB complex and binds AtzE in the same position as GatC binds

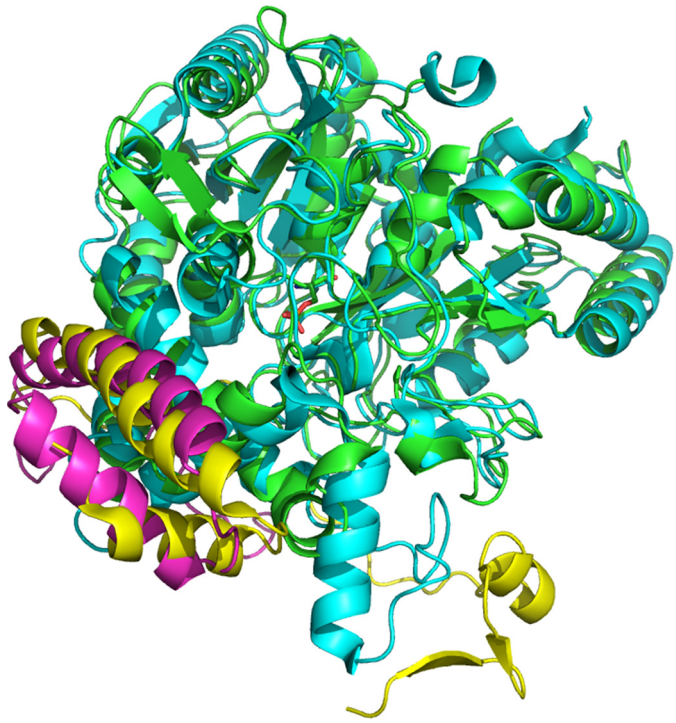


Figure 5. Superposition of the AtzEG heterodimer and the GatA–GatC complex. AtzE and AtzG are shown in *green* and *magenta*, respectively. GatA and GatC are shown in *cyan* and *yellow*, respectively. The phosphoserine from the AtzE X-ray crystal structure (Ser-174) is shown in *stick* representation.

GatA in the GatCAB complex. The function of GatC is to coordinate the complex by forming appropriate interactions with both GatA and GatB. AtzG appears to fulfill a similar role, gluing together the AtzE dimer resulting in a tightly associated heterotetramer. According to the program PISA (23), the AtzEG heterotetramer sequesters over 12,500 Å^2 of buried surface area, giving a ΔG of interaction of −108 kcal/mol. AtzG has extensive interactions with both AtzE and AtzG of the dimerization partner (Fig. 4). Notably, AtzG is 28 amino acids shorter than GatC (68 *versus* 96 amino acids), with the additional 28 amino acids of GatC forming the majority of its interactions with GatB (Fig. 5) (21). GatC also participates in the formation of the ammonia tunnel in the GatCAB complex. Interestingly, the ammonia tunnel is retained in AtzEG (Fig. 6). The retention of the ammonia tunnel may facilitate a high-reaction rate by provide rapid egress of the ammonia product.

When AtzE and AtzG were co-expressed in *E. coli*, a soluble (Fig. S1*B*), active enzyme that was catalytically indistinguishable from AtzEG purified from *Pseudomonas* sp. strain ADP was obtained (Table 1). This confirmed that AtzG promotes the production of soluble, active AtzE. Notably, GatA cannot be overproduced in the absence of GatC (24, 25).

Substrate specificity and reaction mechanism

Molecular dynamics were used to understand the behavior of 1-carboxybiuret in the active site. 1-Carboxybiuret binds via a series of hydrogen-bonding interactions: the terminal carboxylate of the substrate binds to Tyr-125, Asn-172, and Gln-402; the main chain carbonyl of Gly-126 binds the terminal adjacent amines, and the main chain NH of Gln-402 binds the carbonyl between those amines (Fig. 7*A*). The interactions with the ter-

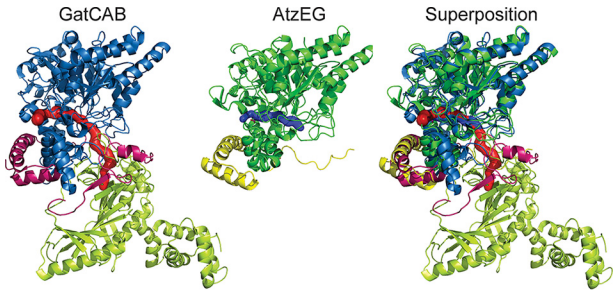


Figure 6. Conservation of the transamidosome ammonia tunnel in AtzEG. The ammonia tunnel is shown in the GatCAB complex, and a comparable tunnel is seen in AtzEG. Superposition of the two protein complexes show that the tunnels overlap for the whole length of the AtzEG tunnel. CAVER (30) was used to generate this figure.

minal carboxylate substantially strengthen the binding of 1-carboxybiuret relative to biuret and would explain the lack of activity with biuret. From a 500-ns MD simulation, we found that the ΔG total as a function of time (Fig. 7*C*) for 1-carboxybiuret reached significantly lower values than biuret (−33 and −12 kcal/mol, respectively). Carboxybiuret was retained by the AtzE active site for the duration of the simulation, whereas biuret was observed to leave the active site after 400 ns.

The catalytic mechanism of AtzE is likely to resemble that of other enzymes with Ser-*cis*Ser-Lys catalytic triads (*e.g.* malonamidase, allophanate amidohydrolase, DNA polymerase V accessory protein, and signal peptidase) (13, 26–28). The catalytic triad of AtzE, inferred from crystallographic data and homology with other amidases, was composed of Lys-74, *cis*-Ser-150, and Ser-174. In the substrate-docked structure, Ser-174 is positioned for attack at the carbonyl of the terminal amide of 1-carboxybiuret (Fig. 7*B*). Moreover, extra density consistent with a phosphate group was seen associated with Ser-174 in the crystal structure of AtzE treated with PPDI (PDB code 6G6C). This suggests that the phosphodiamidate moiety had been transferred to the catalytic serine (Ser-174) during the crystallization process and that Ser-174 is the nucleophilic residue in the catalytic triad. Mass spectrometry of proteolyzed PPDI-treated AtzE showed that 70% of the enzyme was labeled with monoaminophosphate at Ser-174 and 30% with diaminophosphate. This suggests that the diaminophosphate is labile, releasing an ammonia in an aqueous solution (consistent with previous reports) (1).

There are three sets of potential reaction products from the deamination of 1-carboxybiuret: allophanate and carbamate, diacarboxyammonia and urea, or 1,3-dicarboxyurea and ammonia. These products are difficult to resolve using the GDH-coupled assay that we employed, as carbamate and diacarboxyammonia readily decompose in water to produce ammonia. As GatA produces ammonia, we might expect that AtzE does, too; furthermore, the substrate range of AtzE includes succinamic acid from which carbamate cannot be produced. We therefore propose that AtzE deaminates 1-carboxybiuret, producing 1,3-dicarboxyurea (Fig. 1*B*), which then monodecarboxylates to form allophanate (the substrate for AtzF) (13, 15). Currently, it is unclear whether the formation of allophanate is spontaneous or enzyme-mediated.

The suggested mechanism for AtzE is described in Fig. 8. In the first step, Lys-74, Ser-150, and Ser-174 all form a hydrogen bonding network, where Lys-74 acts as a general base to activate the catalytic Ser-174 through the *cis*-Ser-150–bridging ligand. The activated Ser-174 performs a nucleophilic attack on the terminal amide end of 1-carboxybiuret, leading to the formation of a covalent acyl-enzyme intermediate and ammonia. Lys-74 then acts as a general base, activating water and leading to the hydrolysis of the acyl enzyme, releasing 1,3-dicarboxyurea, and restoring the active site to its original state.

It is interesting that AtzE has no detectable activity with GatA substrates (*e.g.* α-amino acids), given their high degree of structural similarity. We compared the structures of AtzE with docked 1-carboxybiuret and GatA with glutamine bound in its active site (Fig. 9). The AtzE active site is highly conserved when compared with that of GatA, albeit with some differences (Fig. 9). The catalytic triad is conserved (Lys-79, Ser-154, and Ser-178 in GatA), as are Gly-126 (GatA Gly-130), Ser-169 (GatA Ser-173), Thr-171 (GatA Thr-175), and Gly-173 (GatA Gly-177). Among the key differences between the two enzymes are Phe-127 and Tyr-125 (Gly-131 and Met-129 in GatA), which appear to stabilize 1-carboxybiuret in the AtzE active site via π-stacking interactions to the conjugated carbonyl system in the substrate. Phe-127 would also sterically prevent the binding of glutamine and other α-amino acids, which cannot adopt a planar orientation like 1-carboxybiuret, and therefore likely has a prominent role in substrate specificity. Additionally, Asp-425 in GatA, which is required for hydrogen bonding to the backbone nitrogen of the substrates, is Gly-398 in AtzE.

Discussion

Previously, AtzE had been reported to be a biuret hydrolase (1, 5, 14, 19), largely through comparison with the nonhomologous biuret hydrolase of *Rhizobium* sp. (Table 1). However, our findings demonstrate that AtzE is a 1-carboxybiuret hydrolase and that the terminal carboxylate is essential for correct substrate binding. Sequence and structural homology suggest that AtzEG may have been "repurposed" from the bacterial glutamine transamidosome, in which a similar complex, GatAC, is essential in channeling a solvent-labile intermediate between the complex's two active sites (22). AtzG does not appear to fulfill the same complex co-ordinating function as GatC, instead it may be that the obligate AtzEG complex is simply a vestige of their shared evolutionary past. Interestingly, *Pseudomonas* sp. strain ADP possesses genes that encode a predicated GatCAB complex (*gatA*, GenBank™ KSW28066.1; *gatB*, GenBank™ KSW26274.1; *gatC*, KSW26275.1). However, as *atzEG* are encoded by a gene cluster on a transposable element, carried on a self-transmissible plasmid (5, 29), it seems unlikely that *atzEG* was recruited from this specific gene cluster.

Despite the structural and mechanistic similarities between AtzEG and its likely ancestor (a GatAC-like protein), the substrate range for AtzEG excludes the GatAC substrates asparagine and glutamine. GatCAB fulfills an essential role in the production of asparagine- and glutamine-charged tRNA at the expense of glutamine (31), and it is plausible that AtzEG rapidly evolved away from these substrates under selective pressure to eliminate perturbations in core cellular functions (*i.e.* aminoacyl tRNA production and amino acid pools).

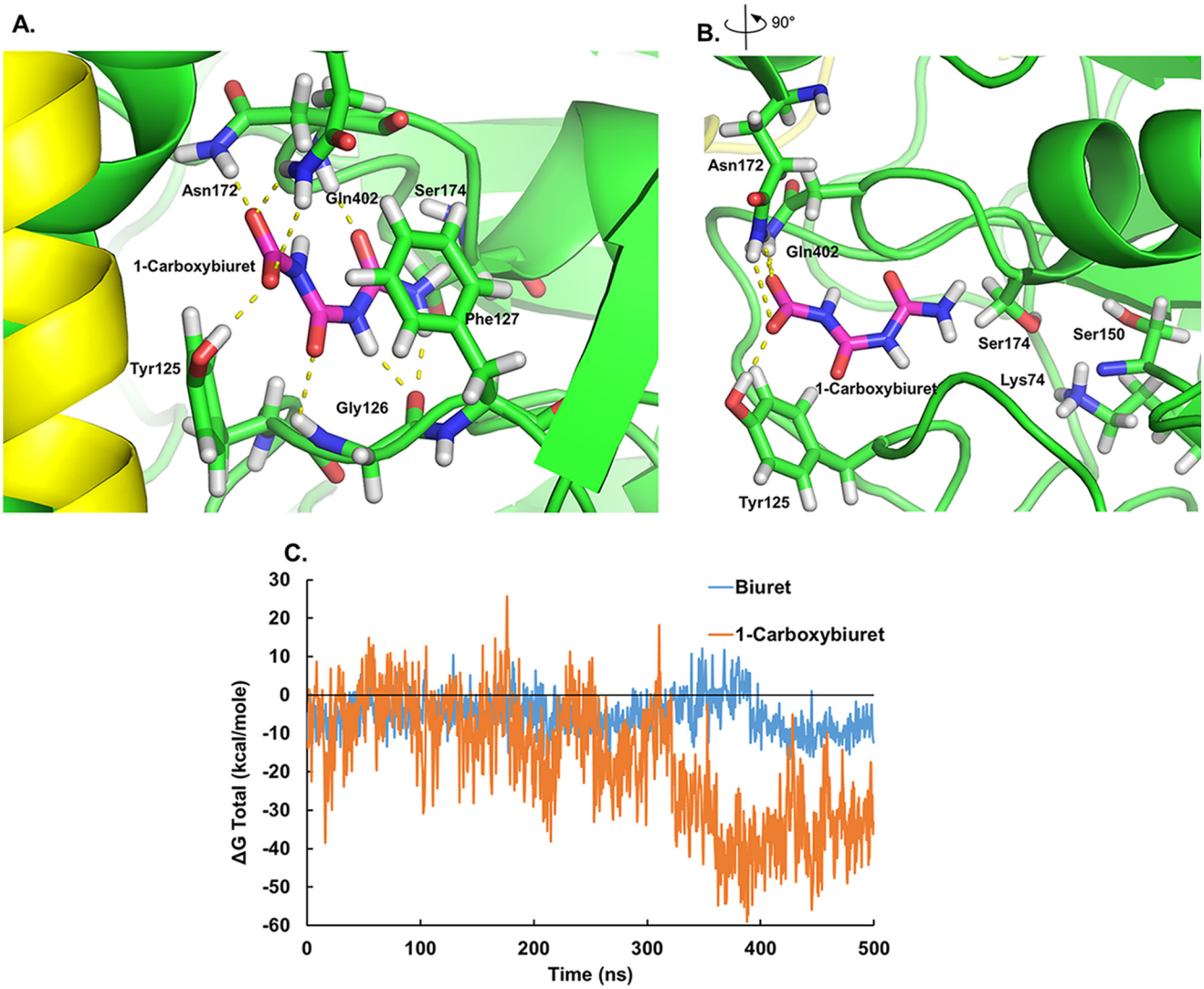


Figure 7. Docking and molecular dynamics of substrate in the AtzE active site. *A,* 1-carboxybiuret docked into the AtzE active site in its most stable configuration (Δ*G* minimal). 1-Carboxybiuret is shown in *magenta*; AtzE is shown in *green*, and AtzG is shown in *yellow*. Hydrogen bonds between the protein and substrate are shown as *dashed lines. B,* 90° rotation of *A* showing the position of the catalytic amino acids Ser-174, Ser-150, and Lys-74; for clarity, the stick representations of amino acids Gly-126 and Phe-127 have been removed. *C,* variation in Δ*G* (kcal/mol) as a function of time (nanoseconds) during a molecular dynamics simulation of biuret and 1-carboxybiuret docked into the active site of AtzEG. The lowest Δ*G* total reached during the 1-carboxybiuret MM-PBSA is −59.0 at 388 ns of simulation, whereas the lowest Δ*G* total for biuret is only of −16.7 at 425 ns of simulation. Although 1-carboxybiuret adopts a stable conformation over time, biuret does not and escapes the active site at $t = 400$ ns.

Although not itself catalytic, AtzG is essential for correct AtzE function, and previous attempts to express AtzE in heterologous hosts were unsuccessful because the gene encoding AtzG had been overlooked by automated annotation. There is an increasing body of evidence that small, overlooked open-reading frames often encode functional low-molecular-weight proteins (32, 33). It may be that a subset of these have structural roles, like that of AtzG, and are required to promote or enhance soluble protein expression or stability of their partner proteins.

Surprisingly, the cyanuric acid-mineralizing operon contains a second unannotated ORF between the genes encoding AtzE and AtzF (Fig. 1*B*) that is predicted to encode a 129-amino acid protein (GenBank™ WP_064987550.1), tentatively named AtzH, which belongs to the DUF3225 family of uncharacterized proteins. Preliminary proteomics with *Pseudomonas* sp. strain ADP indicate that AtzH is expressed under the same conditions as AtzD, -E, -F, and -G. Investigations are underway to determine the role of AtzH.

Experimental procedures

Purification of AtzE from Pseudomonas sp. strain ADP

Pseudomonas sp. strain ADP's growth conditions were optimized to induce expression of the cyanuric acid catabolism operon. 40 liters of *Pseudomonas* sp. strain ADP was grown in 500-ml cultures in 2-liter flasks. Cultures were grown in minimal medium at 28 °C and shaken at 200 rpm until an OD_{600} of 0.4 was reached. The minimal medium contained 10 mM cyanuric acid as a sole nitrogen source, 26 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.5 mM NaCl, 200 μM $MgSO_4$, 2.9 mM sucrose, 3.4 mM trisodium citrate, 44 μM $CaCl_2$, 20 ml/liter of vitamin stock, and 1 ml/liter of trace elements, adapted from Balotra *et al.* (13). The cells were harvested by centrifugation (8000 × *g* for 15

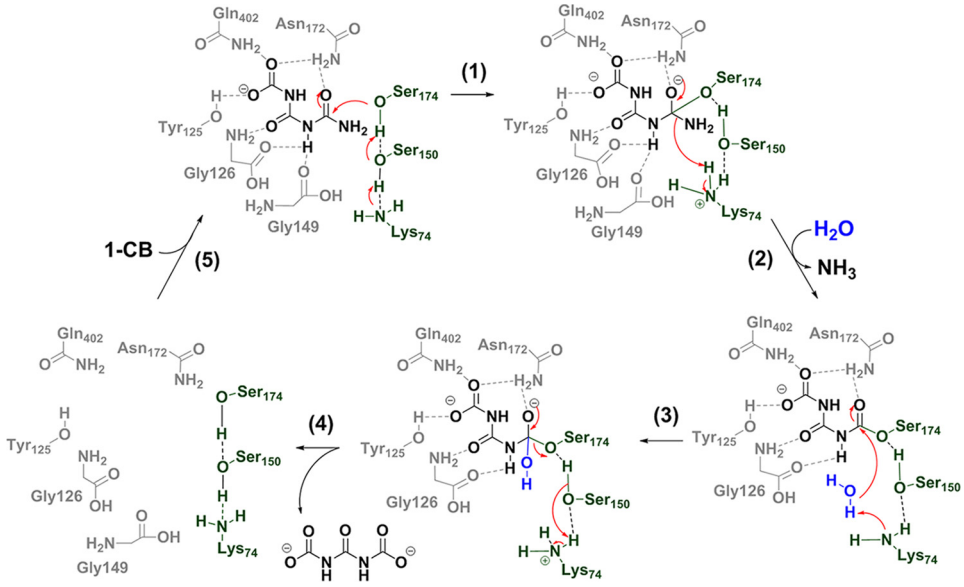


Figure 8. AtzE reaction mechanism. The catalytic triad (Lys-74, Ser-150, and Ser-174; shown in *green*) forms a hydrogen-bonding network with the substrate; Lys-74 activates the catalytic (Ser-174) via the *cis*-Ser-150–bridging ligand. Ser-174 performs a nucleophilic attack on the terminal amide of the substrate, forming the covalent acyl–enzyme intermediate and releasing ammonia. Lys-74 then acts as a general base, activating water (*blue*) and leading to the hydrolysis of the acyl enzyme, releasing 1,3-dicarboxyurea and regenerating the active site.

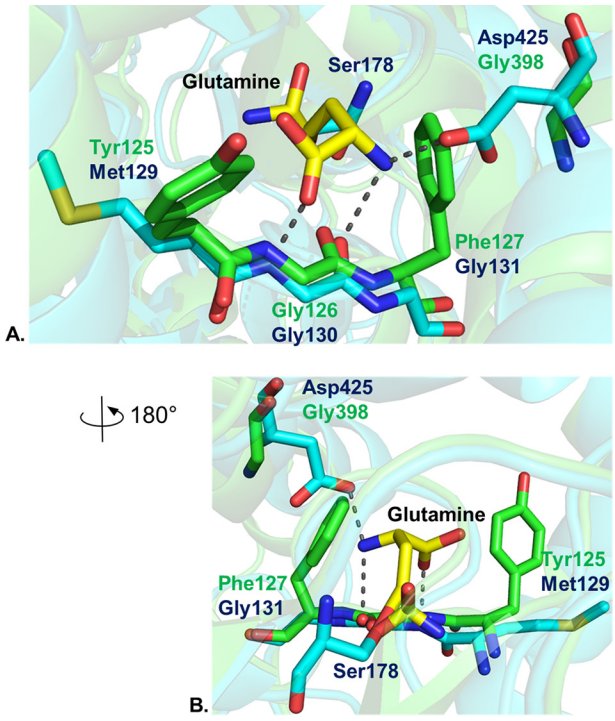


Figure 9. Comparison of the active sites of AtzE and GatA bound to 1-carboxybiuret and glutamine, respectively. The AtzE active site (*green*) is shown superposed on the GatA active site (*cyan*). The GatA substrate (glutamine; *yellow*) is also shown. Hydrogen bonds between GatA and glutamine are shown as *dotted lines*. *A* and *B* are rotated 180° relative to each other to highlight the positions of the α-amine of glutamine and Phe-127 in AtzE.

min) and lysed as described in Ref. 13, and the soluble supernatant was used for further purification.

AtzE was purified by a four-step purification process: ammonium sulfate (AS) precipitation, followed by hydrophobic interaction, anion-exchange and size-exclusion chromatography. A solution of concentrated AS was added to the supernatant to give a final concentration of 0.6 M and stirred for 4 h at 4 °C. After centrifugation (18,000 × *g* for 30 min), the soluble fraction was syringe-filtered through a 0.22-μm filter.

The supernatant from the AS cut was purified by hydrophobic interaction chromatography using a 65-ml column packed with phenyl-Sepharose preparation-grade resin (GE Healthcare), equilibrated with 25 mM potassium phosphate and 0.6 M ammonium sulfate, pH 7.5. After running 4 column volumes (CV) of 0.4 M AS, a reverse gradient from 0.4 to 0.09 M AS over 10 CV was used to elute the protein. Fractions containing a 48-kDa band (SDS-PAGE) eluted between 0.27 and 0.14 M AS. These fractions were pooled and concentrated to 10 ml volume (Amicon Ultra-15 centrifugal filter unit, Ultracel-30 membrane). A desalting step was performed using a 53-ml HiPrep C26/10 desalting column (GE Healthcare), equilibrated with 25 mM potassium phosphate, pH 8.5, for 1.5 CV.

The crude protein fraction was purified by anion-exchange chromatography using a 77-ml column packed with Q-Sepharose preparation grade resin (GE Healthcare), equilibrated with 25 mM potassium phosphate, pH 8.5. The protein was eluted with a gradient of 0 – 0.25 M NaCl over 11 CV. Fractions eluted between 0.16 and 0.23 M NaCl were found to contain a 48-kDa band (SDS-PAGE). After pooling, these fractions were concentrated to 12 ml using an Amicon Ultra-15 centrifugal filter unit.

This combined sample was applied to a 130-ml size-exclusion column packed with Superdex 200 preparation grade resin (GE Healthcare), equilibrated with 25 mM potassium phosphate and 200 mM NaCl, pH 7.5, and eluted over 1.5 CV with the same buffer. All chromatography steps were performed using an ÄKTA purifier UPC 10 (GE Healthcare).

Cloning

A synthetic version of the *atzE* gene, coding for the protein NP_862538 (NP_862538.1), was obtained from GenScript Corp. (Piscataway, NJ), inserted into pUC57, and flanked by the restriction sites NdeI and BamHI. The genes were subcloned

into the NdeI and BamHI sites of pETcc2, described in Peat *et al.* (11). Following the lack of soluble AtzE expression, we reamplified *atzE* flanked with NcoI and AvrII with the primers (5′ to 3′: GTACACCCATGGGAATGAAGACAGTAGAAATTATTGAAGG and TTTTTTGAGCTCCCACATTTCAGTCGGGCGATAC). *atzE* was subcloned into the first multiple cloning site of the pACYCDuet-1 vector (Novagen), a low copy number expression vector for co-expression of two genes, with an in-frame N-terminal thrombin-cleavable His_6 tag. *AtzG* coding for the protein WP_082996223 (WP_082996223.1) was cloned directly from *Pseudomonas* sp. ADP genomic DNA with the primers (5′ to 3′: CGACGACATATGCTCGAGATGACGGAAACTG and TGCTGCGAGCTCCCTAGGTCAGATATCTTCTGC), before being subcloned into the second cloning site of the pACYCDuet-1 vector using the enzymes NdeI and XhoI.

The *atzD* coding for the protein NP_862537 (NP_862537.1) had been cloned into pETcc2 previously (11, 13, 18). Ligations were performed using T_4 ligase. Restriction enzymes and T_4 ligase were obtained from Thermo Fisher Scientific. The biuret hydrolase from *R. leguminasorum* bv. *viciae* 3841 (BiuH) (AM236084.1) and barbituric acid hydrolase from *Nocardioides* sp. JS614 coding for the protein ABL81019 (ABL81019.1) were expressed in *E. coli* as described elsewhere (12, 19).

Heterologous protein expression

The expression vectors were used to transform *E. coli* BL21 (λDE3) cells (Invitrogen). Bacteria were grown on Luria-Bertani (LB) medium containing 100 μg/ml ampicillin for the pETcc2 constructs or 34 μg/ml chloramphenicol for the pACYCDuet-1 construct. Cells were grown with shaking at 200 rpm at 28 °C. Protein expression was induced at an OD_{600} of 0.8 by addition of isopropyl β-D-1-thiogalactopyranoside (1 mM final concentration).

Cells were harvested 24 h after induction by centrifugation at 5000 × *g* for 15 min using an Aventi J-E centrifuge (Beckman Coulter, Indianapolis, IN), resuspended in lysis buffer (25 mM potassium phosphate, 5 mM imidazole, pH 7.5), and lysed by passage through a Microfluidics homogenizer M-110P five times at 15,000 p.s.i. The lysis was followed by centrifugation at 18,000 × *g* for 45 min to pellet the cellular debris, and the soluble fraction was used for further purification.

The soluble fraction was syringe-filtered throughout a 0.22-μm filter. The filtrate was purified using a 5-ml nickel-nitrilotriacetic acid superflow cartridge (Qiagen) with a gradient from 5 mM imidazole to 500 mM over 10 CV. SDS-polyacrylamide gel analysis was performed to assess the purity of the fractions.

Enzyme assays

Cyanuric acid hydrolysis by AtzD was followed by UV spectrophotometry at 214 nm (13) using 0–0.6 mM cyanuric acid in 25 mM phosphate buffer, pH 9.

A GDH (Sigma-Aldrich)-coupled reaction was used to measure ammonia release and determine the activity rate in the BiuH and AtzE-dependent reactions. GDH catalyzes the NADH-dependent amination of α-ketoglutarate. Ammonia production by BiuH and AtzE was followed using the decrease of absorbance by UV spectrophotometry at 340 nm that was due to the oxidation of NADH by GDH. 1.25 units of GDH were used in a 250-μl reaction volume, and the final concentrations of α-ketoglutarate and NADH were 3.5 and 0.2 mM, respectively (Fig. S3) (13).

Monitoring the AtzE activity during purification steps from *Pseudomonas* sp. strain ADP was initially done with its reported substrate, biuret (1, 5, 13). However, as no ammonia could be detected, the fractions were supplemented with 5.1 nM AtzD and 0.2 mM cyanuric acid, in 25 mM phosphate buffer, pH 9, to generate the substrate *in situ*. There was a significant background of ammonia in the early steps/cruder fractions; however, AtzE-dependent ammonia production was detectable after the background rate (measured in the absence of cyanuric acid) had been subtracted.

Steady-state kinetics for AtzD were obtained by using 41 nM AtzD, in the presence of various amounts of cyanuric acid substrate ranging from 0 to 0.6 mM. Cyanuric acid degradation was measured by following the decrease in absorbance at 214 nm, but above 0.2 mM AtzD activity was inhibited. To determine whether the substrate or the product was inhibiting AtzD activity, we pooled 41 nM AtzD, 24 nM AtzE, and 5 units/μl GDH in the presence of 0–0.6 mM cyanuric acid. This removed the inhibition previously observed at cyanuric acid concentrations above 0.2 mM and allowed for the determination of steady-state kinetics for AtzD. Steady-state kinetics for AtzD were determined using 41 nM AtzD, 24 nM AtzE, and 5 units/μl GDH, after ensuring that in these proportions the AtzD-mediated reaction was the rate-limiting reaction and with the assumption that the rate of ammonia production recorded was proportional to the AtzD activity rate.

Steady-state kinetics for AtzE with the product of AtzD or barbituric acid hydrolase (BAH) were obtained under conditions where the AtzE-mediated reaction was the rate-limiting reaction (*i.e.* 3.5 times slower than either the GDH, AtzD, or barbituric acid hydrolase-mediated reactions). AtzE steady-state kinetics were obtained by using 41 nM AtzD/48 nM barbituric acid hydrolase, 12 nM AtzE, and 5 units/μl GDH. Substrate (cyanuric acid or barbituric acid) was added over the range of 0–0.6 mM. As AtzE was rate-limiting, it was assumed to be a good estimate of the rate of AtzE.

Other potential AtzE substrates tested were as follows: 1-nitrobiuret; 1-carboxymalonamide; succinamic acid; 2-amino-3-ureidopropionic (Albizziin); citrulline; lysine; glutamine; asparagine; biuret; and malonamide. Specific activities were obtained with 12–400 nM AtzE and 5 units/μl GDH and 1.5 mM substrates in 25 mM phosphate buffer, pH 9. Steady-state kinetics were obtained with 1-nitrobiuret using 12 nM AtzE and 5 units/μl GDH. Substrate, 1-nitrobiuret, was added over the range of 0–2.5 mM. Steady-state kinetics were obtained with 1-carboxy-malonamide using 12 nM AtzE, 48 nM BAH, and 5 units/μl GDH. Substrate (barbituric acid) was added over the range of 0–4 mM.

Biuret hydrolase kinetic data were obtained by using 11 nM biuret hydrolase and 5 units/μl GDH in the presence of various amounts of biuret ranging from 0 to 0.7 mM. All the kinetics constants were calculated with Hyper32 software, fitting the rate data to Michaelis-Menten Equation 1,

$$\frac{d[P]}{dt} = \frac{V_{max}[S]}{K_m + [S]} \quad \text{(Eq. 1)}$$

The steady-state kinetic values for AtzE with 1-nitro-biuret were estimated using Lineweaver-Burk Equation 2,

$$\frac{1}{V_0} = \frac{K_m}{V_{max}}\frac{1}{[S]} + \frac{1}{V_{max}} \quad \text{(Eq. 2)}$$

LC-MS/MS multiple reaction monitoring (MRM)-based enzyme assay was used to follow the fate of cyanuric acid and biuret when in the presence of either BiuH, AtzD, and/or AtzE. Reaction samples were run in triplicate. Four reactions were composed of the following: 10 mM biuret with 580 nM BiuH, 10 mM biuret with 374 nM AtzE, 10 mM cyanuric acid with 420 nM AtzD, and 10 mM cyanuric acid with 420 nM AtzD and 374 nM AtzE. 5 μl of reaction mix was injected every 2 min for 30–60 min. Analysis was performed using a Waters Alliance 2695 separation module coupled to a Waters TQD detector, but injections were made using a manual injector with a 5-μl injection loop. The HPLC column used was a Phenomenex Develosil RPAQEOUS-AR 5u C30 (dimension: 250 × 4.6 mm), which was heated to 35 °C. A flow rate of 0.7 ml/min was used (15:1 flow splitter was used after separation). The eluent used was 15% (v/v) acetonitrile and 0.1% formic acid in water. Two simultaneous but independent MRM analyses were performed as follows: MRM parameters for biuret (parent m/z = 103.9424): channel 1 (ES+): daughter (m/z): 17.9654, dwell (s): 0.025, cone (V): 20, collision (V): 10; and channel 2 (ES+): daughter (m/z): 43.8956, dwell (s): 0.025, cone (V): 20, collision (V): 20. MRM parameters for cyanuric acid (parent m/z = 127.8577) were channel 1 (ES−): daughter (m/z): 41.9357, dwell (s): 0.025, cone (V): 28, collision (V): 14. MRM peaks were integrated using Waters MassLynx version 4.1 software.

Proteomics

In-gel digestion was performed by resuspending pieces of bis-acrylamide gel containing 50–100 μg of protein in 100 μl of 25 mM ammonium bicarbonate. After addition of 5 μl of 15% DTT, the mix was incubated at room temperature for 30 min, followed by addition of 3.5 μl of 40% acrylamide and another 30 min incubation at room temperature. The sample was digested with 0.5 μg of trypsin at 37 °C overnight. To stop the digestion, formic acid was added to a final concentration of 0.1–1% and then the trypsin-digested peptides were eluted from gel pieces by sonication for 15 min and further incubated at room temperature for 30 min. Typically, 1–5 μl of sample was used for LC-MS analysis.

The trypsin-digested peptides were separated by reversed-phase HPLC using an Agilent NanoFlow LC system (1260 Infinity). The peptides were loaded onto a NanoLC trap column (3 μm, ChromXP C18CL, 120 Å 0.5 mm × 350 μm from Eksignet Technologies) by autosampler at a flow rate of 5 μl/min and continuously desalted for 5 min. The desalted peptides were eluted from the trap column and separated on a NanoLC column (3 μm, ChromXP C18CL, 120 Å, 15 cm × 75 μm from Eksigent) with a 0–40% gradient in 100 min consisting of buffer A (0.1% formic acid in water) and buffer B (0.1% formic acid in acetonitrile). The flow rate of Nano pump was set to 300 nl/min.

Peptides were analyzed using positive and high-sensitive mode on an AB Sciex Tripletof 5500 mass spectrometer. The voltage of Nano Spray II was set to 2300 kV. In data-dependent acquisition mode, the mass window for precursor ions of the quadrupole mass analyzer was set to ± 1 m/z. The precursor irons were fragmented by nitrogen collision gas. The MS1 survey scan (250 ms; mass 350–1500) was carried out. 50 of the most abundant precursors were selected for MS/MS scan (50 ms; mass 100–1800). MS/MS spectra were obtained for product ions that had a charge state of 2–5 and were above 10 counts/s. Rolling collision energy was used, and automatic calibration was carried out after every five sample runs.

Mass spectrum data were analyzed using the Paragon algorithm of ProteinPilot (AB Sciex). The detected protein threshold was set to >0.05(10%). The matching peptide sequences were identified against the *Pseudomonas* sp. ADP proteome found in Uniprot. The false discovery rate (FDR) analysis tool algorithm of ProteinPilot provided a global FDR of 1% and a local FDR at 1% in all cases.

PPDI-treated AtzE samples were digested with trypsin using the SP3 digestion protocol (34). Peptides were separated using an UltiMate nanoUPLC system, utilizing a 60-min gradient on an Acclaim Pepmap 100 column (25 cm × 75 μm inner diameter with 3-μm particles). High-resolution MS/MS data were obtained on an Orbitrap Fusion Lumos mass spectrometer and top 20 multiply charged species selected for fragmentation in high-high mode with the Orbitrap resolution set at 75,000.

Orbitrap MS/MS data were searched against a focused decoy database containing AtzE, *E. coli,* and common contaminant protein sequences using the Byonic search engine (Protein Metrics) with tolerance of 5 ppm for precursor ions and 10 ppm for product ions. Enzyme specificity was tryptic and allowed for up to two missed cleavages per peptide. Variable modifications were set for N-terminal acetylation or protein N termini, oxidation of methionine or tryptophan, deamidation of asparagine or glutamine, and dehydration of cysteine. A Wildcard search with a range of +75 to +80 Da facilitated confident peptide identification (<1% FDR) and spectrum counting of monoaminophosphate- or diaminophosphate-modified serine residues.

Crystallization and structural determination

The stability of purified AtzE in the phosphate buffer used in the size-exclusion chromatography was assessed using differential scanning fluorimetry as implemented in the Collaborative Crystallisation Centre (35). Protein was used at 2.3 mg/ml and showed a T_m of 64.6 ± 1.4 °C in the phosphate buffer, but it also showed that the protein was equally stable in all buffers tested with pH between 6.5 and 8.5, and the stability was essentially independent of salt concentration. An initial bank of 384 crystallization trials (PACT, shotgun and PS gradient at 20 °C, shotgun at 8 °C; see http://www.csiro.au/C3[4] for details of the screens) showed the formation of crystals overnight in many PEG-containing conditions. All of the crystals were small and

[4] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

rod-shaped, and all diffracted poorly and had a pathological packing problem, which made them unsuitable for diffraction analyses. Over the course of a year, over 10,000 droplets were set up, trying various combinations of seeding, *in situ* proteolysis, concentrations, differing formulations, vapor diffusion, microbatch, and temperatures. All crystals showed the same pathological form as the original crystals that grew overnight. Finally, it was found that the addition of 0.05% agarose to the protein solution before setup, coupled with seeding from the poor form crystals gave a new crystal form that showed diffraction to 2 Å, and did not have the same packing problems. The well-behaved crystal form appeared overnight and grew to full size in 5 days and grew at both 8 and 20 °C. The crystals used in the structural determination were grown from sitting drop trials with 250 nl of 1.1 mg/ml protein in 25 mM Hepes, pH 7.5, 100 mM NaCl, with 0.05% agarose, 250 nl of crystallant (0.1 M bis-tris, pH 6.04, 0.276 M $MgCl_2$, 17.6% w/v PEG 8000), incubated at 8 °C. Crystals of the protein treated with inhibitor PPDI grew under similar conditions to the native protein (reservoir 0.1 M bis-tris, pH 5.46, 0.128 M $MgCl_2$, 21% w/v PEG 8000) at 20 °C. A crystal of the protein–PPDI complex was cryoprotected with reservoir solution supplemented with glycerol to 20% final concentration and flash-cooled in liquid nitrogen. This crystal was used to collect 360° of data to 2.1 Å at beamline MX1 of the Australian Synchrotron. The data were initially processed with XDS (36), and the structure was solved by molecular replacement using MoRDa (37), which output two copies each of the A domains from 3ip4 and 3dha in the asymmetric unit. After some rebuilding, the new model was used in Phaser with reprocessed data (using Xia2 and DIALS) (38–42), and this was followed by manual building (COOT) and refinement with Refmac (43). Two complete chains of AtzE (residues 1–457) were located in the asymmetric unit. Additional protein density was clearly visible in the maps, and another protein chain was modeled into the density. The sequence was confirmed by searching the *Pseudomonas* database and finding that the AtzG sequence matched the sequence of the density (with 66 of 68 residues visible). Subunit interactions were analyzed using PISA (23).

A crystal of the native AtzE without inhibitor was harvested, and data were collected that extended to just beyond 2 Å at the MX2 beamline of the Australian Synchrotron. These data were processed using Xia2/DIALS, and the AtzE/PPDI model was used to phase the data with Phaser.

Computational methods

Models of AtzEG containing biuret and 1-carboxybiuret were prepared in Accelrys Discovery Studio version 4.1 using the experimentally derived structure as a starting point. Models for the substrates were relaxed using the Full Minimization tool in Discovery Studio version 4.1 using the default settings (CHARMm force field) and positioned in the receptor cavity (active site) using CDOCKER with the default parameters.

Ligands were prepared for MD using the Antechamber module in AMBER16 (44) using the GAFF2 force field. The protein models were prepared for MD simulations using xLeap applying the ff14SB force field and charge-neutralized by the addition of Na^+ ions. The proteins were solvated in a TIP3P-truncated octahedral solvent box with a minimum 12-Å periodic boundary distance from the solute.

Initial minimization of both systems was performed using AMBER 16 over 10,000 steps under a constant pressure of 1 bar (Berendsen barostat). Bonds lengths on bonds involving hydrogen were constrained using SHAKE, and force evaluation on these bonds was not performed. MD simulations of 500 ns with a step-size of 0.002 ps were performed at 310 K and 1 bar pressure with a 1-ps relaxation time. Long range electrostatic interactions were treated with the particle mesh Ewald method beyond 12 Å. Simulations were analyzed using VMD (version 1.9.2) (45).

Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) continuum solvation models were calculated using the MM-PBSA module in AMBER 16 over the course of the entire simulation. Entropy approximations were calculated with ptraj.

The ammonia tunnel in the GatCAB complex (PDB code 2G5H) and the equivalent tunnel in AtzEG were visualized using CAVER 3 (30) using the default parameters (clustering threshold of 3.5; probe radius of 0.9; approximation of 12).

Author contributions—L. E., T. S .P., M. W., J.-W. L., H. O., T. N., and J. N. formal analysis; L. E., T. S. P., M. W., J.-W. L., N. G. F., H. O., T. N., and J. N. investigation; L. E., T. S. P., J. N., and C. S. writing-original draft; L. E., T. S. P., C. J. H., C. J. E., J. N., and C. S. writing-review and editing; M. W., C. J. H., C. J. E., and C. S. supervision; J. N. data curation; C. S. conceptualization; C. S. funding acquisition; C. S. project administration.

Acknowledgments—We thank the beamline scientists at the Australian Synchrotron for their help in data collection. Crystals were grown in the CSIRO Collaborative Crystallisation Centre (http://www.csiro.au/C3)[4] at CSIRO Manufacturing. We also thank Drs. Andrew Warden and Thomas Walsh (CSIRO Land & Water) for their constructive comments while preparing this manuscript.

References

1. Cameron, S. M., Durchschein, K., Richman, J. E., Sadowsky, M. J., and Wackett, L. P. (2011) A new family of biuret hydrolases involved in *s*-triazine ring metabolism. *ACS Catal.* **2011,** 1075–1082 Medline
2. Udiković-Kolić, N., Scott, C., and Martin-Laurent, F. (2012) Evolution of atrazine-degrading capabilities in the environment. *Appl. Microbiol. Biotechnol.* **96,** 1175–1189 CrossRef Medline
3. Jutzi, K., Cook, A. M., and Hutter, R. (1982) The degradative pathway of the *s*-triazine melamine: the steps to ring cleavage. *Biochem. J.* **208,** 679–684 Medline
4. Mandelbaum, R. T., Allan, D. L., and Wackett, L. P. (1995) Isolation and characterization of a *Pseudomonas* sp that mineralizes the *s*-triazine herbicide atrazine. *Appl. Environ. Microbiol.* **61,** 1451–1457 Medline
5. Martinez, B., Tomkins, J., Wackett, L. P., Wing, R., and Sadowsky, M. J. (2001) Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. *J. Bacteriol.* **183,** 5684–5697 CrossRef Medline
6. de Souza, M. L., Wackett, L. P., Boundy-Mills, K. L., Mandelbaum, R. T., and Sadowsky, M. J. (1995) Cloning, characterization, and expression of a gene region from *Pseudomonas* sp. strain ADP involved in the dechlorination of atrazine. *Appl. Environ. Microbiol.* **61,** 3373–3378 Medline
7. Peat, T. S., Newman, J., Balotra, S., Lucent, D., Warden, A. C., and Scott, C. (2015) The structure of the hexameric atrazine chlorohydrolase AtzA. *Acta Crystallogr D Biol. Crystallogr.* **71,** 710–720 CrossRef Medline
8. Seffernick, J. L., Aleem, A., Osborne, J. P., Johnson, G., Sadowsky, M. J., and Wackett, L. P. (2007) Hydroxyatrazine *N*-ethylaminohydrolase (AtzB): an

amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *J. Bacteriol.* **189,** 6989–6997 CrossRef Medline
9. Balotra, S., Warden, A. C., Newman, J., Briggs, L. J., Scott, C., and Peat, T. S. (2015) X-ray structure and mutagenesis studies of the *N*-isopropylammelide isopropylaminohydrolase, AtzC. *PLoS ONE* **10,** e0137700 CrossRef Medline
10. Sadowsky, M. J., Tong, Z., de Souza, M., and Wackett, L. P. (1998) AtzC is a new member of the amidohydrolase protein superfamily and is homologous to other atrazine-metabolizing enzymes. *J. Bacteriol.* **180,** 152–158 Medline
11. Peat, T. S., Balotra, S., Wilding, M., French, N. G., Briggs, L. J., Panjikar, S., Cowieson, N., Newman, J., and Scott, C. (2013) Cyanuric acid hydrolase: evolutionary innovation by structural concatenation. *Mol. Microbiol.* **88,** 1149–1163 CrossRef Medline
12. Peat, T. S., Balotra, S., Wilding, M., Hartley, C. J., Newman, J., and Scott, C. (2017) High-resolution X-ray structures of two functionally distinct members of the cyclic amide hydrolase family of Toblerone fold enzymes. *Appl. Environ. Microbiol.* **83,** e03365-16 Medline
13. Balotra, S., Newman, J., Cowieson, N. P., French, N. G., Campbell, P. M., Briggs, L. J., Warden, A. C., Easton, C. J., Peat, T. S., and Scott, C. (2015) X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. *Appl. Environ. Microbiol.* **81,** 470–480 CrossRef Medline
14. Cheng, G., Shapir, N., Sadowsky, M. J., and Wackett, L. P. (2005) Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism. *Appl. Environ. Microbiol.* **71,** 4437–4445 CrossRef Medline
15. Shapir, N., Sadowsky, M. J., and Wackett, L. P. (2005) Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp. strain ADP. *J. Bacteriol.* **187,** 3731–3738 CrossRef Medline
16. Eaton, R. W., and Karns, J. S. (1991) Cloning and comparison of the DNA encoding ammelide aminohydrolase and cyanuric acid amidohydrolase from three *s*-triazine-degrading bacterial strains. *J. Bacteriol.* **173,** 1363–1366 CrossRef Medline
17. Seffernick, J. L., Erickson, J. S., Cameron, S. M., Cho, S., Dodge, A. G., Richman, J. E., Sadowsky, M. J., and Wackett, L. P. (2012) Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. *J. Bacteriol.* **194,** 4579–4588 CrossRef Medline
18. Balotra, S., Newman, J., French, N. G., Briggs, L. J., Peat, T. S., and Scott, C. (2014) Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp. strain ADP. *Acta Crystallogr F Struct. Biol. Commun.* **70,** 310–315 CrossRef Medline
19. Esquirol, L., Peat, T. S., Wilding, M., Lucent, D., French, N. G., Hartley, C. J., Newman, J., and Scott, C. (2018) Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841. *PLoS ONE* **13,** e0192736 CrossRef Medline
20. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* **28,** 235–242 Medline
21. Suzuki, T., Nakamura, A., Kato, K., Söll, D., Tanaka, I., Sheppard, K., and Yao, M. (2015) Structure of the *Pseudomonas aeruginosa* transamidosome reveals unique aspects of bacterial tRNA-dependent asparagine biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 382–387 CrossRef Medline
22. Nakamura, A., Yao, M., Chimnaronk, S., Sakai, N., and Tanaka, I. (2006) Ammonia channel couples glutaminase with transamidase reactions in GatCAB. *Science* **312,** 1954–1958 CrossRef Medline
23. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372,** 774–797 CrossRef Medline
24. Curnow, A. W., Hong, Kw., Yuan, R., Kim, Si., Martins, O., Winkler, W., Henkin, T. M., and Söll, D. (1997) Glu-tRNA(Gln) amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 11819–11826 CrossRef Medline
25. Skouloubris, S., Ribas de Pouplana, L., De Reuse, H., and Hendrickson, T. L. (2003) A noncognate aminoacyl-tRNA synthetase that may resolve a missing link in protein evolution. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 11297–11302 CrossRef Medline
26. Shin, S., Lee, T. H., Ha, N. C., Koo, H. M., Kim, S. Y., Lee, H. S., Kim, Y. S., and Oh, B. H. (2002) Structure of malonamidase E2 reveals a novel Ser-*cis*Ser-Lys catalytic triad in a new serine hydrolase fold that is prevalent in nature. *EMBO J.* **21,** 2509–2516 CrossRef Medline
27. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (1998) Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor. *Nature* **396,** 186–190 CrossRef Medline
28. Peat, T. S., Frank, E. G., McDonald, J. P., Levine, A. S., Woodgate, R., and Hendrickson, W. A. (1996) Structure of the UmuD′ protein and its regulation in response to DNA damage. *Nature* **380,** 727–730 CrossRef Medline
29. de Souza, M. L., Wackett, L. P., and Sadowsky, M. J. (1998) The atzABC genes encoding atrazine catabolism are located on a self-transmissible plasmid in *Pseudomonas* sp. strain ADP. *Appl. Environ. Microbiol.* **64,** 2323–2326 Medline
30. Pavelka, A., Sebestova, E., Kozlikova, B., Brezovsky, J., Sochor, J., and Damborsky, J. (2016) CAVER: Algorithms for analyzing dynamics of tunnels in macromolecules. *IEEE-ACM Trans. Comput. Biol. Bioinform.* **13,** 505–517 CrossRef Medline
31. Min, B., Pelaschier, J. T., Graham, D. E., Tumbula-Hansen, D., and Söll, D. (2002) Transfer RNA-dependent amino acid biosynthesis: an essential route to asparagine formation. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 2678–2683 CrossRef Medline
32. Baek, J., Lee, J., Yoon, K., and Lee, H. (2017) Identification of unannotated small genes in *Salmonella*. *G3* **7,** 983–989 CrossRef Medline
33. Hücker, S. M., Ardern, Z., Goldberg, T., Schafferhans, A., Bernhofer, M., Vestergaard, G., Nelson, C. W., Schloter, M., Rost, B., Scherer, S., and Neuhaus, K. (2017) Discovery of numerous novel small genes in the intergenic regions of the *Escherichia coli* O157:H7 Sakai genome. *PLoS ONE* **12,** e0184119 CrossRef Medline
34. Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., and Krijgsveld, J. (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol. Syst. Biol.* **10,** 757 CrossRef Medline
35. Rosa, N., Ristic, M., Seabrook, S. A., Lovell, D., Lucent, D., and Newman, J. (2015) Meltdown: a tool to help in the interpretation of thermal melt curves acquired by differential scanning fluorimetry. *J. Biomol. Screen.* **20,** 898–905 CrossRef Medline
36. Kabsch, W. (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66,** 125–132 CrossRef Medline
37. 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., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67,** 235–242 CrossRef Medline
38. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 CrossRef Medline
39. McCoy, A. J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D Biol. Crystallogr.* **63,** 32–41 CrossRef Medline
40. Parkhurst, J. M., Winter, G., Waterman, D. G., Fuentes-Montero, L., Gildea, R. J., Murshudov, G. N., and Evans, G. (2016) Robust background modelling in DIALS. *J. Appl. Crystallogr.* **49,** 1912–1921 CrossRef Medline
41. Waterman, D. G., Winter, G., Gildea, R. J., Parkhurst, J. M., Brewster, A. S., Sauter, N. K., and Evans, G. (2016) Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr D Struct. Biol.* **72,** 558–575 CrossRef Medline
42. Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* **43,** 186–190 CrossRef
43. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53,** 240–255 CrossRef Medline
44. Case, D. A., Cerutti, D. S., Cheatham, I., Darden, R. E., Duke, T. J., Giese T. J., Gohlke, H., Goetz, A. W., Greene, D., Homeyer, N., Izadi, S., Kovalenko, A., Lee, T. S., LeGrand, S., Li, P., *et al.* (2017) *AMBER16*, University of California, San Francisco, CA
45. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–38, 27–8 CrossRef Medline

An unexpected vestigial protein complex reveals the evolutionary origins of an *s*-triazine catabolic enzyme

Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Jian-Wei Liu, Nigel G. French, Carol J. Hartley, Hideki Onagi, Thomas Nebl, Christopher J. Easton, Janet Newman and Colin Scott

J. Biol. Chem. 2018, 293:7880-7891.
doi: 10.1074/jbc.RA118.001996 originally published online March 9, 2018

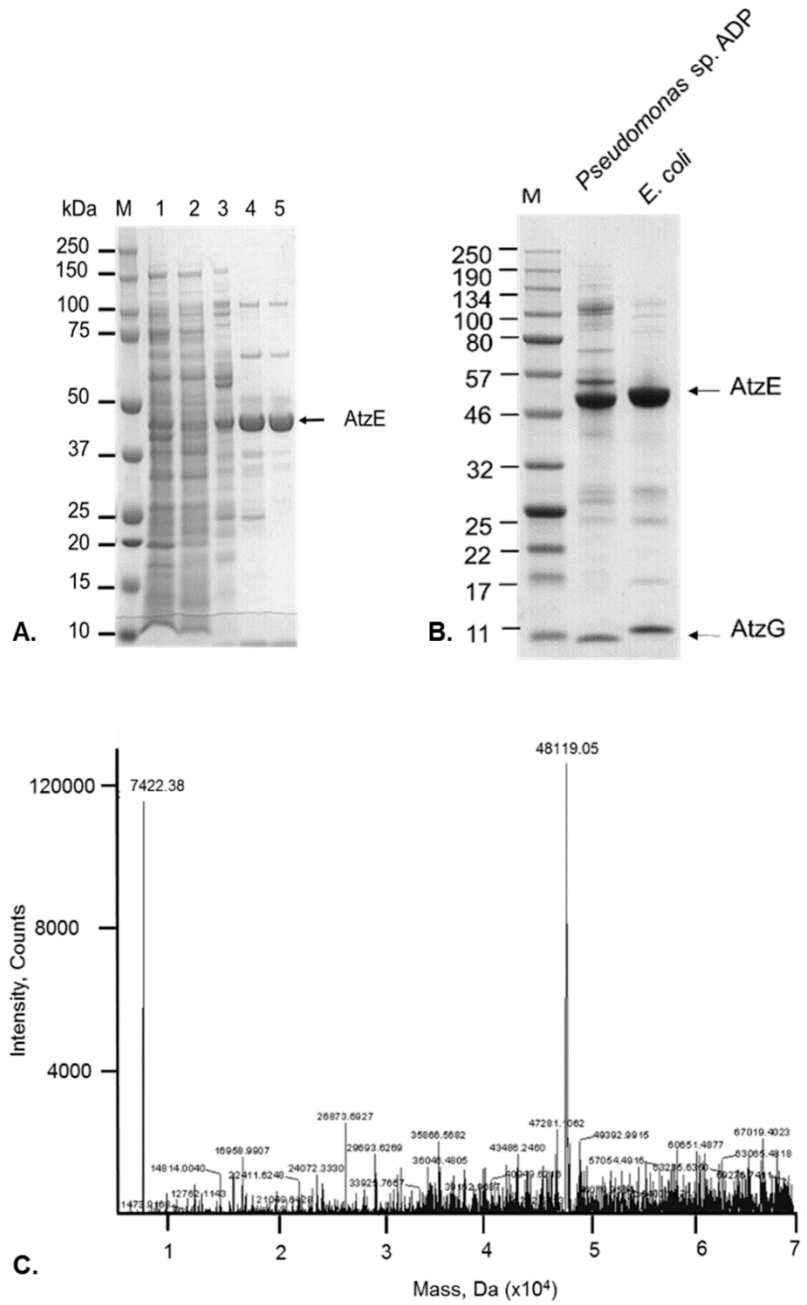


Figure S1: Purification of the AtzEG complex.

A) SDS-PAGE analysis of the five step purification of native AtzE from *Pseudomonas* sp. strain ADP: M, protein size marker; 1, ammonium sulfate precipitation; 2, hydrophobic interaction chromatography; 3, desalting; 4, anion exchange chromatography; and, 5, size exclusion.

B) SDS-PAGE analysis of AtzE prurified from *Pseudomonas* sp. strain ADP and from *E. coli* in which his-tagged AtzE and AtzG were copurified.

C) Deconvoluted mass spectrum of AtzE purified from *Pseudomonas* sp. strain ADP showing two major peaks of 48,119 Da and 7,422 Da, which is in agreement with the predicted molecular wieght of AtzE (48,120 Da) and AtzG (7,422 Da).

A.

MTETEIFAYIEAASIAIGIPLEPARARAVAHHFSRTALLAEMLESVPLSPESELAEIYRPAPFPAEDI

B.

MKTVEIIEGIASGRTSARDVCEEALATIGATDGLINAFTCRTVERARAEADAIDVRRARGEVLPPLAGLP
YAVKNLFDIEGVTTLAGSKINRTLPPARADAVLVQRLKAAGAVLLGGLNMDEFAYGFTTENTHYGPTRNP
HDTGRIAGGSSGGSGAAIAAGQVPLSLGSDTNGSIRVPASLCGVWGLKPTFGRLSRRGTYPFVHSIDHLG
PLADSVEGLALAYDAMQGPDPLDPGCSASRIQPSVPVLSQGIAGLRIGVLGGWFRDNAGPAARAAVDVAA
LTLGASEVVMWPDAEIGRAAAAFVITASEGGCLHLDDLRIRPQDFEPLSVDRFISGVLQPVAWYLRAQRFR
RVYRDKVNALFRDWDILIAPATPISAPAIGTEWIEVNGTRHPCRPAMGLLTQPVSFAGCPVVAAPTWPGE
NDGMPIGVQLIAAPWNESLCLRAGKVLQDTGIARLKC

Figure S2. Tryptic digest peptides identified by LCMS from AtzE purified from *Pseudomonas* sp. strain ADP. The primary sequence of AtzE (B) and the 68 amino acid long protein associated with AtzE (i.e., the newly identified AtzG; A) are shown. The tryptic peptides detected by LCMS that were used to identify the two proteins are underlined. A detailed report for the peptides detected is given in Table S3.

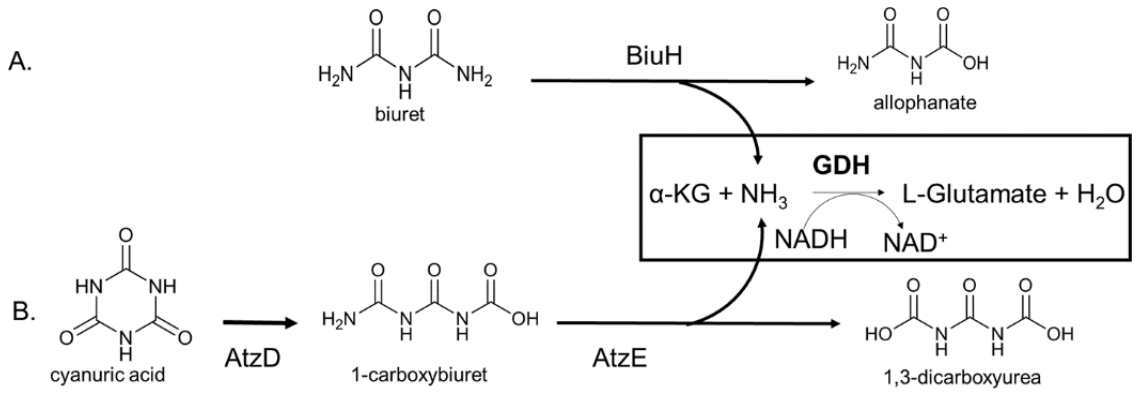


Figure S3: GDH-coupled assay. A glutamate dehydrogenase (GDH) coupled reaction was used to measure ammonium release in the biuret hydrolase (A) and AtzE (B) assays, following the decrease in absorbance at 340 nm, for each mole of ammonium produced, the GDH oxidizes one mole of NADH.

Table S1. Purification of AtzE from *Pseudomonas* sp. strain ADP. 20 L of culture were used to produce 2.4 mg of AtzE. (n=4). AS: ammonium sulfate, HIC: hydrophobic interaction column, AEC: anion exchange column.

	Total protein (mg)	**Rate of reaction[1] (nmole. min^{-1})**	**Total activity (nmole. $min^{-1}.mL^{-1}$)**	**Specific activity (nmole. $min^{-1}.mg^{-1}$)**	**Yield (%)**	**Purification (Fold)**
Soluble extract	9250	0.6 ± 0.35	147.5	3 ± 2	100	1
AS precipitation	8060	ND	ND	ND	ND	ND
HIC	36	ND	ND	ND	ND	ND
Desalt	50	1.9 ± 0.02	39.3	1557 ± 13	27	486
AEC	6	2.8 ± 0.2	9.8	3099 ± 198	7	968
Size Exclusion	2	9.9 ± 0.07	4.4	3654 ± 80	3	1142

ND: not determined.

[1]Rate of production of ammonia when incubated with purified AtzD and cyanuric acid.

4 A novel decarboxylating amidohydrolase involved in avoiding metabolic dead ends during cyanuric acid catabolism in *Pseudomonas* sp. strain ADP

4.1 Overview

This chapter describes the study of AtzH, a previously unknown small protein encoded by a gene located in the *Pseudomonas* sp. strain ADP's cyanuric acid degradation operon. The structural characterisation of AtzH determined it belonged to the versatile NTF2 protein superfamily. We used a combination of structural modelling and mutagenesis studies to provide evidence that AtzH is an allophanate forming, 1,3-dicarboxyurea amidohydrolase. Mutagenesis also indicated that Tyr22 and Arg46 may play an essential role in the catalysis of 1,3-dicarboxyurea.

Finally, a comparison of the genomic context of thirty different bacterial strains containing AtzH showed that the most frequent organisation is: *atzG-atzE-atzH*, suggesting that beyond its role in the cyanuric acid degradation operon, AtzH might be involved more broadly in the catabolism of nitrogenous compounds in Proteobacteria. Moreover, this observation also suggests that the *atzG-atzE-atzH* cluster predates the formation of the cyanuric acid catabolism operon.

4.2 Contribution

I performed the cloning, site-directed mutagenesis, protein purification and enzyme activity measurements. I also performed the AtzG-AtzE-AtgF gene cluster analysis in more than thrity different bacterial genomes. Docking was performed by Dr Matthew Wilding. The crystallography and structural characterisation were obtained by Dr Janet Newman and Dr Tom Peat respectively.

Status of the publication: Submitted for review 07/2018

Australian National University

This thesis is submitted as a Thesis by Compilation in accordance with https://policies.anu.edu.au/ppl/document/ANUP_003405

I declare that the research presented in this Thesis represents original work that I carried out during my candidature at the Australian National University, except for contributions to multi-author papers incorporated in the Thesis where my contributions are specified in this Statement of Contribution.

Title and authors:
Structural and genomic insights into AtzH, a protein dynamo for cyanuric acid catabolism in *Pseudomonas* sp. strain ADP. Esquirol L*, Peat TS*, Wilding M, Hartley CJ, Newman J, Scott C.

Current status of paper:
Submitted 05/2018

Contribution to paper:
I performed the cloning, site-directed mutagenesis, protein purification as well as enzyme activity measurements. Docking was performed by Dr Matthew Wilding. The crystallography and structural characterisation were obtained by Dr Janet Newman and Dr Tom Peat respectively.

Senior author's endorsement:
Lygie's assessment of her contribution to this work is accurate.

Dr. Colin Scott, senior author

LYGIE ESQUIROL
Candidate – Print Name

Signature _____ 25/05/2018 Date

Endorsed

Prof. C.J. Easton
Chair of Supervisory Panel – Print Name

Signature _____ 29 May 2018 Date

G. Otting
Delegated Authority – Print Name

Signature _____ 30 May 2018 Date

4.3 Publication: A novel decarboxylating amidohydrolase involved in avoiding metabolic dead ends during cyanuric acid catabolism in *Pseudomonas* sp. strain ADP

Erratum to this publication:

Page 3, Section Methods and material, Cloning and mutagenesis, Line 4:

… before being subcloned into the NdeI and SacI (not *Bam*HI as stated)…

Full details are given in the appendix on page 180.

PLOS ONE

RESEARCH ARTICLE

A novel decarboxylating amidohydrolase involved in avoiding metabolic dead ends during cyanuric acid catabolism in *Pseudomonas* sp. strain ADP

Lygie Esquirol[1,2☯], Thomas S. Peat[3☯], Matthew Wilding[2,3], Carol J. Hartley[1], Janet Newman[3], Colin Scott[1,4]*

1 Biocatalysis and Synthetic Biology Team, CSIRO Land & Water, Canberra, ACT, Australia, **2** Research School of Chemistry, Australian National University, Canberra, ACT, Australia, **3** CSIRO Biomedical Manufacturing, Parkville, Melbourne, VIC, Australia, **4** Synthetic Biology Future Science Platform, CSIRO Land & Water, Canberra, ACT, Australia

☯ These authors contributed equally to this work.
* colin.scott@csiro.au

Citation: Esquirol L, Peat TS, Wilding M, Hartley CJ, Newman J, Scott C (2018) A novel decarboxylating amidohydrolase involved in avoiding metabolic dead ends during cyanuric acid catabolism in *Pseudomonas* sp. strain ADP. PLoS ONE 13(11): e0206949. https://doi.org/10.1371/journal.pone.0206949

Editor: Monika Oberer, Karl-Franzens-Universitat Graz, AUSTRIA

Received: August 29, 2018

Accepted: October 21, 2018

Published: November 6, 2018

Data Availability Statement: New protein structures are lodged in the PDB with the following accession numbers: 6D63, 6BJT, 6BJU.

Funding: CSIRO (Commonwealth Scientific and Industrial Reasearch Organisation) is the sole funder of the work (paying the salaries of the authors and operating for the project), and that the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Cyanuric acid is a common environmental contaminant and a metabolic intermediate in the catabolism of *s*-triazine compounds, including atrazine and other herbicides. Cyanuric acid is catabolized *via* a number of bacterial pathways, including one first identified in *Pseudomonas* sp. strain ADP, which is encoded by a single, five-gene operon (*atzDGEHF*) found on a self-transmissible plasmid. The discovery of two of the five genes (*atzG* and *atzH*) was reported in 2018 and although the function of *atzG* was determined, the role of *atzH* was unclear. Here, we present the first *in vitro* reconstruction of the complete, five-protein cyanuric acid catabolism pathway, which indicates that AtzH may be an amidase responsible for converting 1,3-dicarboxyurea (the AtzE product) to allophanate (the AtzF substrate). We have solved the AtzH structure (a DUF3225 protein from the NTF2 superfamily) and used it to predict the substrate-binding pocket. Site-directed mutagenesis experiments suggest that two residues (Tyr22 and Arg46) are needed for catalysis. We also show that *atzH* homologs are commonly found in Proteobacteria associated with homologs of the *atzG* and *atzE* genes. The genetic context of these *atzG-atzE-atzH* clusters imply that they have a role in the catabolism of nitrogenous compounds. Moreover, their presence in many genomes in the absence of homologs of *atzD* and *atzF* suggests that the *atzG-atzE-atzH* cluster may pre-date the evolution of the cyanuric acid catabolism operon.

Introduction

The symmetrical triazine cyanuric acid (1,3,5-triazine-2,4,6-triol) is a common anthropogenic compound that is used in the synthesis of a variety of disinfectants (e.g., trichlorocyanuric acid) and herbicides (e.g., atrazine). It is catabolized and used as a nitrogen source by a number

Competing interests: The authors have declared that no competing interests exist.

of bacteria and can be mineralized to carbon dioxide and ammonia [1]. Cyanuric acid is also a common intermediate in the bacterial metabolism of other *s*-triazine compounds such as melamine, atrazine and simazine [1].

There are at least two pathways for the mineralization of cyanuric acid, a biuret-dependent pathway first described in *Rhizobium leguminosarum* bv. *viciae* 3841 (Fig 1A) [2,3], and a 1-carboxybiuret-dependent pathway observed in *Pseudomonas* sp. strain ADP (Fig 1B) [4,5]. The *Rhizobium leguminosarum* bv. *viciae* 3841 pathway employs cyanuric acid amidohydrolase (E.C. 3.5.2.15)[2,6,7] to ring-open cyanuric acid and the product (1-carboxybiuret) then undergoes solvent-mediated hydrolysis under physiological conditions to form biuret. Biuret is deaminated by biuret amidohydrolase (E.C. 3.5.1.84) to form allophanate and ammonia. The genes encoding the cyanuric acid hydrolase (E.C. 3.5.2.15) and biuret hydrolase (E.C. 3.5.1.84) are co-located on a large plasmid and may form an operon (although this has not been demonstrated). The fate of allophanate produced in this pathway remains unknown, although there are several genes in the *Rhizobium leguminosarum* bv. *viciae* 3841 genome that could degrade allophanate or its autohydrolysis product, urea [2,3].

The *Pseudomonas* sp. strain ADP cyanuric acid mineralization pathway is encoded by a single, five-gene operon: *atzDGEHF* [8]. The *atzG* and *atzH* genes were discovered recently and have been shown to encode proteins that are functionally expressed in *Pseudomonas* sp. strain ADP [8]. AtzG is non-catalytic and forms an obligate heterotetramer with the 1-carboxybiuret amidohydrolase, AtzE. In the absence of AtzG, AtzE is not soluble. The AtzEG complex is thought to have evolved from the GatCA component of the bacterial transamidosome (GatCAB) [8]. Like the *Rhizobium leguminosarum* bv. *viciae* 3841 pathway, *Pseudomonas* sp. strain ADP also ring opens cyanuric acid using cyanuric acid amidohydrolase (AtzD) [7,9]. In this pathway, it is 1-carboxybiuret that is deaminated by the next enzymatic step, which is catalyzed

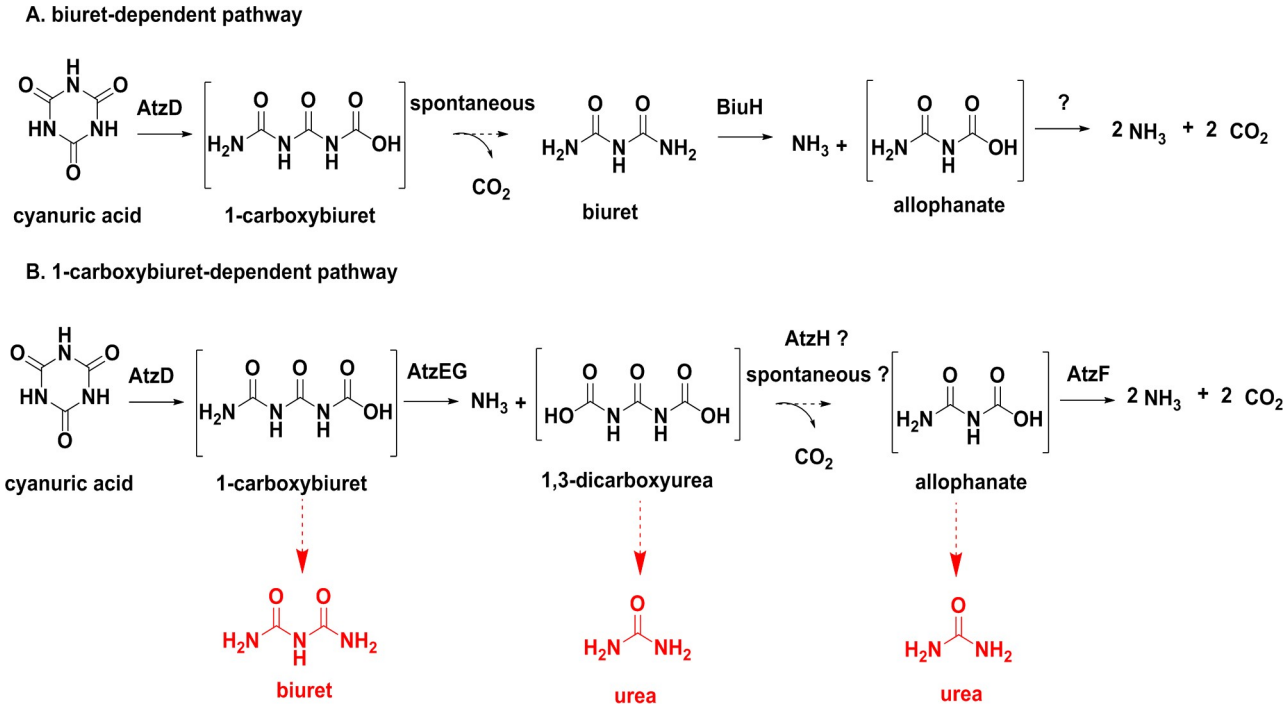


Fig 1. Bacterial cyanuric acid catabolism. A. The 1-carboxbiuret-dependent pathway first described in *Pseudomonas* sp. ADP; and, B. the biuret-dependent degradation pathway described in *Rhizobium leguminosarum* bv. *viciae* 3841. Compounds in square brackets are unstable under physiological conditions: the terminal amides hydrolyze to liberate CO_2.

https://doi.org/10.1371/journal.pone.0206949.g001

by 1-carboxybiuret amidohydrolase (AtzE) to form 1,3-dicarboxyurea [8]. 1,3-Dicarboxyurea must then undergo a single decarboxylation step to deliver allophanate to AtzF, which is again deaminated to ultimately produce ammonia and carbon dioxide [10–13].

Several of the metabolic intermediates in the 1-carboxybiuret-dependent pathway are unstable under physiological conditions: 1-carboxybiuret forms biuret and CO_2, 1,3-dicarboxyurea and allophanate both form urea and CO_2, and dicarboxyammonia (from allophanate deamination) forms ammonia and CO_2. *Pseudomonas* sp. strain ADP is unable to degrade biuret or urea, and these solvent-mediated reactions effectively form metabolic 'dead-ends' [2,10]. Presumably, the enzyme-catalyzed reactions have rates that are sufficient to avoid this outcome, effectively using the enzyme-mediated rate enhancement to control which hydrolyses occur. In this context, it is perhaps counter-intuitive that the conversion of 1,3-dicarboxyurea to allophanate would be uncontrolled, which would potentially lead to urea accumulation *via* the solvent-mediated process; however, to date no enzyme has been identified that catalyzes this step [8].

Previously, proteomic analysis had demonstrated that AtzH is expressed from the cyanuric acid catabolism operon. However, its function had not been explored in previous studies [8]. Herein, we present the structure of AtzH and explore its potential function.

Methods and materials

Cloning and mutagenesis

The *AtzH* gene, coding for the protein WP_006390399.1 (KSW21442.1), was cloned directly from *Pseudomonas* sp. strain ADP genomic DNA with appropriate primers (5' to 3': `CGACGA CATATGCTCGAGATGCAAATTAATCTACC` and `TGCTGCGAGCTCCCTAGGTCAGGAAACG GGC`), before being subcloned into the *Nde*I and *Bam*HI sites of the pETcc2 [9], which enabled production of AtzH with an N-terminal hexa-his-tag with thrombin cleavage site (S1 Fig). Cloning of the other genes used in this work has been described elsewhere [9,12,13].

The AtzF Ser189Ala variant and fourteen variants of AtzH (Tyr22Ala, Tyr22Phe, Arg46Ala, Arg46Lys, Arg63Ala, Arg63Lys, Arg66Ala, Arg66Lys, Arg73Ala, Arg73Lys, Arg96Ala, Arg96Lys, Gln106Ala, Gln108Ala) were produced using the QuikChange Lightning Multi Site-Directed mutagenesis kit from Agilent Technologies (Santa Clara, CA) with the mutagenic primers listed in S1 Table.

Heterologous protein expression and purification

The expression vectors were used to transform *Escherichia coli* BL21 (λDE3) cells (Invitrogen). Bacteria were grown on Luria-Bertani (LB) medium containing ampicillin 100 μg/mL for the pETcc2 constructs or chloramphenicol 34 μg/mL for the pACYCDuet-1 construct (used to co-express AtzE and AtzG). Cells were grown with shaking at 200 rpm at 28˚C. Protein expression was induced at an OD_{600} of 0.8 by addition of isopropyl-β–D-1-thiogalactopyranoside (IPTG; 1 mM final concentration).

Cells were harvested 24 hours after induction by centrifugation at 5000 x *g* for 15 minutes using an Aventi J-E centrifuge (Beckman Coulter, Indianapolis, USA), resuspended in lysis buffer (25 mM potassium phosphate, 5 mM imidazole, pH 7.5) and lysed by passage through a Microfluidics homogenizer M-110P (Massachusetts, USA) five times at 15000 PSI. The lysis was followed by centrifugation at 18000 x *g* for 45 minutes to pellet the cellular debris, and the soluble fraction was used for further purification.

The soluble fraction was syringe filtered through a 0.22 μm filter (Merck Millipore, USA). The His_6-tagged proteins were isolated from the filtrate using a 5 mL Ni-NTA HisTrap HP

(Healthcare Life Sciences) with an imidazole gradient from 5 mM to 500 mM over 10 column volumes (CV). SDS-PAGE gel analysis was performed to assess the purity of the fractions.

Fractions containing AtzH were pooled and concentrated to 8 mg/mL for crystallography using an Amicon Ultra-15 centrifugal filter unit. Finally, size exclusion chromatography was performed using a 130 mL column packed with Superdex 76 preparation grade resin (GE Healthcare Life Sciences), equilibrated with 25 mM HEPES and 200 mM NaCl, pH 7.5, over 1.5 CV.

All chromatography steps were performed using an ÄKTA purifier UPC 10 (GE Healthcare Life Sciences).

Differential Scanning Fluorimetry (DSF)

Initially, buffer screening by DSF was performed on the wild-type protein–thirteen different combinations of salt/buffer at different pHs were tested in triplicate (Buffer screen 9 protocol) [14]. Protein was tested at 8 mg/mL in gel filtration buffer; 300 nL protein was diluted into 19.4 μL of each buffer, and 300 nL of a 1:20 dilution of sypro orange (Sigma S5692) was added. The temperature was increased from 20 to 100˚C in increments of 0.5˚/ 5 sec. This gave a T_m of 55.8 +/- 0.1˚C in the gel filtration HEPES buffer. A slight improvement was seen when the protein was in 50 mM Tris pH 8, 200 mM NaCl (T_m of 57.1 +/- 0.1˚C). The fourteen mutants were also tested for stability using DSF. Each sample was diluted to 1 mg/mL in 50 mM HEPES pH 7.5, 200 mM NaCl, and 1 μL of protein was added to 19 μL of mix of 4 mL 50 mM HEPES pH 7.5, 200 mM NaCl, 2.5 μL Sypro Dye. Each variant was run in 6-fold replicates, and compared to the wild-type protein which was included as a control (under these conditions the wild type protein had a T_m of 55.0 +/- 0.1˚C). A complete table of the results is shown in S2 Table.

Crystallography

Purified AtzH at 8 mg/mL protein was used to set up a PCT test (Pre-Crystallization Test [15]), and immediately showed microcrystal formation in the PEG based condition of the PCT. The protein was diluted to 4 mg/mL and then set up in 3 initial PEG-rich screens (shotgun, PACT and C3_3; conditions available from c6.csiro.au [16]), in SD2 sitting drop plates (Molecular Dimensions, UK) with an equal volume of crystallant (150 nL protein plus 150 nL reservoir) at 20˚C. Crystals appeared overnight in many conditions. Further PEG based screens (C3_1->C3_4, Morpheus_C3) were also set up, with the protein at 4 mg/mL with and without added thrombin. Crystals continued to appear in many PEG-based conditions for up to 3 weeks after setting up the plates. The AtzH wild-type protein and the variants were also set up in co-crystallization trials with the PPDi inhibitor. 1 μL of a 40 mM solution of the inhibitor in protein buffer (50 mM HEPES, 200 mM NaCl pH 7.5) was added to 40 μL of the 4 mg/mL protein. The variants were each set up in two screens (Shotgun, C3_3) at 4 mg/mL (if accessible) with and without added inhibitor and incubated at 20˚C. The mutants in general crystallized poorly, and only a few variants gave crystals suitable for harvesting. Crystals were cryoprotected with addition of glycerol to a final concentration of 20% and cryo-cooled in liquid nitrogen prior to data collection. Crystals used in data collection were harvested directly from the screening plates.

The native AtzH P1 crystal form crystal was harvested from a 300 nL drop consisting of a 1:1 ratio of protein and reservoir, where the reservoir consisted of 200 mM ammonium acetate, 30% PEG 4000 and 100 mM sodium acetate buffer pH 5.0. The native $C222_1$ crystal form was harvested from a 300 nL drop consisting of a 1:1 ratio of protein and reservoir, where the

reservoir contained 20% PEG 6000, 2.5% v/v tert-butanol with 100 mM sodium citrate buffer at pH 5.5.

The R73K mutant (which crystallized in space group $P2_1$) was harvested from a 300 nL drop consisting of a 1:1 ratio of protein/inhibitor and reservoir, where the reservoir contained 0.2 M diammonium tartrate, 20% w/v PEG 3350. Data sets were obtained at the Australian Synchrotron beamline MX2 using a Dectris Eiger 16M detector and a full 360 degrees of data were taken. The data were indexed using DIALS [17–21], scaled using Aimless [22] and the original structure solution was determined by molecular replacement using Phaser [20,21] with PDB code 2RCD as the model, after sequence alignment with Clustal-Omega and pruning using Chainsaw [23,24] to give an appropriate starting model. The model was rebuilt manually using Coot [25] and refined using Refmac [24,26].

In silico substrate docking was done using Accelrys Discovery Studio (v4.1). The Define and Edit Binding Site program was used to identify cavities in AtzH. Models of 1,3-dicarboxyurea were prepared using the same client in doubly protonated, deprotonated and singly protonated states and docked into the cavity identified previously using the CDOCKER program with default parameters.

Enzyme assays

Reaction rates were monitored by detecting the production of ammonium in μmoles per second using the glutamate dehydrogenase GDH-coupled assay, described previously [3]. GDH catalyzes the NADH-dependent amination of alpha-ketoglutarate. Ammonia production was followed using the decrease of absorbance by UV spectrophotometry at 340 nm, due to the oxidation of NADH by GDH. 1.25 U of GDH was used in a 250 μL reaction volume, the final concentrations of alpha-ketoglutarate and NADH were 3.5 mM and 0.2 mM, respectively. The buffer was 25 mM potassium phosphate, pH 8.5. Cyanuric acid hydrolysis by AtzD was followed by UV spectrophotometry at 214 nm [13]. In all the reaction mixes, 12 μM of AtzD, 0.17 μM of AtzE, 0.17 μM AtzF and 3.3 μM of AtzH or its variants were used in presence of 1 mM of cyanuric acid.

Results and discussion

Structural characterization of the *Pseudomonas* sp. strain ADP AtzH

We obtained the X-ray crystal structure of the AtzH. His_6-tagged AtzH was produced in *E. coli* and purified, yielding 5 mg of AtzH per 1 L of culture. Size exclusion chromatography performed on AtzH after removal of the his-tag, suggested that the native protein had a molecular mass of ~ 29 kDa, consistent with a dimer (2 x 14,655 kDa) (S2 Fig). Differential Scanning Fluorimetry (DSF) was used to estimate the stability of the protein and showed that the native protein had a T_m of 57.2˚C. Well diffracting crystals of AtzH grew readily, and two different space groups were observed: P1 and $C222_1$. X-ray diffraction data were collected for both.

The structure of AtzH was solved by molecular replacement, using the structure of a functionally uncharacterized protein from *Pectobacterium atrosepticum* strain SCRI 1043 (PDB accession number: 2RCD) [27] as the model. 2RCD is 56% identical to AtzH and there is a 1.0 Å rmsd between Cα atoms when superposing the structures (x-ray data are in Table 1). The structures exhibit minor differences at the N- and C-termini and in some loop regions (including the loop with weak density in AtzH, residues 66 to 73). Additionally, 2RCD has a dimer in the asymmetric unit, which mirrors the dimer seen in AtzH where two, six-stranded β-sheets sit face-on at an angle to form the interface. One long helix and two short helices sit on the outside of this interface and the overall dimer structure is very compact (Fig 2). Analysis with

Table 1. Crystal structure parameters.

	6BJU Wild-type	6C6G Wild-type	6D63 Arg73Lys
Data collection			
Space group	P1	C2221	P21
Cell dimensions			
a, *b*, *c* (Å)	51.8, 51.8, 59.7	71.5, 75.0, 96.2	99.7, 74.7, 103.6
α, β, γ (˚)	99.8, 102.2, 92.5	90, 90, 90	90, 96.7, 90
Resolution (Å)	1.64 (1.67–1.64)	1.80 (1.84–1.80)	2.30 (2.35–2.30)
R_{merge}	0.087 (0.674)	0.159 (1.581)	0.267 (1.402)
R_{pim}	0.054 (0.412)	0.045 (0.438)	0.109 (0.561)
I / σI	7.8 (1.9)	11.2 (2.2)	6.6 (1.6)
CC1/2	0.995 (0.733)	0.998 (0.837)	0.986 (0.761)
Completeness (%)	97.2 (95.3)	100 (100)	100 (100)
Redundancy	3.6 (3.7)	13.3 (13.7)	6.9 (7.1)
Refinement	2 dimers	1 dimer	6 dimers
Resolution (Å)	50.5–1.64	37.5–1.80	50.1–2.30
Unique reflections	67,428	23,104	64,253
Rwork / Rfree	19.9 / 22.8	16.2 / 20.6	24.9 / 28.7
No. atoms	4538	2307	12,009
Protein	4029	2110	11,791
Inhibitor	n/a	n/a	30
Water	509	178	288
B-factors (Å2)	21.6	26.2	25.3
Protein	21.7	26.6	26.4
Inhibitor	n/a	n/a	40.9
Water	31.4	31.1	17.2
R.m.s deviations			
Bond lengths (Å)	0.020	0.017	0.011
Bond angles (˚)	1.856	1.677	1.363

Note: Values in parenthesis are for the highest resolution shell.

https://doi.org/10.1371/journal.pone.0206949.t001

PISA [28] gives an interfacial surface area of almost 3,200 Å^2 (or about 29% of the total surface area of the protein).

AtzH belongs to NTF2-fold superfamily, which contains 38 protein families [29], including the AtzH family: DUF3225 (DUF: domain of unknown function). Proteins of the NTF2 superfamily fulfil a number of functions, including enzymatic functions such as limonene-1,2-epoxide hydrolase, scytalone dehydratase, δ5-3-ketosteroid isomerase and polyketide cyclase [30–33] and non-enzymatic roles including nuclear transport and bacterial secretion [34–39]. Among the 170 protein structures belonging to NTF2-fold superfamily available on the PBD database, 24 are assigned as DUF3225 [40]; no DUF3225 protein had been functionally characterized prior to this study.

A physiological role for AtzH

We investigated the effect of AtzH on the cyanuric acid catabolic pathway by reconstituting the complete cyanuric acid mineralization pathway *in vitro* using purified enzymes; i.e., AtzD, AtzEG and AtzF with and without AtzH. Inclusion of AtzH led to an increased rate of ammonia production from cyanuric acid by the pathway (Fig 3C). This is consistent with an

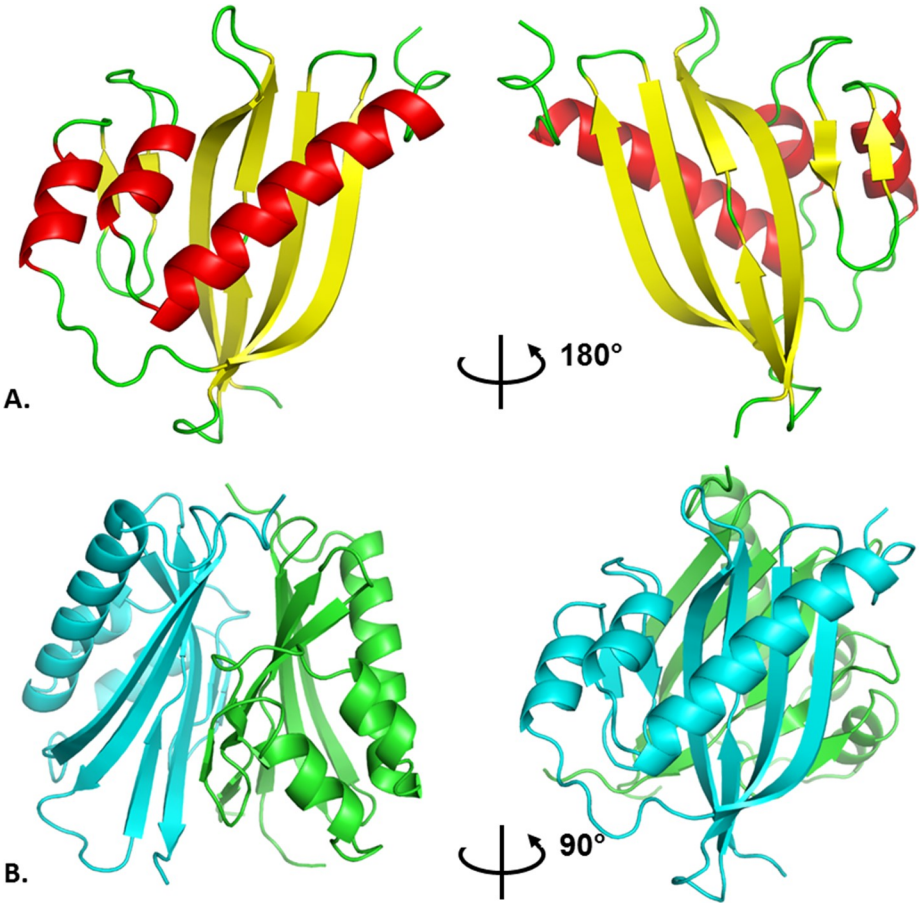


Fig 2. The AtzH structure. A) The AtzH monomer with the seven-stranded anti-parallel β-sheet highlighted in yellow and the three helices are shown in red; B) The AtzH dimer with the two protomers interacting through the β-sheet secondary structure with the helices sitting on the outside. The two anti-parallel β-sheets sit at an approximate 30 degree angle to each other.

https://doi.org/10.1371/journal.pone.0206949.g002

enzymatic role for AtzH, potentially fulfilling the role of a decarboxylating 1,3-dicarboxyurea amidohydrolase. Unfortunately, both the substrate and product for the putative AtzH-mediated reaction are unstable under the conditions required for enzyme function (i.e., aqueous buffer at physiological pH) and so the reaction could not be followed directly. However, it was possible to probe the role of AtzH by varying the composition of the *in vitro* pathway.

The AtzD substrate (cyanuric acid) is stable and its hydrolysis easily followed by measuring the reduction in absorbance at 214 nm caused by AtzD-mediated ring-opening [9]. It was therefore possible to directly test the influence of AtzH on this step of the pathway by incubating AtzD and its substrate in the presence or absence of AtzH, and AtzH did not influence the rate of this step (Fig 3A). The influence of AtzH on the AtzEG-catalyzed reaction was also investigated (Fig 3B). The substrate for AtzEG (1-carboxybiuret) is unstable and was therefore produced *in situ via* AtzD-mediated ring opening of cyanuric acid. One product of the AtzEG reaction is ammonia, which was used to determine the rate of the reaction through a GDH-coupled assay. *In vitro* cascades containing AtzD and AtzEG evolved ammonia from cyanuric acid at the same rate whether or not AtzH was included in the incubation (Fig 3B).

These observations suggest that the AtzH-dependent increase in the rate of ammonia production from the pathway was also AtzF-dependent; i.e., AtzH likely increases the rate at

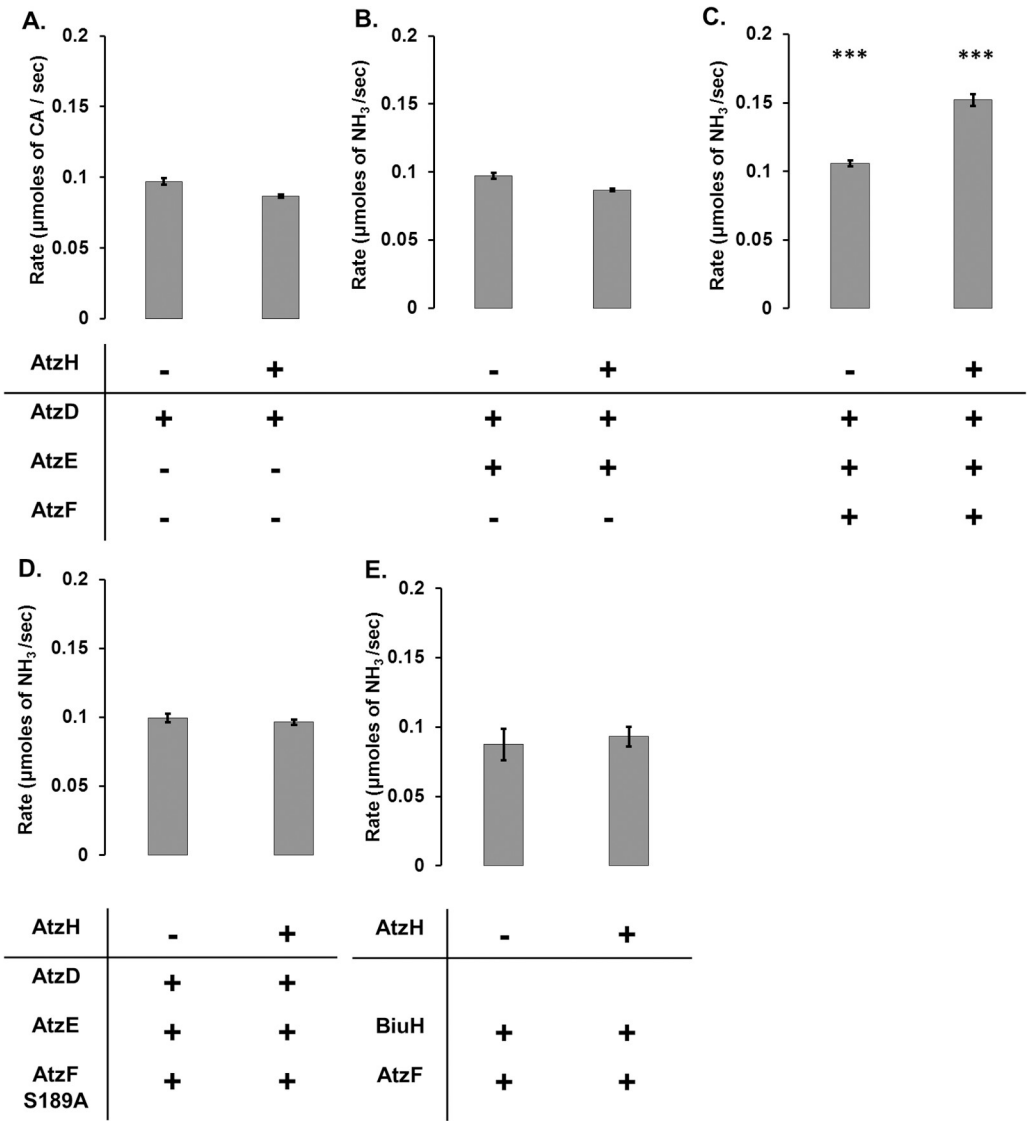


Fig 3. Impact of AtzH on the rate of the cyanuric acid pathway *in vitro*. A) Rate of cyanuric acid (CA) degradation in presence of AtzD with/without AtzH; B) Rate of ammonia production when CA is put in presence of AtzD and AtzEG with/without AtzH; C) Rate of ammonia production when CA is put in presence of AtzD, AtzEG and AtzF with/without AtzH; D) Rate of ammonia production when CA is put in presence of AtzD, AtzEG and an inactive variant of AtzF Ser189Ala with/without AtzH; E) Rate of ammonia production when biuret is put in presence of BiuH and AtzF with/without AtzH. *** indicates a significant difference between the rate measured in the presence of AtzH and the rate measured without the AtzH (pval < 0.005, measured by Student's t-test; n = 3).

https://doi.org/10.1371/journal.pone.0206949.g003

which AtzF produces ammonia, possibly by increasing the rate at which allophanate is fed to AtzF. To further investigate this, an *in vitro* pathway comprised of AtzD, AtzEG and a variant of AtzF that lacked the active-site nucleophile (AtzF Ser189Ala) was also tested. In this reaction, the presence of AtzH also failed to improve the rate of ammonia production (Fig 3D), confirming that the AtzH-mediated effect observed in the fully reconstituted pathway is dependent on the presence of active AtzF. These observations imply that AtzH is either increasing the rate of allophanate production from 1,3-dicarboxyurea, or enhancing the activity of AtzF itself.

To test whether AtzH increases the catalytic activity of AtzF, an alternative *in vitro* pathway was tested in which biuret was incubated with the biuret amidohydrolase BiuH [2] and AtzF. BiuH produces allophanate directly, and so an increase in the rate of ammonia production in this pathway would indicate that AtzH is an allosteric activator of AtzF. However, the presence of AtzH made no measurable difference to the rate of ammonia production (Fig 3E), suggesting that AtzH does not 'activate' AtzF. The simplest explanation for these observations is that AtzH mediates the decarboxylation of 1,3-dicarboxyurea, the product from AtzEG, to form allophanate, the substrate for AtzF.

Defining the active site of AtzH

The AtzH active site was sought by examining the X-ray structure. As 1,3-dicarboxyurea is unstable, it was not possible to produce crystals of AtzH in its presence. However, an *in silico* docking simulation between the X-ray crystal structure of the AtzH dimer and 1,3-dicarboxyurea elucidated a potential active site (Fig 4A and 4B). The putative substrate bound with good complementarity to an arginine-rich pocket comprised of Tyr22, Arg46, Arg63, Arg66, Arg73, Phe94, Arg96, Gln106 and Gln108 (Fig 4A and 4B). The arginine residues form an extensive hydrogen bonding network with the carbonyl oxygens of 1,3-dicarboxyurea. Tyr22 is positioned within 3 Å of the carbonyl carbon of one of the two terminal carboxylic acids, and Arg46 is within hydrogen-bonding distance of the hydroxyl group of Tyr22. Tyr22 and Arg46 are therefore well positioned to form a catalytic dyad, albeit one of atypical composition.

Guided by the results of the docking experiment, fourteen variants of AtzH were produced: Tyr22Ala, Tyr22Phe, Arg46Ala, Arg46Lys, Arg63Ala, Arg63Lys, Arg66Ala, Arg66Lys, Arg73Ala, Arg73Lys, Arg96Ala, Arg96Lys, Gln106Ala, Gln108Ala (S1 Table). All of the variants were correctly folded, as assessed by DSF (S2 Table). Furthermore, X-ray structures of the R73K variant (PDB: 6D63) was also obtained and closely resembled the wild-type enzyme (PDB: 6BJT, 6BJU) except for the intended amino acid substitution. The AtzH variants were added to *in vitro* pathways comprised of AtzD, AtzEG and AtzF and incubated with cyanuric acid. Five of the variants (Tyr22Ala, Tyr22Phe, Arg46A, Arg46Lys and Arg63Lys) were found to enhance the rate of ammonia production, albeit to a significantly lower extent than the wild-type AtzH (Fig 5). The loss of function of Arg63Lys, but not of Arg63Ala, may suggest that Arg63 does not have a catalytic role. However, AtzH Tyr22Ala/Phe was indistinguishable from the AtzH-free cascade (Fig 5). The Arg46Ala and Arg46Lys variants both increased the rate of ammonia production by the cascade, but to a lesser degree than that of the wild-type AtzH (Fig 5). This may suggest that Tyr22 is the catalytic amino acid residue in AtzH, possibly acting as a proton donor (consistent with the catalytic role of tyrosine in other enzymes) [41], with Arg46 fulfilling a non-critical function As the substrate analog binds between Arg46 and Tyr22, it seems unlikely that Arg46 has a role in activating Tyr22 (noting that the substrate analog may not adopt exactly the same conformation as the bound substrate). However, this positioning of the active site residues relative to the bound analog is similar to that of the catalytic residues in inverting β-glycosidases that use an acid-base catalytic mechanism [42], and it is plausible that that a similar mechanism may be used here. Further work is needed to test this hypothesis.

The wild-type and variants were co-crystallized with 1,3-dicarboxyacetone, a stable structural analogue of 1,3-dicarboxyurea. In the AtzH Arg73Lys structure (PDB 6D63) 1,3-dicarboxyacetone was found to occupy the predicted active site of AtzH in a subset of the protomers (25%), indicating that this pocket is accessible by 1,3-dicarboxyurea-like small molecules (Fig 4C and 4D). The hydrogen bonding network is largely as predicted for 1,3-dicarboxyurea, and Tyr22 is well positioned as the catalytic residue.

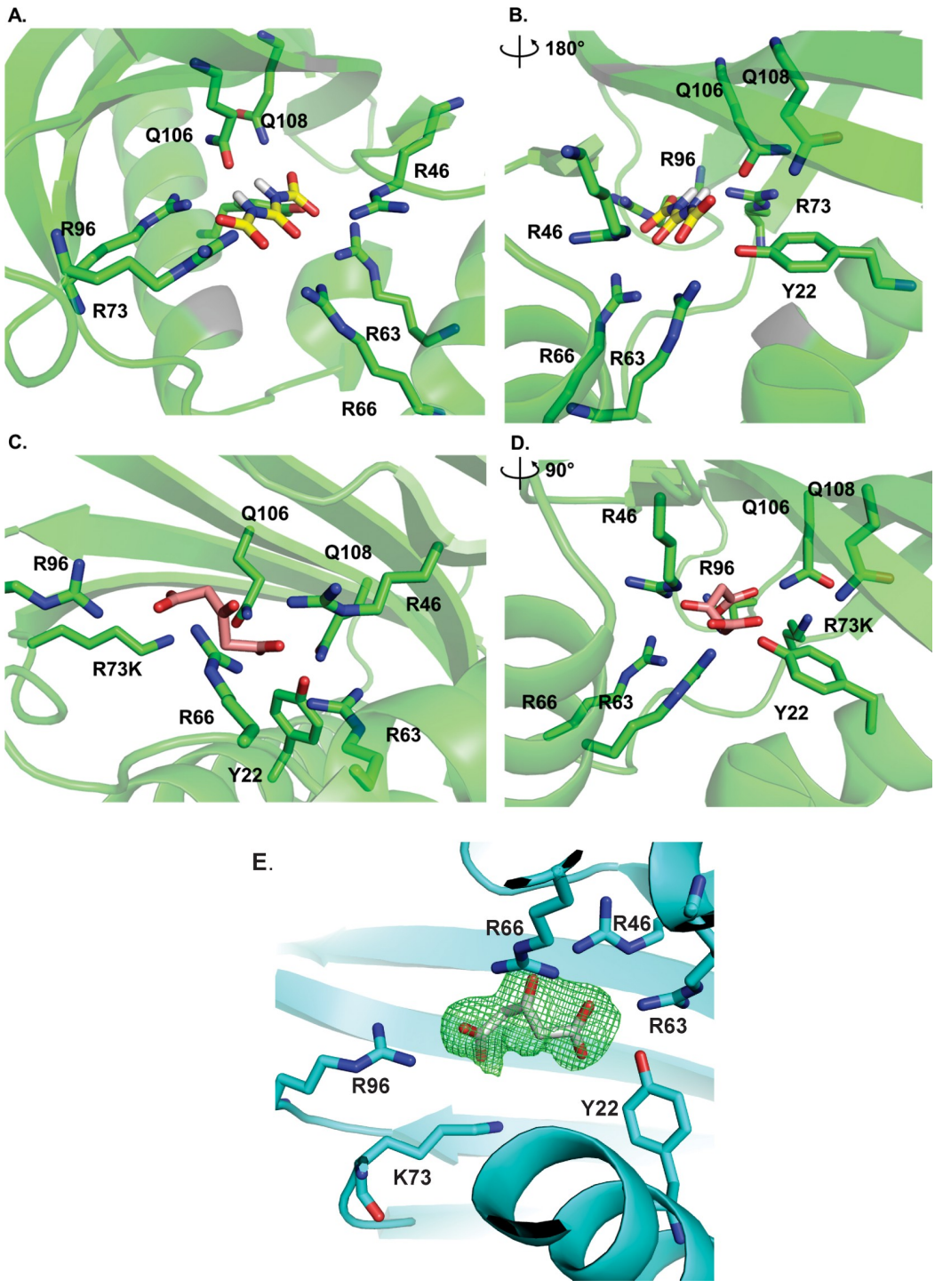


Fig 4. Substrate and substrate analog binding by AtzH. A-B) Simulated binding of 1,3-dicarboxyurea (yellow) by AtzH. Amino acids residues involved in binding 1,3-dicarboxyurea are shown as green sticks; C-D) X-ray structure of AtzH Arg73Lys co-crystalized with the substrate analog 1,3 dicarboxyacetone (pink). The distance between Tyr22 and the carbon of the carboxyl group is <3 Å; E) Polder map (3 sigma density contour) of 1,3 dicarboxyacetone bound to the AtzH Arg73Lys variant.

https://doi.org/10.1371/journal.pone.0206949.g004

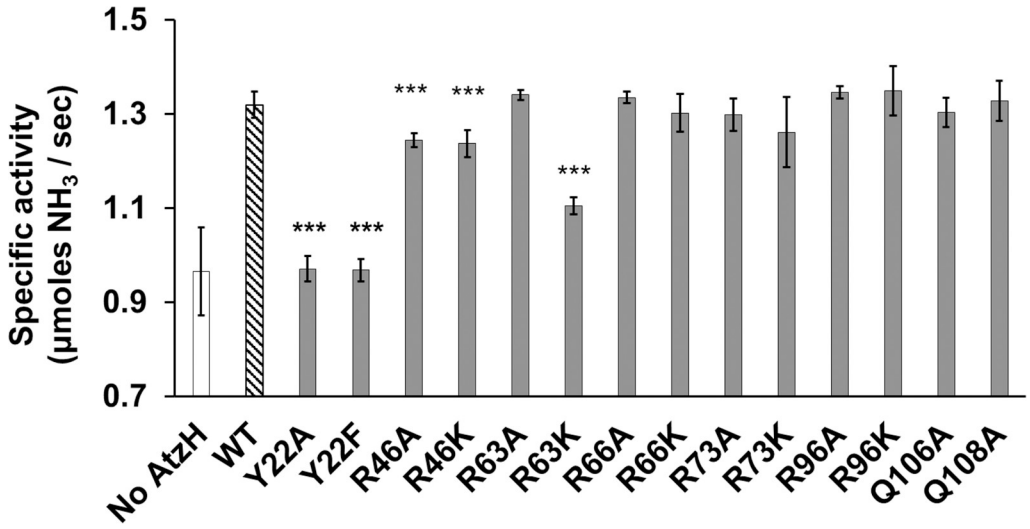


Fig 5. Impact of the presence of AtzH variants on the ammonium production rate of the *in vitro* cyanuric acid catabolism pathway. The rates of ammonia production by the cyanuric acid catabolism pathway (AtzD, AtzEG and AtzF) reconstituted *in vitro* with no AtzH (white), wild-type AtzH (striped) or an AtzH variant (gray) are shown. *** indicates an ammonia production rate that is significantly different from a reaction in the presence of wild-type AtzH (pval < 0.005, measured by Student's t-test; n = 3).

https://doi.org/10.1371/journal.pone.0206949.g005

AtzH homologs in other bacterial genomes

AtzH is the first DUF3225 family protein that has been shown to possess a catabolic function, likely decarboxylating the product of AtzEG-mediated deamination of 1-carboxybiuret as part of the cyanuric acid mineralization pathway. Homologs of the other members of the pathway are found in different metabolic contexts: the AtzD homolog barbiturase participates in pyrimidine catabolism [43], the AtzEG homologs GatCA are involved with acyl-tRNA metabolism [44] and the AtzF homolog allophanate hydrolase is used in urea catabolism [45]. We were therefore interested to investigate whether AtzH homologs may also participate in alternative metabolic contexts.

A thorough search of microbial genomes in the GenBank database revealed more than thirty examples of genomes containing homologs of *atzG*, *atzE*, and *atzH* organized as a single gene cluster with the same organization as that of the *Pseudomonas* sp. strain ADP cyanuric acid catabolism operon (Fig 6). Perhaps surprisingly, the *atzG-atzE-atzH* clusters were rarely associated with *atzD* or *atzF* homologs.

The most common arrangement, found in species of Enterobacteriaceae (*Pectobacterium*, *Serratia*, *Enterobacter*, *Klebsiella*, *Kosakonia*, *Pantoea*, *Yersinia* and *Erwinia*), has the *atzG-atzE-atzH* cluster immediately downstream of a gene predicted to encode γ-glutamyl transferase (*ggt*, E.C. 2.3.2.2). In these genomes, the *ggt-atzG-atzE-atzH* cluster is found downstream and on the opposite strand from a gene cluster encoding a LysR-family regulator, an ABC transporter, a transaminase and an allantoate amidohydrolase. The overall genetic organization of these clusters is similar to that of the cyanuric acid catabolism operon of *Pseudomonas* sp. strain ADP, albeit lacking *atzF* and with *atzD* replaced by *ggt*. (Fig 6B, S3 Fig)

γ-Glutamyl transferase, which also acts as glutathione hydrolase (E.C. 3.4.19.13), is responsible for the amide hydrolysis of glutathione to produce L-glutamate and L-cysteinylglycine (Fig 7). Further, the substrate for allantoate hydrolase, an enzyme found encoded in close proximity to the Enterobacteriaceae *atzG-atzE-atzH* clusters described here, contains two

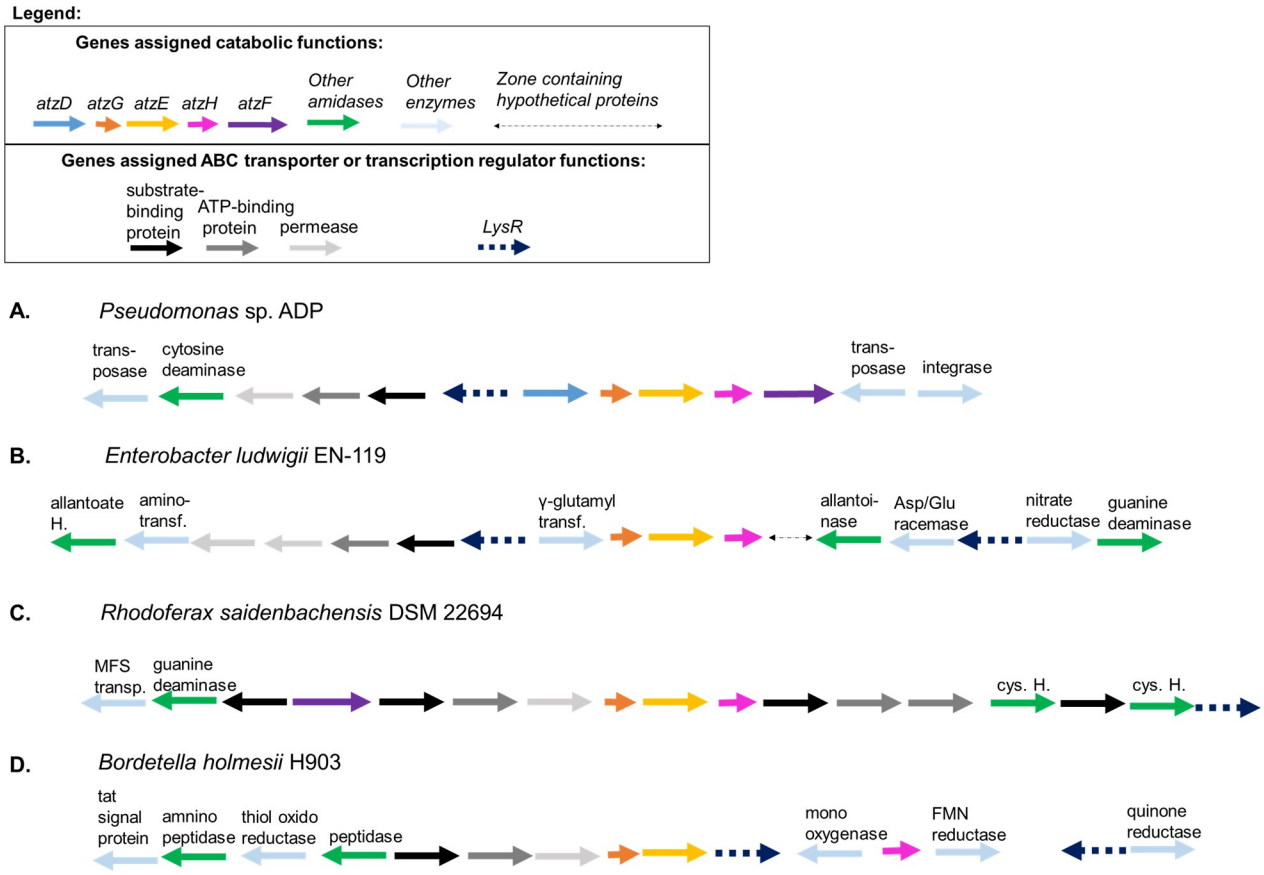


Fig 6. Representative genomic organization of *atzG-atzE-atzE* clusters in Proteobacteria. The organization of gene regions that contain the *atzG-atzE-atzH* clusters from A) *Pseudomonas* sp. strain ADP and, B-D) three bacteria that illustrate the most common genomic organizations associated with the cluster are shown. A comprehensive version is available (S3 Fig).

https://doi.org/10.1371/journal.pone.0206949.g006

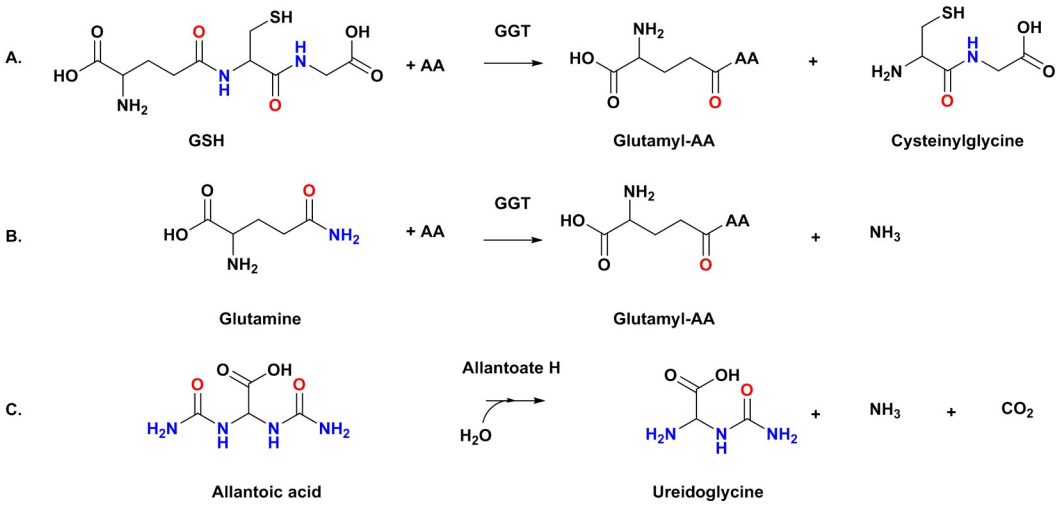


Fig 7. Hydrolyses catalyzed by enzymes encoded by genes associated with *atzG-atzE-atzH* clusters. A-B) Reactions performed by a gamma-glutamyltransferase (GGT) and C) an allantoate hydrolase. Ureido and amide groups are shown in colors.

https://doi.org/10.1371/journal.pone.0206949.g007

ureido groups (Fig 7) [46]. It is possible that *atzG-atzE-atzH* clusters form components of catabolic pathways in which the substrates possess amide groups.

There are also gene clusters found in β-proteobacter from the orders Comamonadacaea (*Rhodoferax* and *Variovorax*) and Burkholderiales (*Rhizobacter*, *Leptothrix*, *Methylibium* and *Rhizobiales*) that contain an allophanate hydrolase encoding gene. In all cases, the allophanate hydrolase gene is upstream of the *atzG-atzE-atzH* cluster and encoded on the same strand; in about half of the genomes there are genes encoding an ABC transporter separating the allophanate gene from the *atzG-atzE-atzH* cluster. Downstream, and encoded on the same strand as these *atzG-atzE-atzH* clusters, are genes encoding ABC transporters and amidase enzymes. This may indicate that these clusters are for the uptake and catabolism of nitrogenous compounds. (Fig 6C, S3 Fig)

Other Burkholderiales (*Bordetella* and *Achromobacter*) possess *atzG-atzE* and *atzH* clusters, but where the cluster also includes genes of unknown function that are annotated by homology as a transcriptional regulator, a monooxygenase and an FMN-reductase. In these clusters *atzH* is located between the genes encoding the monooxygenase and reductase, rather than immediately downstream of *atzG* and *atzE*. Finally, a number of diverse α-, β-, and γ-proteobacteria (*Pseudorhodoplanes*, *Methylobacterium*, *Thiomonas*, *Hydrogenophaga*, *Comamomonas*, *Pseudomonas*, and *Halothiobacillus*) have *atzG-atzE-atzH* clusters in unique genetic contexts, including the cyanuric acid catabolic operon of *Pseudomonas* sp. strain ADP. (Fig 6D, S3 Fig).

The occurrence of the *atzG-atzE-atzH* cluster appears to be far more wide-spread than the cyanuric acid catabolic pathway, and it is frequently associated with genes that may be involved in the catabolism of diverse nitrogen-containing compounds. However, the genetic contexts in which *atzH*-like genes are found provides little further evidence to suggest the specific function of the AtzH protein.

Conclusion

The data presented here imply that AtzH is a decarboxylating, 1,3-dicarboxyurea amidohydrolase. As yet its catalytic mechanism is not known, but our results suggest it involves the action of Tyr22 and Arg46, and that these amino acid residues are positioned relative to the substrate in a manner consistent with an acid-base mechanism. AtzH also facilitates the only previously uncontrolled step in the catabolic pathway, and as such, we believe that we have now identified the last remaining enzyme in the *Pseudomonas* sp. strain ADP cyanuric acid catabolism pathway (Fig 1B). The pathway appears to be dependent on enzymatic steps that outcompete solvent mediated hydrolysis that would otherwise lead to the accumulation of dead-end metabolites (biuret and urea). With no endogenous biuret or urea hydrolases, formation of these metabolites effectively prevent access to two of the three nitrogen atoms of the triazine ring.

Unexpectedly, *atzG-atzE-atzH* -like clusters appear to be relatively common within the Proteobacteria. Given the amidase function of AtzE and proposed amidase activity of AtzH, it is perhaps unsurprising that the *atzG-atzE-atzH*-like clusters tend to be associated with genes predicted to be involved in nitrogen compound metabolism. Previously, we had demonstrated that *atzG* and *atzE* were likely to have co-evolved from components of the bacterial GatCAB (transamidosome) complex. Here, we report the discovery of *atzG-atzE-atzH*-like clusters associated with diverse genetic contexts and metabolic functions, implying that the formation of the *atzG-atzE-atzH* cluster may predate the evolution of the cyanuric acid catabolism operon.

Supporting information

S1 Fig. AtzH protein sequence. The AtzH sequence is shown in black and the supplementary hexa-his-tag with thrombin cleavage site in red.
(DOCX)

S2 Fig. Purification of heterologous AtzH from *E. coli*. A. SDS-PAGE analysis of the fractions 2–11 of the Ni-NTA purification of His_6-AtzH (left), UV absorbance trace of the Ni-NTA purification (right), fractions 7 to 11 were concentrated and further purified; B. SDS-PAGE analysis of the size exclusion, showing the content of fractions 7–11 (left); UV absorbance trace of the size exclusion (right). On the traces, the green line in A. represents the % buffer B, the blue line shows the UV absorbance at 280 nm in arbitrary units; M: molecular marker, W: whole cell, S: soluble.
(DOCX)

S3 Fig. Genetic organization found flanking *atzE* homologues in other types of bacteria. A. Bacteria containing an allantoate hydrolase and glutamyl-transferase enzymes upstream of the *atzG-atzE-atzH* segment; B. Bacteria found to be containing an allophanate hydrolase homologue of AtzF found upstream of the *atzG-atzE-atzH* segment; C. Bacteria containing a non-adjacent AtzH homologue, located between a monooxygenase and reductase enzymes, downstream of the *atzG-atzE* segment; D. Bacteria presenting the *atzG-atzE-atzH* cluster within unique gene organizations. A legend is shown at the top of the figure. Green arrows represents amidase type enzymes, black and grey tones represent ABC transporters, discontinued dark blue arrows represent the presence of a *LysR* type regulator, genes encoding for AtzG, AtzE and AtzH homologues are shown by orange, yellow and pink arrows and light blue arrows represent other types of enzymes, respectively. This figure is not to scale.
(DOCX)

S1 Table. Mutagenic primers used for introducing point mutation in AtzH's sequence.
(DOCX)

S2 Table. DSF of the AtzH variants.
(DOCX)

Acknowledgments

We thank the Australian Synchrotron and MX beamline scientists for help with data collection and the CSIRO Collaborative Crystallisation Centre for DSF and crystallogenesis. We would also like to thank Drs Hafna Ahmed and Andrew Warden (CSIRO) for their critical comments during the preparation of this manuscript.

Author Contributions

Conceptualization: Colin Scott.

Formal analysis: Thomas S. Peat, Matthew Wilding, Janet Newman.

Funding acquisition: Colin Scott.

Investigation: Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Janet Newman.

Resources: Thomas S. Peat.

Supervision: Matthew Wilding, Carol J. Hartley, Colin Scott.

Writing – original draft: Lygie Esquirol, Thomas S. Peat, Janet Newman, Colin Scott.

Writing – review & editing: Lygie Esquirol, Thomas S. Peat, Matthew Wilding, Carol J. Hartley, Janet Newman, Colin Scott.

References

1. Udikovic-Kolic N, Scott C, Martin-Laurent F (2012) Evolution of atrazine-degrading capabilities in the environment. Appl Microbiol Biotechnol 96: 1175–1189. https://doi.org/10.1007/s00253-012-4495-0 PMID: 23076592
2. Cameron SM, Durchschein K, Richman JE, Sadowsky MJ, Wackett LP (2011) A new family of biuret hydrolases involved in *s*-triazine ring metabolism. ACS Catal 2011: 1075–1082. https://doi.org/10.1021/cs200295n PMID: 21897878
3. Esquirol L, Peat TS, Wilding M, Lucent D, French NG, Hartley CJ, et al. (2018) Structural and biochemical characterization of the biuret hydrolase (BiuH) from the cyanuric acid catabolism pathway of *Rhizobium leguminasorum* bv. *viciae* 3841. PLoS One 13: e0192736. https://doi.org/10.1371/journal.pone.0192736 PMID: 29425231
4. Mandelbaum RT, Allan DL, Wackett LP (1995) Isolation and characterization of a *Pseudomonas* sp that mineralizes the *s*-triazine herbicide atrazine. Applied and Environmental Microbiology 61: 1451–1457. PMID: 16534995
5. Martinez B, Tomkins J, Wackett LP, Wing R, Sadowsky MJ (2001) Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. J Bacteriol 183: 5684–5697. https://doi.org/10.1128/JB.183.19.5684-5697.2001 PMID: 11544232
6. Robinson SL, Badalamenti JP, Dodge AG, Tassoulas LJ, Wackett LP (2018) Microbial biodegradation of biuret: defining biuret hydrolases within the isochorismatase superfamily. Environ Microbiol.
7. Seffernick JL, Erickson JS, Cameron SM, Cho S, Dodge AG, Richman JE, et al. (2012) Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. J Bacteriol 194: 4579–4588. https://doi.org/10.1128/JB.00791-12 PMID: 22730121
8. Esquirol L, Peat TS, Wilding M, Liu JW, French NG, Hartley CJ, et al. (2018) An unexpected vestigial protein complex reveals the evoluionary history of an *s*-triazine catabolic enzyme. Journal of Biological Chemistry.
9. Peat TS, Balotra S, Wilding M, French NG, Briggs LJ, Panjikar S, et al. (2013) Cyanuric acid hydrolase: evolutionary innovation by structural concatenation. Mol Microbiol 88: 1149–1163. https://doi.org/10.1111/mmi.12249 PMID: 23651355
10. Cheng G, Shapir N, Sadowsky MJ, Wackett LP (2005) Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism. Appl Environ Microbiol 71: 4437–4445. https://doi.org/10.1128/AEM.71.8.4437-4445.2005 PMID: 16085834
11. Shapir N, Sadowsky MJ, Wackett LP (2005) Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp. strain ADP. J Bacteriol 187: 3731–3738. https://doi.org/10.1128/JB.187.11.3731-3738.2005 PMID: 15901697
12. Balotra S, Newman J, French NG, Briggs LJ, Peat TS, Scott C (2014) Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp. strain ADP. Acta Crystallogr F Struct Biol Commun 70: 310–315. https://doi.org/10.1107/S2053230X13034705 PMID: 24598916
13. Balotra S, Newman J, Cowieson NP, French NG, Campbell PM, Briggs LJ, et al. (2015) X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. Appl Environ Microbiol 81: 470–480. https://doi.org/10.1128/AEM.02783-14 PMID: 25362066
14. Rosa N, Ristic M, Seabrook SA, Lovell D, Lucent D, Newman J (2015) Meltdown: A Tool to Help in the Interpretation of Thermal Melt Curves Acquired by Differential Scanning Fluorimetry. Journal of Biomolecular Screening 20: 898–905. https://doi.org/10.1177/1087057115584059 PMID: 25918038
15. Watson AA, O'Callaghan CA (2005) Crystallization and X-ray diffraction analysis of human CLEC-2. Acta Crystallographica Section F-Structural Biology and Crystallization Communications 61: 1094–1096.
16. Newman J, Fazio VJ, Lawson B, Peat TS (2010) The C6 Web Tool: A Resource for the Rational Selection of Crystallization Conditions. Crystal Growth & Design 10: 2785–2792.
17. Winter G (2010) xia2: an expert system for macromolecular crystallography data reduction. Journal of Applied Crystallography 43: 186–190.
18. Waterman DG, Winter G, Gildea RJ, Parkhurst JM, Brewster AS, Sauter NK, et al. (2016) Diffraction-geometry refinement in the DIALS framework. Acta Crystallographica Section D-Structural Biology 72: 558–575.
19. Parkhurst JM, Winter G, Waterman DG, Fuentes-Montero L, Gildea RJ, Murshudov GN, et al. (2016) Robust background modelling in DIALS. Journal of Applied Crystallography 49: 1912–1921. https://doi.org/10.1107/S1600576716013595 PMID: 27980508

20. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallographica Section D-Biological Crystallography 63: 32–41.
21. Mccoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. Journal of Applied Crystallography 40: 658–674. https://doi.org/10.1107/S0021889807021206 PMID: 19461840
22. Evans PR, Murshudov GN (2013) How good are my data and what is the resolution? Acta Crystallographica Section D-Biological Crystallography 69: 1204–1214.
23. Stein N (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. Journal of Applied Crystallography 41: 641–643.
24. Bailey S (1994) The Ccp4 Suite—Programs for Protein Crystallography. Acta Crystallographica Section D-Biological Crystallography 50: 760–763.
25. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. Acta Crystallographica Section D-Biological Crystallography 66: 486–501.
26. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallographica Section D-Biological Crystallography 53: 240–255.
27. (JCSG) JCfSG (2007) Crystal structure of uncharacterized protein conserved in bacteria with a cystatin-like fold (YP_051588.1) from Erwinia carotovora subsp. atroseptica SCRI1043 at 2.32 A resolution. To be published.
28. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 372: 774–797. https://doi.org/10.1016/j.jmb.2007.05.022 PMID: 17681537
29. pfam clan: NTF2 (CL0051), NTF2-like superfamily.
30. Sultana A, Kallio P, Jansson A, Wang JS, Niemi J, Mantsala P, et al. (2004) Structure of the polyketide cyclase SnoaL reveals a novel mechanism for enzymatic aldol condensation. Embo Journal 23: 1911–1921. https://doi.org/10.1038/sj.emboj.7600201 PMID: 15071504
31. Nakasako M, Motoyama T, Kurahashi Y, Yamaguchi I (1998) Cryogenic X-ray crystal structure analysis for the complex of scytalone dehydratase of a rice blast fungus and its tight-binding inhibitor, carpropamid: The structural basis of tight-binding inhibition. Biochemistry 37: 9931–9939. https://doi.org/10.1021/bi980321b PMID: 9665698
32. Arand M, Hallberg BM, Zou JY, Bergfors T, Oesch F, van der Werf MJ, et al. (2003) Structure of Rhodococcus erythropolis limonene-1,2-epoxide hydrolase reveals a novel active site. Embo Journal 22: 2583–2592. https://doi.org/10.1093/emboj/cdg275 PMID: 12773375
33. Kim SW, Cha SS, Cho HS, Kim JS, Ha HC, Cho MJ, et al. (1997) High-resolution crystal structures of Delta(5)-3-ketosteroid isomerase with and without a reaction intermediate analogue. Biochemistry 36: 14030–14036. https://doi.org/10.1021/bi971546+ PMID: 9369474
34. Griffith LC, Lu CS, Sun XX (2003) CaMKII, an enzyme on the move: regulation of temporospatial localization. Mol Interv 3: 386–403. https://doi.org/10.1124/mi.3.7.386 PMID: 14993460
35. Ott M, Prestele M, Bauerschmitt H, Funes S, Bonnefoy N, Herrmann JM (2006) Mba1, a membrane-associated ribosome receptor in mitochondria. EMBO J 25: 1603–1610. https://doi.org/10.1038/sj.emboj.7601070 PMID: 16601683
36. Goessweiner-Mohr N, Grumet L, Arends K, Pavkov-Keller T, Gruber CC, Gruber K, et al. (2013) The 2.5 A structure of the enterococcus conjugation protein TraM resembles VirB8 type IV secretion proteins. J Biol Chem 288: 2018–2028. https://doi.org/10.1074/jbc.M112.428847 PMID: 23188825
37. Porter CJ, Bantwal R, Bannam TL, Rosado CJ, Pearce MC, Adams V, et al. (2012) The conjugation protein TcpC from Clostridium perfringens is structurally related to the type IV secretion system protein VirB8 from Gram-negative bacteria. Mol Microbiol 83: 275–288. https://doi.org/10.1111/j.1365-2958.2011.07930.x PMID: 22150951
38. Terradot L, Bayliss R, Oomen C, Leonard GA, Baron C, Waksman G (2005) Structures of two core subunits of the bacterial type IV secretion system, VirB8 from Brucella suis and ComB10 from Helicobacter pylori. Proc Natl Acad Sci U S A 102: 4596–4601. https://doi.org/10.1073/pnas.0408927102 PMID: 15764702
39. Ribbeck K, Lipowsky G, Kent HM, Stewart M, Gorlich D (1998) NTF2 mediates nuclear import of Ran. EMBO J 17: 6587–6598. https://doi.org/10.1093/emboj/17.22.6587 PMID: 9822603
40. Eberhardt RY, Chang YY, Bateman A, Murzin AG, Axelrod HL, Hwang WC, et al. (2013) Filling out the structural map of the NTF2-like superfamily. Bmc Bioinformatics 14.
41. Holliday GL, Mitchell JBO, Thornton JM (2009) Understanding the Functional Roles of Amino Acid Residues in Enzyme Catalysis. Journal of Molecular Biology 390: 560–577. https://doi.org/10.1016/j.jmb.2009.05.015 PMID: 19447117

42. Zechel DL, Withers SG (2001) Dissection of nucleophilic and acid-base catalysis in glycosidases. Current Opinion in Chemical Biology 5: 643–649. PMID: 11738173

43. Loh KD, Gyaneshwar P, Papadimitriou EM, Fong R, Kim KS, Parales R, et al. (2006) A previously undescribed pathway for pyrimidine catabolism. Proceedings of the National Academy of Sciences of the United States of America 103: 5114–5119. https://doi.org/10.1073/pnas.0600521103 PMID: 16540542

44. Esquirol L, Peat TS, Wilding M, Liu JW, French NG, Hartley CJ, et al. (2018) An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme. Journal of Biological Chemistry 293: 7880–7891. https://doi.org/10.1074/jbc.RA118.001996 PMID: 29523689

45. Zhao J, Zhu L, Fan C, Wu Y, Xiang S (2018) Structure and function of urea amidolyase. Bioscience Reports 38: 12.

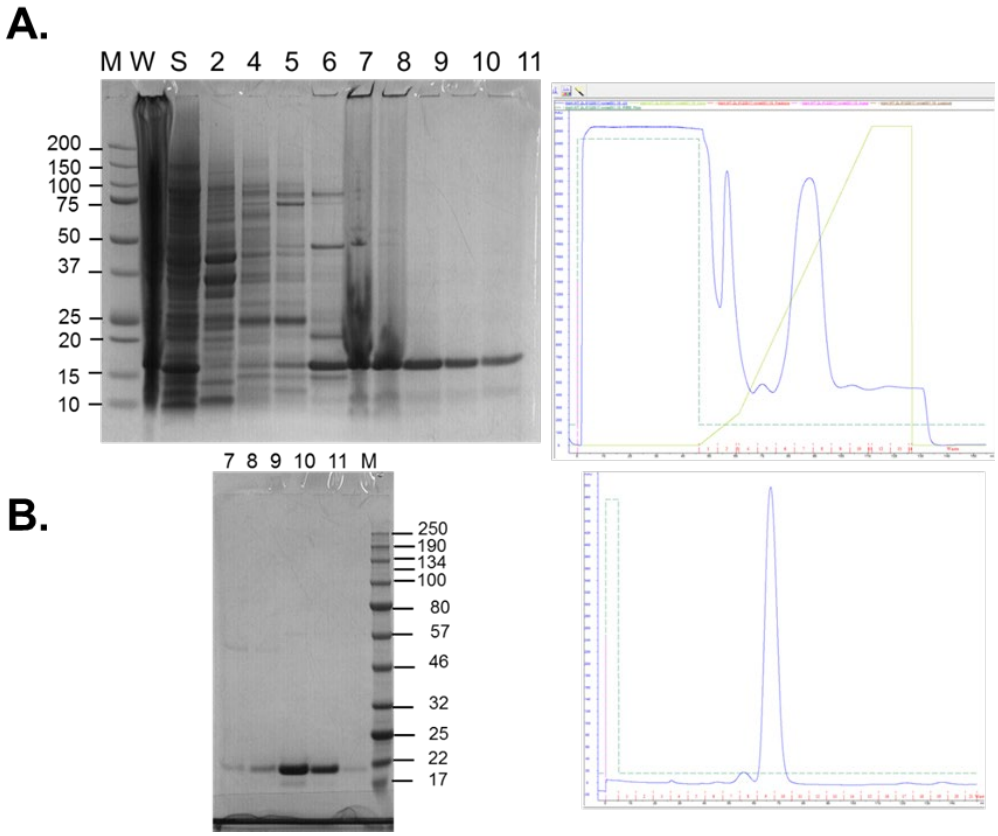
46. Agarwal R, Burley SK, Swaminathan S (2007) Structural analysis of a ternary complex of allantoate amidohydrolase from Escherichia coli reveals its mechanics. Journal of Molecular Biology 368: 450–463. https://doi.org/10.1016/j.jmb.2007.02.028 PMID: 17362992

MGSSHHHHHHSSGLVPRGSHMQINLPEVHAEVTAQFVRYEKALTSNDTAVLNELF

WNSPQTLRYGATENLYGYEAIAGFRATRSPNNLEREIVRTVITTYGHDFATANIEFR

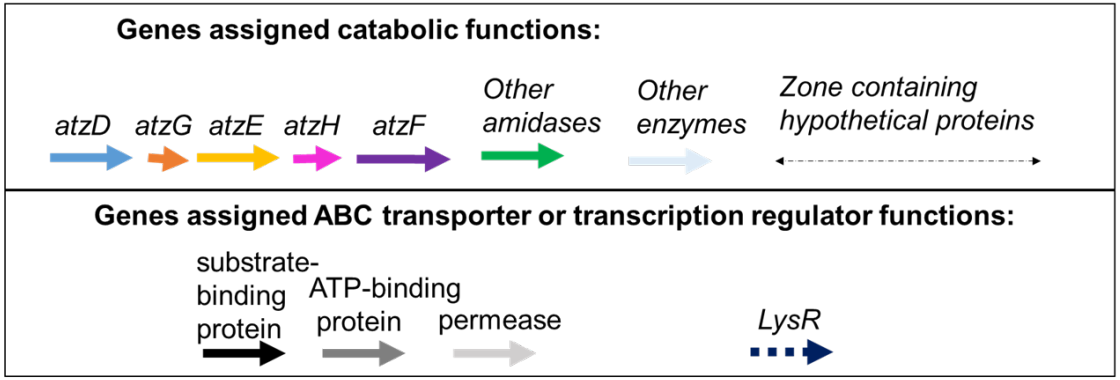
RLSHSQLTGRQSQTWMRTSQGWRVVAAHVSLIALPVS

S1 Fig: AtzH protein sequence. The AtzH sequence is shown in black and the supplementary hexa-his-tag with thrombin cleavage site in red.

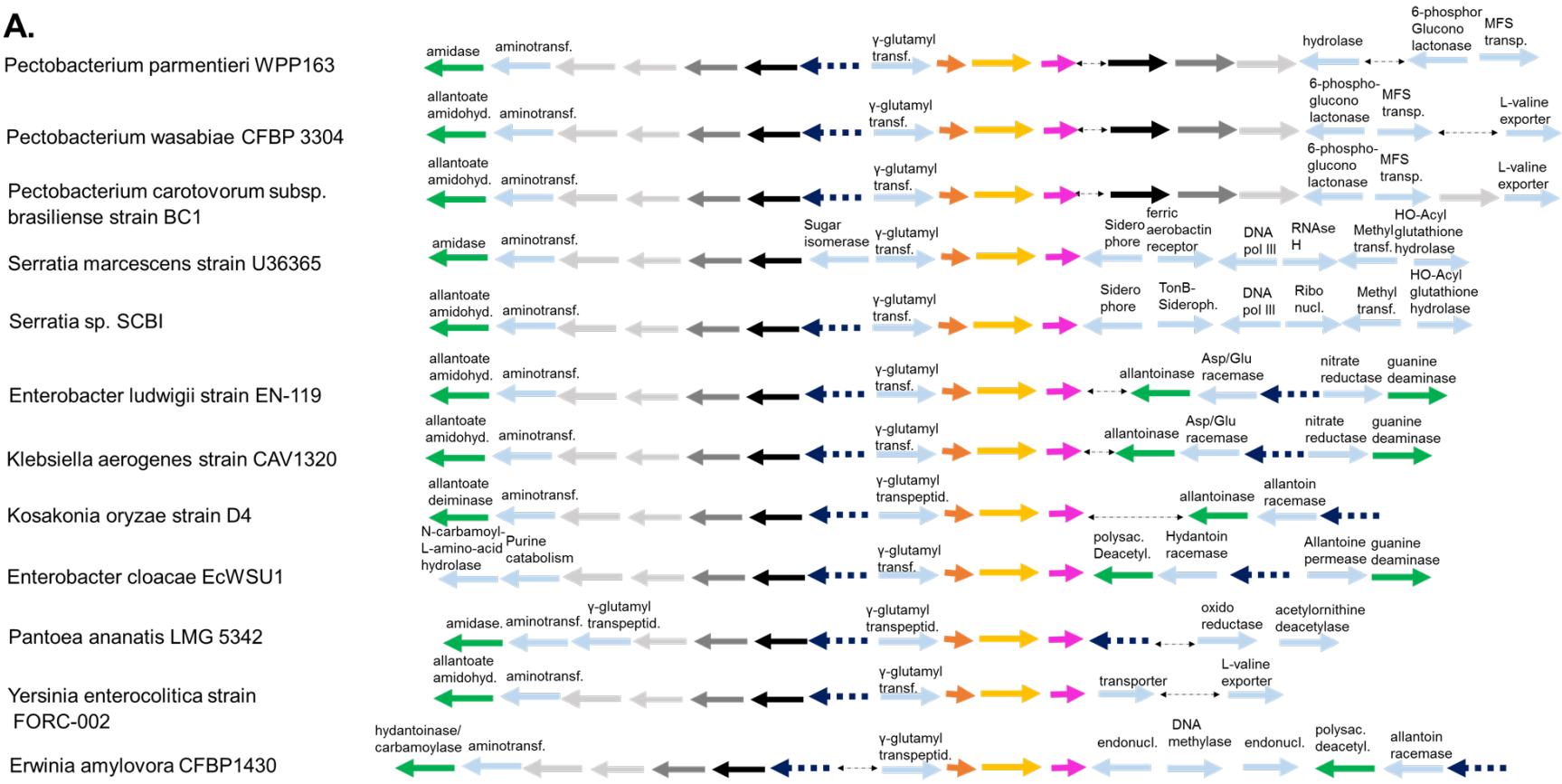


S2 Fig: Purification of heterologous AtzH from *E. coli*. A. SDS-PAGE analysis of the fractions 2-11 of the Ni-NTA purification of His_6-AtzH (left), UV absorbance trace of the Ni-NTA purification (right), fractions 7 to 11 were concentrated and further purified; B. SDS-PAGE analysis of the size exclusion, showing the content of fractions 7-11 (left); UV absorbance trace of the size exclusion (right). On the traces, the green line in A. represents the % buffer B, the blue line shows the UV absorbance at 280 nm in arbitrary units; M: molecular marker, W: whole cell, S: soluble.

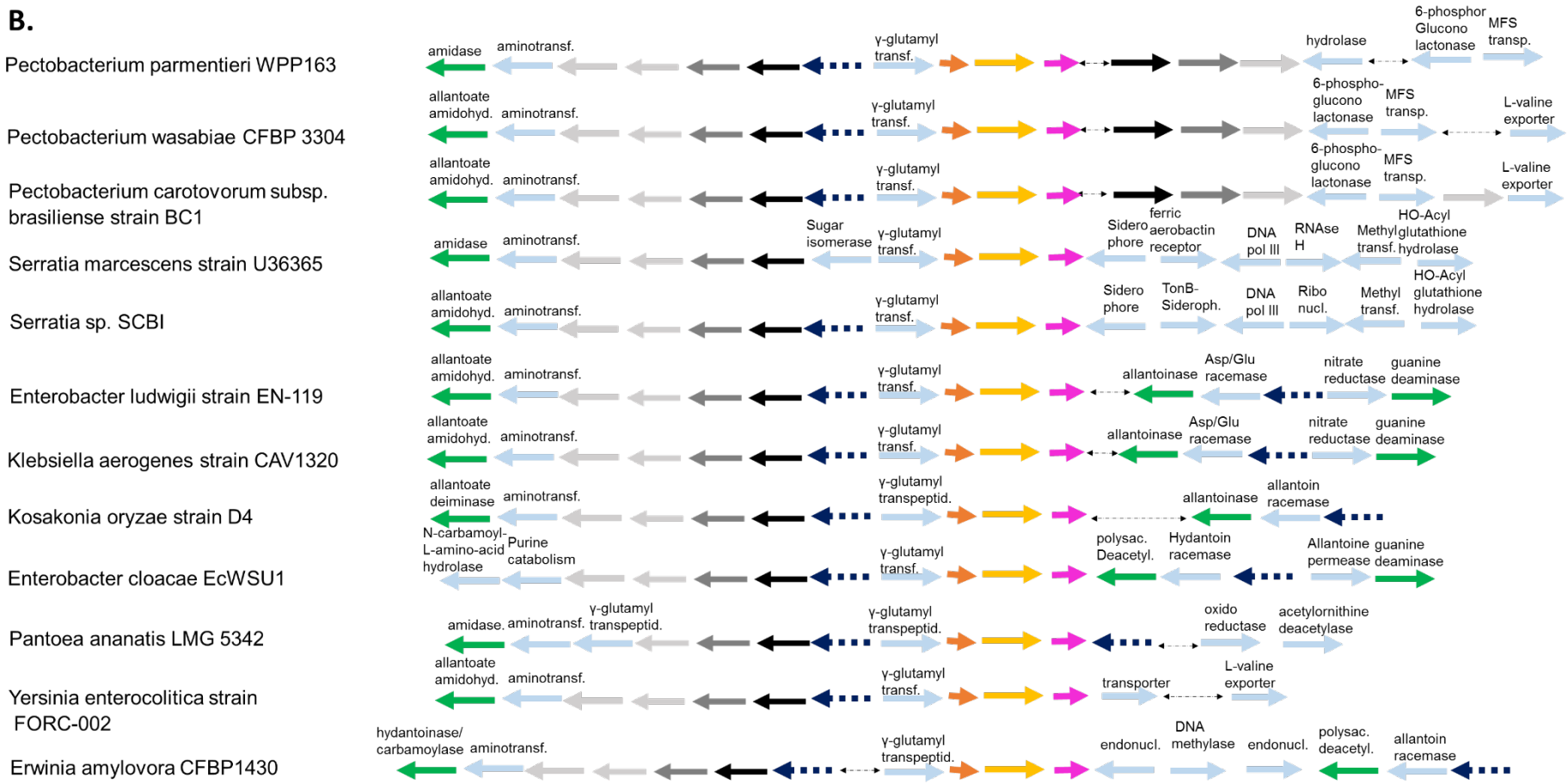
Legend:


Genes assigned catabolic functions:
atzD
atzG
atzE
atzH
atzF
Other amidases
Other enzymes
Zone containing hypothetical proteins
Genes assigned ABC transporter or transcription regulator functions:
substrate-binding protein
ATP-binding protein
permease
LysR

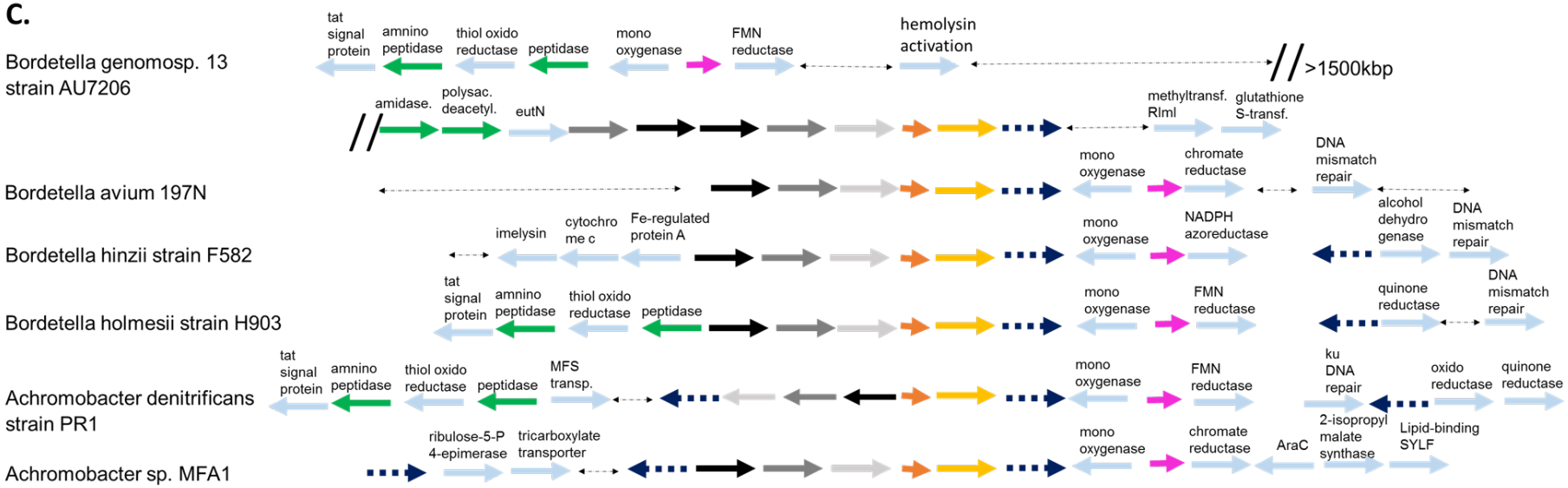
A.

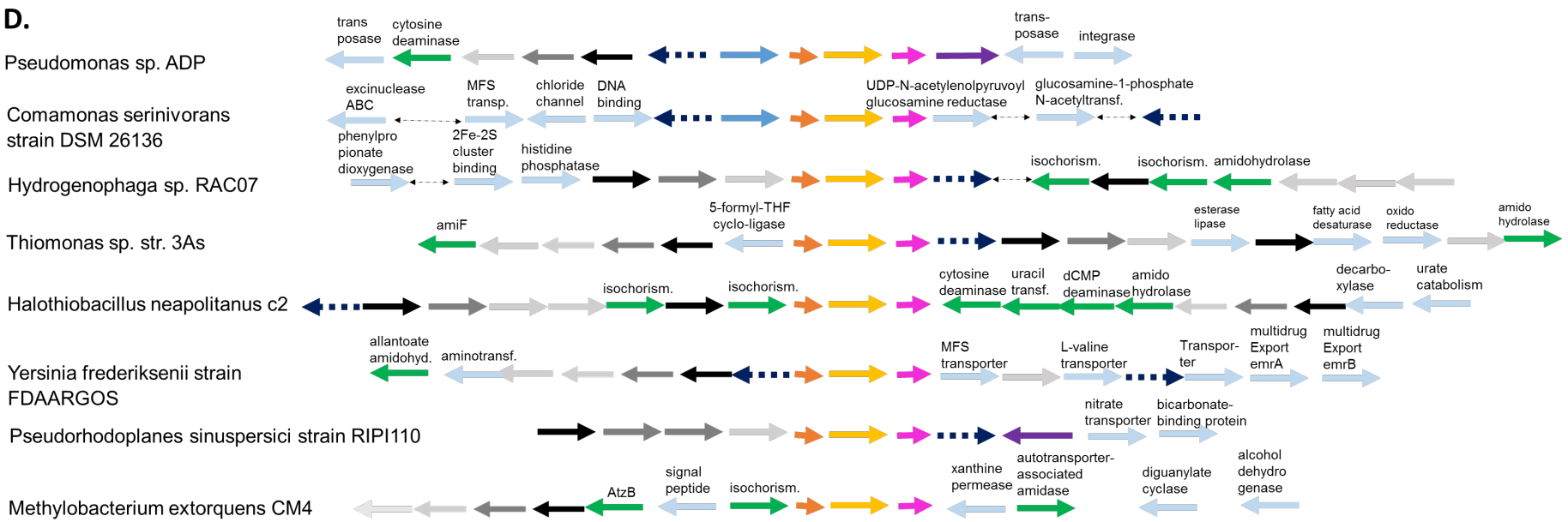

Pectobacterium parmentieri WPP163
amidase
aminotransf.
γ-glutamyl transf.
hydrolase
6-phosphor Glucono lactonase
MFS transp.
Pectobacterium wasabiae CFBP 3304
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
6-phospho-glucono lactonase
MFS transp.
L-valine exporter
Pectobacterium carotovorum subsp. brasiliense strain BC1
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
6-phospho-glucono lactonase
MFS transp.
L-valine exporter
Serratia marcescens strain U36365
amidase
aminotransf.
Sugar isomerase
γ-glutamyl transf.
Sidero phore
ferric aerobactin receptor
DNA pol III
RNAse H
Methyl transf.
HO-Acyl glutathione hydrolase
Serratia sp. SCBI
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
Sidero phore
TonB-Sideroph.
DNA pol III
Ribo nucl.
Methyl transf.
HO-Acyl glutathione hydrolase
Enterobacter ludwigii strain EN-119
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
allantoinase
Asp/Glu racemase
nitrate reductase
guanine deaminase
Klebsiella aerogenes strain CAV1320
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
allantoinase
Asp/Glu racemase
nitrate reductase
guanine deaminase
Kosakonia oryzae strain D4
allantoate deiminase
aminotransf.
γ-glutamyl transpeptid.
allantoinase
allantoin racemase
Enterobacter cloacae EcWSU1
N-carbamoyl-L-amino-acid hydrolase
Purine catabolism
γ-glutamyl transf.
polysac. Deacetyl.
Hydantoin racemase
Allantoine permease
guanine deaminase
Pantoea ananatis LMG 5342
amidase.
aminotransf.
γ-glutamyl transpeptid.
γ-glutamyl transpeptid.
oxido reductase
acetylornithine deacetylase
Yersinia enterocolitica strain FORC-002
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
transporter
L-valine exporter
Erwinia amylovora CFBP1430
hydantoinase/ carbamoylase
aminotransf.
γ-glutamyl transpeptid.
endonucl.
DNA methylase
endonucl.
polysac. deacetyl.
allantoin racemase

B.


Pectobacterium parmentieri WPP163
amidase
aminotransf.
γ-glutamyl transf.
hydrolase
6-phosphor Glucono lactonase
MFS transp.
Pectobacterium wasabiae CFBP 3304
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
6-phospho-glucono lactonase
MFS transp.
L-valine exporter
Pectobacterium carotovorum subsp. brasiliense strain BC1
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
6-phospho-glucono lactonase
MFS transp.
L-valine exporter
Serratia marcescens strain U36365
amidase
aminotransf.
Sugar isomerase
γ-glutamyl transf.
Sidero phore
ferric aerobactin receptor
DNA pol III
RNAse H
Methyl transf.
HO-Acyl glutathione hydrolase
Serratia sp. SCBI
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
Sidero phore
TonB-Sideroph.
DNA pol III
Ribo nucl.
Methyl transf.
HO-Acyl glutathione hydrolase
Enterobacter ludwigii strain EN-119
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
allantoinase
Asp/Glu racemase
nitrate reductase
guanine deaminase
Klebsiella aerogenes strain CAV1320
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
allantoinase
Asp/Glu racemase
nitrate reductase
guanine deaminase
Kosakonia oryzae strain D4
allantoate deiminase
aminotransf.
γ-glutamyl transpeptid.
allantoinase
allantoin racemase
Enterobacter cloacae EcWSU1
N-carbamoyl-L-amino-acid hydrolase
Purine catabolism
γ-glutamyl transf.
polysac. Deacetyl.
Hydantoin racemase
Allantoine permease
guanine deaminase
Pantoea ananatis LMG 5342
amidase.
aminotransf.
γ-glutamyl transpeptid.
γ-glutamyl transpeptid.
oxido reductase
acetylornithine deacetylase
Yersinia enterocolitica strain FORC-002
allantoate amidohyd.
aminotransf.
γ-glutamyl transf.
transporter
L-valine exporter
Erwinia amylovora CFBP1430
hydantoinase/ carbamoylase
aminotransf.
γ-glutamyl transpeptid.
endonucl.
DNA methylase
endonucl.
polysac. deacetyl.
allantoin racemase

C.


Bordetella genomosp. 13 strain AU7206
tat signal protein
amnino peptidase
thiol oxido reductase
peptidase
mono oxygenase
FMN reductase
hemolysin activation
>1500kbp
amidase.
polysac. deacetyl.
eutN
methyltransf. RlmI
glutathione S-transf.
Bordetella avium 197N
mono oxygenase
chromate reductase
DNA mismatch repair
Bordetella hinzii strain F582
imelysin
cytochro me c
Fe-regulated protein A
NADPH azoreductase
alcohol dehydro genase
DNA mismatch repair
Bordetella holmesii strain H903
quinone reductase
Achromobacter denitrificans strain PR1
MFS transp.
ku DNA repair
oxido reductase
quinone reductase
Achromobacter sp. MFA1
ribulose-5-P 4-epimerase
tricarboxylate transporter
AraC
2-isopropyl malate synthase
Lipid-binding SYLF



S3 Fig: Genetic organization found flanking *atzE* homologues in other types of bacteria. A. Bacteria containing an allantoate hydrolase and glutamyl-transferase enzymes upstream of the *atzG-atzE-atzH* segment; B. Bacteria found to be containing an allophanate hydrolase homologue of AtzF found upstream of the *atzG-atzE-atzH* segment; C. Bacteria containing a non-adjacent AtzH homologue, located between a monooxygenase and reductase enzymes, downstream of the *atzG-atzE* segment; D. Bacteria presenting the *atzG-atzE-atzH* cluster within unique gene organizations. A legend is shown at the top of the figure. Green arrows represents amidase type enzymes, black and grey tones represent ABC transporters, discontinued dark blue arrows represent the presence of a *LysR* type regulator, genes encoding for AtzG, AtzE and AtzH homologues are shown by orange, yellow and pink arrows and light blue arrows represent other types of enzymes, respectively. This figure is not to scale.

S1 Table: Mutagenic primers used for introducing point mutation in AtzH's sequence.

Amino acid mutated	Mutagenic primer 5' → 3'
Tyr22Ala Fwd	cgcccaattcgttcgcgccgagaaagcgcttacc
Tyr22Phe Rev	gcgctttctcgaagcgaacgaattgggcg
Arg46Ala Fwd	cccccaaaccctggcctacggtgcgacc
Arg46Lys Fwd	agcccccaaaccctgaagtacggtgcgaccgag
Arg63Ala Fwd	gcgtgtggcggcaaagcctgcaatggcctca
Arg60Lys Rev	agagcgtgtggccttaaagcctgcaatggcctcatagc
Arg66Ala Fwd	aggctttcgcgccacagcctctccaaacaacttg
Arg63Lys Rev	ccaagttgtttggagactttgtggcgcgaaagcctgcaatgg
Arg73Ala Fwd	gctctccaaacaacttggaagccgagatcgtccgga
Arg73Lys Rev	ccgtccggacgatctccttttccaagttgtttggagagcgt
Arg96Ala Fwd	ccgccaacatcgaattccgtgcgctcagtcatagtc
Arg96Lys Fwd	taccgccaacatcgaattccgtaagctcagtcatagtcag
Glu106Ala Fwd	gcttacgggccgggcgagccaaacctgg
Glu106Ala Rev	gggccggcagagcgcaacctggatgcgc

S2 Table: DSF of the AtzH variants.

AtzH Variant	T_m (°C)
Y22A	48.8 +/- 0.04
Y22F	57.0 +/- 0.09
R46A	55.3 +/- 0.02
R46K	58.3 +/- 0.05
R63A	55.3 +/- 0.04
R63K	54.2 +/- 0.13
R66A	55.3 +/- 0.04
R66K	55.0 +/- 0.05
R73A	55.5 +/- 0.06
R73K	50.6 +/- 0.09
R96A	57.0 +/- 0.03
R96K	55.6 +/- 0.05
Q106A	50.2 +/- 0.04
Q108A	52.1 +/- 0.06

5 Conclusion

5.1 PhD outcomes

The work presented here has added to our understanding of the partially characterised biuret amidohydrolase (BiuH) from *Rhizobium leguminasorum* bv. *viciae* 3841. The atomic structure of BiuH was solved and site-directed mutagenesis was used to gain a better understanding of the BiuH catalytic mechanism. Additionally, molecular dynamics simulations highlighted the presence of three channels from the active site to the enzyme surface forming potential channels for the substrate, co-product (ammonia) channel and co-substrate (water) channel.

The cyanuric acid degradation pathway in *Pseudomonas* sp. strain ADP1 has been known and studied for more than twenty years; however, the work presented here includes the first purification and and characterisation of AtzE, a putative biuret amidohydrolase. In chapter three the purification of the native AtzE from *Pseudomonas* sp. strain ADP was reported, allowing its biochemical and structural characterisation. The structure revealed the presence of a small, essential protein (AtzG), with which AtzE forms an a_2b_2 heterotetramer. Biochemical characterisation and molecular dynamics experiments revealed AtzE is a 1-carboxybiuret hydrolase, and not a biuret hydrolase as previously thought. Finally, this work implies that AtzEG has probably evolved from the GatCAB transamidosome complex.

The discovery of *atzG* encouraged a more thorough examination of the cyanuric acid catabolic operon and the discovery of a second over looked open-reading frame, *atzH*. The fourth chapter of this PhD thesis details the structural characterisation of AtzH, which belongs to the versatile NTF2 protein superfamily.

A combination of structural modelling and mutagenesis studies was used to provide evidence that AtzH is an allophanate forming, 1,3-dicarboxyurea amidohydrolase. Mutagenesis also indicated that Tyr22 and Arg46 may play an essential role in the catalysis of 1,3-dicarboxyurea. Finally, a comparison of the genomic context suggests AtzH might be involved more broadly in the catabolism of nitrogenous compounds in Proteobacteria. Moreover, this observation also suggests that the *atzG-atzE-atzH* cluster predates the formation of the cyanuric acid catabolism operon.

5.2 Update of the cyanuric acid degradation pathway of *Pseudomonas* sp. strain ADP

The *Pseudomonas* sp. ADP cyanuric acid catabolic pathway requires revision in light of the work presented here (Fig. 5-1). The updated pathway does not include biuret, as previously thought, but rather 1-carboxybiuret. Additionally, a new hydrolytic step has been added: the hydrolysis of 1,3-dicarboxyurea to form allophanate. The updated pathway contains two newly identified proteins: AtzG (Chapter 3, (181)) and AtzH (Chapter 4).

This update provides a clearer picture of the cyanuric acid degradation strategy in *Pseudomonas* sp. strain ADP. The *Pseudomonas* enzymes AtzEG, AtzH and AtzF enable controlled hydrolysis of the solvent labile intermediates 1-carboxybiuret, 1,3-dicarboxyurea and allophanate, respectively. In doing so, *Pseudomonas* sp. strain ADP avoids the competing solvent-mediated hydrolyses, which result in metabolites that *Pseudomonas* sp. strain ADP cannot process further (i.e., urea and biuret). These metabolites are effectively ‘dead end’ products, making the nitrogen they contain unavailable to the bacterium, as *Pseudomonas* lacks either urease or biuret hydrolase.

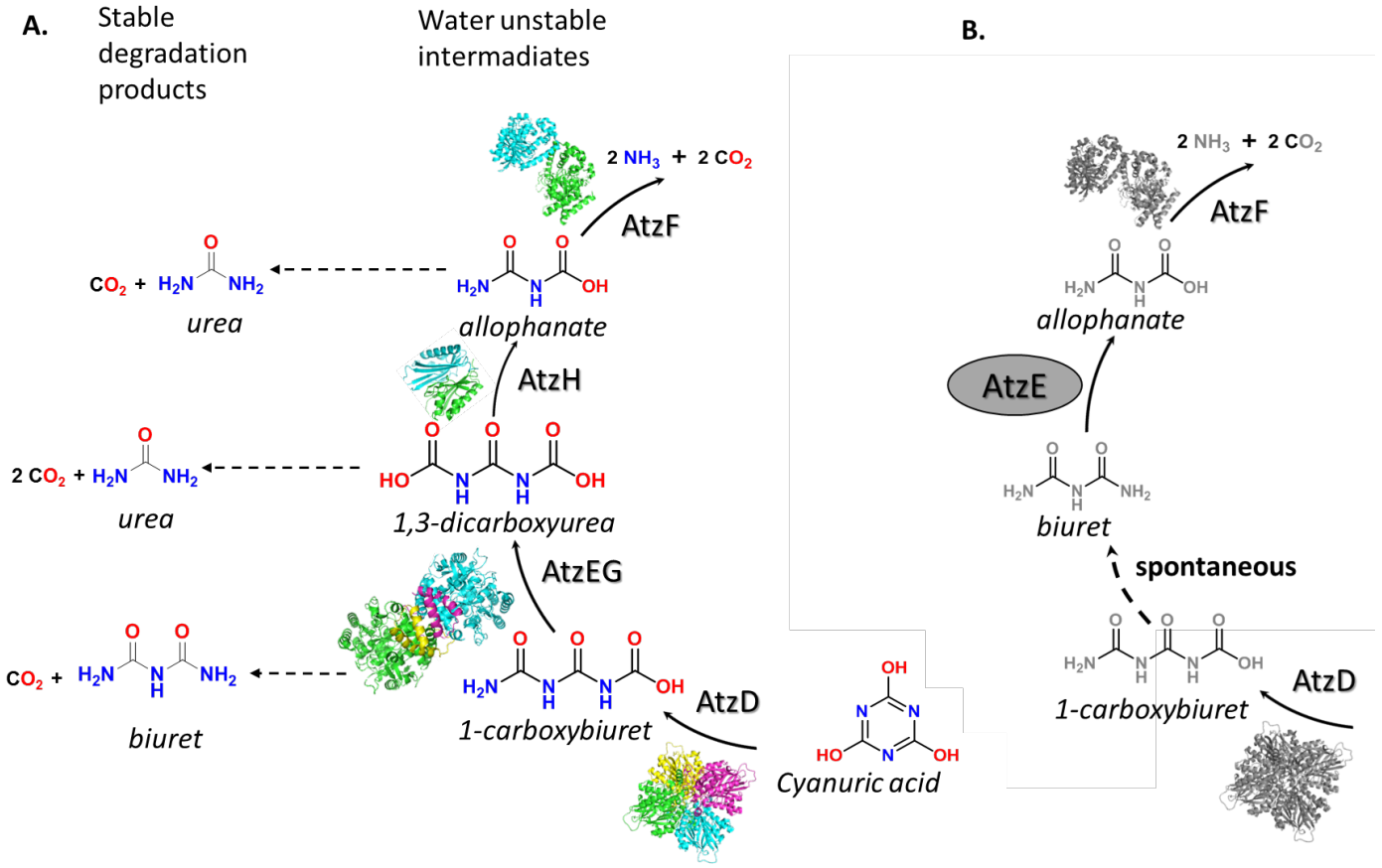


Figure 5-1 Cyanuric acid degradation pathway in* Pseudomonas *sp. strain ADP prevents the degradation of unstable intermediates. A. *New proposed cyanuric acid degradation pathway showing the new structures AtzEG and AtzH and their water sensitive substrate;* ***B.*** *Cyanuric acid degradation pathway as published previously to this work.*

Our work showed that AtzG and AtzH play critical roles in the cyanuric acid catabolic pathway of *Pseudomonas* sp. strain ADP. Often overlooked and unannotated, small proteins have been found to be common and involved in essential and versatile roles (182-185). Genomic analysis suggests that, in addition to its role in the cyanuric acid degradation pathway, AtzH may play a catalytic role in other bacterial metabolic pathways. A more thorough investigation will be required to characterise the role of these AtzH homologues.

5.3 Insights into the evolutionary origin of the cyanuric acid degradation operon

Understanding the origin of each of the genes of the cyanuric acid degradation operon has provided insight into the evolution of the cyanuric acid degradation operon (Fig. 5-2).

AtzD belongs to a unique family of ring-opening hydrolases (the cyclic amide hydrolase family) that includes the cyanuric acid and barbituric acid hydrolases. Recently, a barbituric acid hydrolase with promiscuous cyanuric acid hydrolase activity was described from *Nocardioides* sp. strain JS614 (149). When coupled with the structural similarities between the enzymes and substrates, these observations suggest that cyanuric acid hydrolase enzymes may have evolved from an ancestral barbituric acid hydrolase. There is structural evidence to suggest the origin of the cyclic amide hydrolase family. The trimeric proteins from the YgjF-family have structural homology with each of the three sub-domains of AtzD. This could indicate that the cyclic amide hydrolases are the product of concatenation of three *ygjF* genes, possibly following gene duplication (146, 147) (Fig. 5-2).

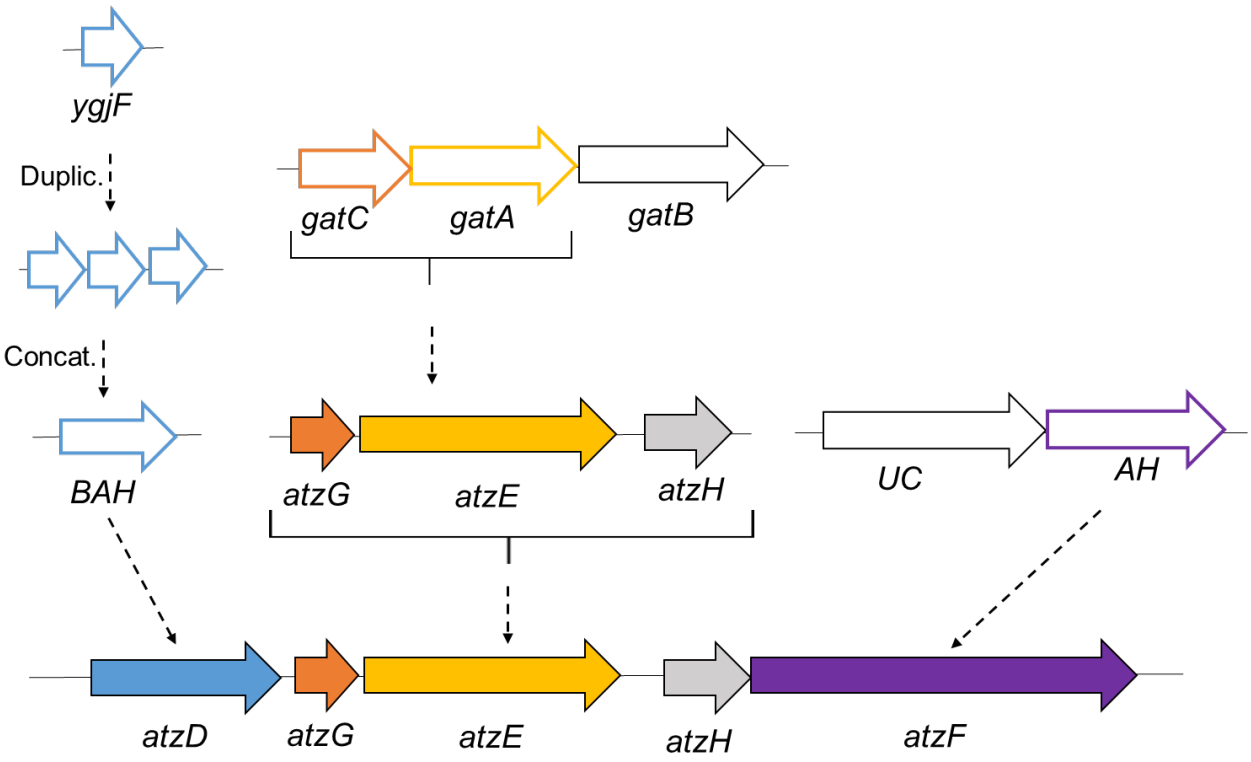


Figure 5-2 Hypothetical evolutionary origin of the* Pseudomonas sp. *strain ADP cyanuric acid degradation operon. *Concat.: concatenation; Duplic.: duplication.*

Our work has also suggested a probable evolutionary origin for the unusual AtzEG complex. Based on the structural alignment of AtzEG with other structures available in the PDB database, the GatCAB complex was identified as the highest structural match for AtzEG. Based on this similarity, we hypothesised that AtzEG might have evolved from the GatCA complex (Chapter 3)(181) (Fig. 5-2). Only five amino acids are different between the active sites of GatCA and AtzEG: Ala301Tyr, Asn172Gly, Asp425Gly, Phe127Gly and Tyr125Met (Fig. 5-3). Like the AtzA/TriA and TrzN systems (157, 161, 163), the small number of differences makes reconstruction of evolutionary trajectories that separate these complexes tractable. Future work could focus on reconstructing, *via* site-directed mutagenesis, the evolutionary pathway leading from GatCA to AtzEG or from AtzEG to GatCA.

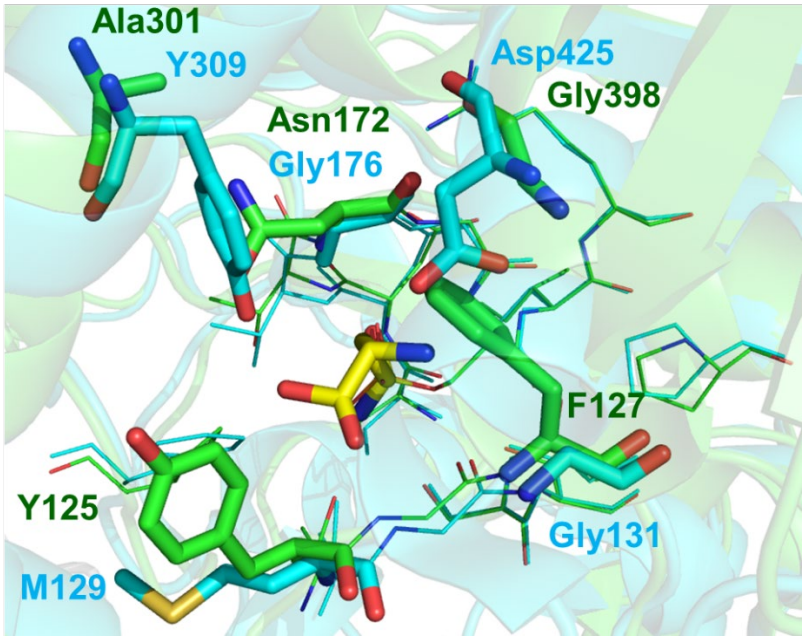


Figure 5-3 Alignment of the active sites of GatA (cyan) and AtzE (green). *The non-conserved residues in both active sites are shown in sticks and the conserved amino acids in lines, the glutamine (substrate of GatA) is shown in yellow stick. These five amino acids could be potential targets for the reconstruction of evolutionary trajectories between AtzEG and GatA by directed mutagenesis.*

AtzEG and AtzH have been found to be expressed as part of the cyanuric acid degradation pathway of *Pseudomonas* sp. strain ADP; however, the *atzG-atzE-atzH* gene cluster is found in the genomes of many different strains of bacteria (Chapter 4) without the presence of the *atzD* and *atzF*. This suggests that the association of these three genes predates the formation of the cyanuric acid operon (Fig. 5-2).

AtzF appears to have been recruited from the urea carboxylases complex (141), in which allophanate hydrolase is complexed with urea carboxylase (178) (Fig. 5-2). Lin *et al.* have demonstrated that although UC and AH are able to function independently, complexation results in a rate enhancement (186). Interestingly,

one third of the bacterial genomes that possess an *atzG-atzE-atzH* gene cluster also have an a*tzF* homologue in the same gene region. However, *atzD* homologues are found in association with only two of the *atzG-atzE-atzH* clusters described and the *atzF* homologue does not appear to form an operon with the *atzG-atzE-atzH* clusters (Chapter 4).

Coupling of the catalytic steps of the *Pseudomonas* sp. strain ADP cyanuric acid catabolism pathway may be required to ensure that 'dead-end' metabolites do not accumulate *via* the competing non-enzymatic hydrolyses of the unstable metabolic intermediates. The emergence of the cyanuric acid operon may have been an evolutionary response that allows co-ordinated expression and maintains stoichiometric production of the enzymes of this pathway (187, 188). It may be possible to use the cyanuric acid degradation operon to test the importance of operon organisation for the fitness of the bacteria.

5.4 An alternative cyanuric acid degradation pathway in Rhizobium *leguminosarum* bv. *viciae* 3841

BiuH and AtzEG have non-overlapping specificities: AtzEG is a 1-carboxybiuret amidohydrolase and BiuH is a biuret amidohydrolase (118, 181). This difference illustrates that there are at least two strategies employed for the catabolism of cyanuric acid: *Rhizobium* exploits solvent-mediated hydrolysis of the metabolic intermediates of the pathway while *Pseudomonas* has adopted a pathway that limits solvent-mediated hydrolysis (Fig. 5-4).

More work is needed in order to understand the fate of the allophanate produced by BiuH in *Rhizobium*. There is no equivalent of *atzF* in close proximity to *atzD* and *biuH (153)*. However, the genome of *Rhizobium leguminosarum* bv. *viciae* 3841 (131) possesses a number of genes that encode enzymes that could degrade allophanate or its hydrolysis product, urea (Fig. 5-4).

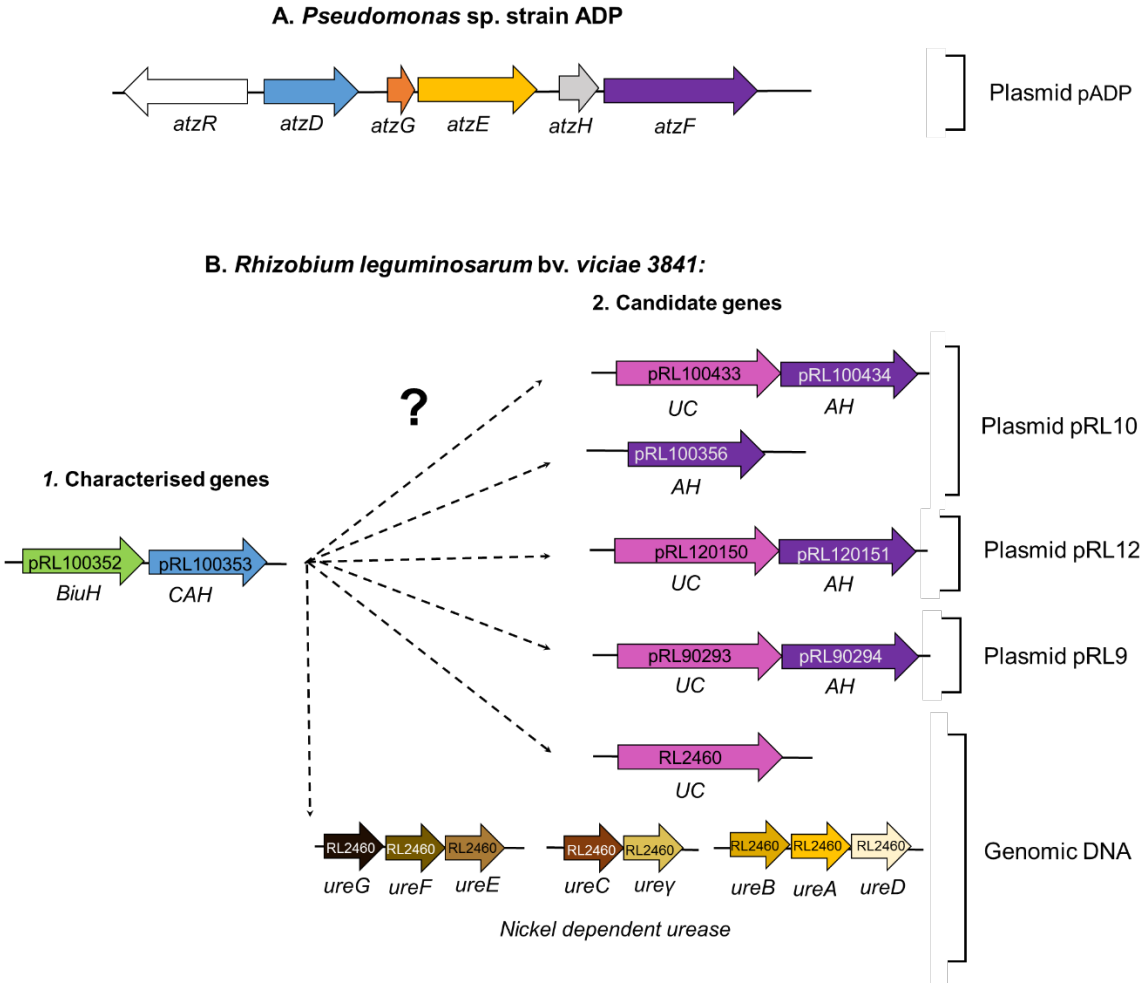


Figure 5-4 Genomic organisation of genes involved in the cyanuric acid degradation pathways. A. *Characterised genes present in* Pseudomonas *sp. strain ADP;* ***B.1*** *Characterised genes present in* Rhizobium leguminosarum *bv.* viciae *3841;* ***B.2.*** *List and location of the genes that could play a role in the degradation of allophanate or urea in* Rhizobium leguminosarum *bv.* viciae *3841. UC: urea carboxylase; AH allophanante hydrolase; CAH; cyanuric acid hydrolase.*

Transcriptomic, metabolomics and proteomic analysis of *Rhizobium leguminosarum* bv. *viciae* 3841 grown on various nitrogen sources (ammonia, cyanuric acid, biuret, urea) could help to determine which enzymes are employed in the cyanuric acid degradation pathway of *Rhizobium leguminosarum* bv. *viciae* 3841.

5.5 Cyanuric acid catabolic pathway: a model to study evolution of new catabolic pathways in bacteria

The cyanuric acid degradation pathway appears to be more diverse than previously thought. At least two different biochemical strategies have been adopted that require different enzymes and have resulted in different genomic organisations. Eventually, the more we understand about the different cyanuric acid catabolic pathway existing across various type of bacteria, the more we will understand about how they have evolved. Cyanuric acid pathways could be used as models to understand more broadly the evolution of catabolic pathways in bacteria. As the discipline of Synthetic Biology matures, it will be essential to have a good understanding of the rules, constraints and design principles for building effective metabolic pathways through effective genetic organisation of their component genes. Ultimately knowledge gained from studying cyanuric acid catabolism could also be applied to building better synthetic metabolic pathways (189).

References

1. **Lewis DE.** 2010. 150 Years of organic structures, p 35-57, Atoms in chemistry: from dalton's predecessors to complex atoms and beyond, vol 1044.
2. **Wirpsza Z.** 1996. Achievements in the use of melamine as a chemical raw material . Part 1. General review and melamine-based coating resins and coatings. Polimery **41:**265-276.
3. **Bachmann WE, Sheehan JC.** 1949. A new method of preparing the high explosive RDX. J Am Chem Soc **71:**1842-1845.
4. **Canelli E.** 1974. Chemical, bacteriological, and toxicological properties of cyanuric acid and chlorinated isocyanurates as applied to swimming pool disinfection - review. Am J Public Health **64:**155-162.
5. **Bojar HP, Walter F, Baumgartner J.** 2017. Joanneumite, Cu(C3N3O3H2)(2)(NH3)(2), a new mineral from Pabellon de Pica, Chile and the crystal structure of its synthetic analogue. Mineral Mag **81:**155-166.
6. **Dherin C, Gasparutto D, O'Connor TR, Cadet J, Boiteux S.** 2004. Excision by the human methylpurine DNA N-glycosylase of cyanuric acid, a stable and mutagenic oxidation product of 8-oxo-7,8-dihydroguanine. Int J Radiat Biol **80:**21-27.
7. **Fuerst EP, Norman MA.** 1991. Interactions of herbicides with photosynthetic electron-transport. Weed Sci **39:**458-464.
8. **Shukla A, Devine MD.** 2008. Basis of crop selectivity and weed resistance to triazine herbicides, p 111-118, Triazine herbicides: 50 years revolutionizing agriculture doi:Doi 10.1016/B978-044451167-6.50012-X.
9. **Trebst A.** 2008. The Mode of Action of Triazine Herbicides in Plants, p 101-110, Triazine herbicides: 50 years revolutionizing agriculture doi:Doi 10.1016/B978-044451167-6.50011-8.
10. **Cox JR.** 1962. Triazine derivatives as non-selective herbicides. J Sci Food Agr **13:**99-103.
11. **LeBaron HM, McFarland JE, Burnside OC.** 2008. The triazine herbicides: a milestone in the development of weed control technology, p 1-12, Triazine herbicides: 50 years revolutionizing agriculture doi:Doi 10.1016/B978-044451167-6.50004-0.
12. **United States Environmental protection Agency.** 2017. Pesticides industry sales and usage: 2008-2012 market estimates. https://www.google.com.au/url?sa=t&rct=j&q=&esrc=s&source=web&cd=7&cad=rja&uact=8&ved=0ahUKEwiH8uf10YjaAhWBw7wKHcYmAJkQFgh3MAY&url=https%3A%2F%2Fwww.epa.gov%2Fsites%2Fproduction%2Ffiles%2F2015-08%2Fdocuments%2Fatra_final_0.pdf&usg=AOvVaw0J3j1qOPEdGbVT7m-LcixW Accessed 26 March
13. **Food and Agriculture Organisation of the United Nation.** 2015. FAOSTAT: pesticides use. http://www.fao.org/faostat/en/#data/RP. Accessed 20 Novembre.
14. **Ackerman F.** 2007. The economics of atrazine. Int J Occup Environ Health **13:**437-445.

15. **Udikovic-Kolic N, Scott C, Martin-Laurent F.** 2012. Evolution of atrazine-degrading capabilities in the environment. Appl Microbiol Biotechnol **96:**1175-1189.
16. **Frank R, Sirons GJ.** 1985. Dissipation of atrazine residues from soils. Bull Environ Contam Toxicol **34:**541-548.
17. **DeLorenzo ME, Scott GI, Ross PE.** 2001. Toxicity of pesticides to aquatic microorganisms: a review. Environ Toxicol Chem **20:**84-98.
18. **Dao TH, Lavy TL, Sorensen RC.** 1979. Atrazine degradation and residue distribution in soil. Soil Sci Soc Am J **43:**1129.
19. **Capriel P, Haisch A.** 1983. Persistance of atrazine and its metabolites in soil after a single herbicide application. Z Pflanz Bodenkunde **146:**474-480.
20. **Arias-Estevez M, Lopez-Periago E, Martinez-Carballo E, Simal-Gandara J, Mejuto JC, Garcia-Rio L.** 2008. The mobility and degradation of pesticides in soils and the pollution of groundwater resources. Agr Ecosyst Environ **123:**247-260.
21. **Solomon KR, Baker DB, Richards RP, Dixon DR, Klaine SJ, LaPoint TW, Kendall RJ, Weisskopf CP, Giddings JM, Giesy JP, Hall LW, Williams WM.** 1996. Ecological risk assessment of atrazine in North American surface waters. Environ Toxicol Chem **15:**31-74.
22. **Vonberg D, Vanderborght J, Cremer N, Putz T, Herbst M, Vereecken H.** 2014. 20 Years of long-term atrazine monitoring in a shallow aquifer in western Germany. Water Res **50:**294-306.
23. **Dobson RL, Motlagh S, Quijano M, Cambron RT, Baker TR, Pullen AM, Regg BT, Bigalow-Kern AS, Vennard T, Fix A, Reimschuessel R, Overmann G, Shan Y, Daston GP.** 2008. Identification and characterization of toxicity of contaminants in pet food leading to an outbreak of renal toxicity in cats and dogs. Toxicol Sci **106:**251-262.
24. **Brown CA, Jeong KS, Poppenga RH, Puschner B, Miller DM, Ellis AE, Kang KI, Sum S, Cistola AM, Brown SA.** 2007. Outbreaks of renal failure associated with melamine and cyanuric acid in dogs and cats in 2004 and 2007. J Vet Diagn Invest **19:**525-531.
25. **Boffetta P, Adami HO, Berry SC, Mandel JS.** 2013. Atrazine and cancer: a review of the epidemiologic evidence. Eur J Cancer Prev **22:**169-180.
26. **Hayes TB, Anderson LL, Beasley VR, de Solla SR, Iguchi T, Ingraham H, Kestemont P, Kniewald J, Kniewald Z, Langlois VS, Luque EH, McCoy KA, Munoz-de-Toro M, Oka T, Oliveira CA, Orton F, Ruby S, Suzawa M, Tavera-Mendoza LE, Trudeau VL, Victor-Costa AB, Willingham E.** 2011. Demasculinization and feminization of male gonads by atrazine: consistent effects across vertebrate classes. J Steroid Biochem Mol Biol **127:**64-73.
27. **Infurna R, Levy B, Meng C, Yau E, Traina V, Rolofson G, Stevens J, Barnett J.** 1988. Teratological evaluations of atrazine technical, a triazine herbicide, in rats and rabbits. J Toxicol Environ Health **24:**307-319.
28. **Jowa L, Howd R.** 2011. Should atrazine and related chlorotriazines be considered carcinogenic for human health risk assessment? J Environ Sci Heal C **29:**91-144.
29. **Lenkowski JR, Reed JM, Deininger L, McLaughlin KA.** 2008. Perturbation of organogenesis by the herbicide atrazine in the amphibian Xenopus laevis. Environ Health Perspect **116:**223-230.

30. **Orton F, Tyler CR.** 2015. Do hormone-modulating chemicals impact on reproduction and development of wild amphibians? Biol Rev Camb Philos Soc **90:**1100-1117.
31. **Scialli AR, DeSesso JM, Breckenridge CB.** 2014. Developmental toxicity studies with atrazine and its major metabolites in rats and rabbits. Birth Defects Res B Dev Reprod Toxicol **101:**199-214.
32. **Oettmeier W.** 1999. Herbicide resistance and supersensitivity in photosystem II. Cell Mol Life Sci **55:**1255-1277.
33. **Traba HM, Dominguez-Morueco N, Barreno E, Catala M.** 2017. Lichen microalgae are sensitive to environmental concentrations of atrazine. J Environ Sci Health B **52:**223-228.
34. **Wood RJ, Mitrovic SM, Lim RP, Kefford BJ.** 2016. How benthic diatoms within natural communities respond to eight common herbicides with different modes of action. Sci Total Environ **557-558:**636-643.
35. **Vandermeer JR, Devos WM, Harayama S, Zehnder AJB.** 1992. Molecular mechanisms of genetic adaptation to xenobiotic compounds. Microbiol Rev **56:**677-694.
36. **Russell RJ, Scott C, Jackson CJ, Pandey R, Pandey G, Taylor MC, Coppin CW, Liu JW, Oakeshott JG.** 2011. The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. Evol Appl **4:**225-248.
37. **Copley SD.** 2000. Evolution of a metabolic pathway for degradation of a toxic xenobiotic: the patchwork approach. Trends Biochem Sci **25:**261-265.
38. **Sugrue E, Hartley CJ, Scott C, Jackson CJ.** 2016. The evolution of new catalytic mechanisms for xenobiotic hydrolysis in bacterial metalloenzymes. Aust J Chem **69:**1383-1395.
39. **Shapir N, Mongodin EF, Sadowsky MJ, Daugherty SC, Nelson KE, Wackett LP.** 2007. Evolution of catabolic pathways: genomic insights into microbial s-triazine metabolism. J Bacteriol **189:**674-682.
40. **Burnside OC, Schmidt EL, Behrens R.** 1961. Dissipation of simazine from the sol. Weeds **9:**477.
41. **Armstrong DE, Chesters G, Harris RF.** 1967. Atrazine hydrolysis in soil. Soil Sci Soc Am Pro **31:**61-67.
42. **Skipper HD, Gilmour CM, Furtick WR.** 1967. Microbial versus chemical degradation of atrazine in soils. Soil Sci Soc Am Pro **31:**653-654.
43. **Nagy I, Compernolle F, Ghys K, Vanderleyden J, Demot R.** 1995. A single cytochrome-P-450 system is involved in degradation of the herbicides eptc (S-ethyl dipropylthiocarbamate) and atrazine by *Rhodococcus* sp. strain Ni86/21. Appl Environ Microbiol **61:**2056-2060.
44. **Cook AM, Hutter R.** 1981. *s*-Triazines as nitrogen-sources for bacteria. J Agr Food Chem **29:**1135-1143.
45. **Cook AM, Hutter R.** 1984. Deethylsimazine - bacterial dechlorination, deamination, and complete degradation. J Agr Food Chem **32:**581-585.
46. **Cook AM, Beilstein P, Grossenbacher H, Hutter R.** 1985. Ring cleavage and degradative pathway of cyanuric acid in bacteria. Biochem J **231:**25-30.
47. **Eaton RW, Karns JS.** 1991. Cloning and comparison of the DNA encoding ammelide aminohydrolase and cyanuric acid amidohydrolase from three *s*-triazine-degrading bacterial strains. J Bacteriol **173:**1363-1366.

48. **Eaton RW, Karns JS.** 1991. Cloning and analysis of *s*-triazine catabolic genes from *Pseudomonas* sp. strain NRRLB-12227. J Bacteriol **173:**1215-1222.
49. **Yanzekontchou C, Gschwind N.** 1994. Mineralization of the herbicide atrazine as a carbon source by a *Pseudomonas* strain. Appl Environ Microbiol **60:**4297-4302.
50. **Mandelbaum RT, Allan DL, Wackett LP.** 1995. Isolation and characterization of a *Pseudomonas* sp. that mineralizes the *s*-triazine herbicide atrazine. Appl Environ Microbiol **61:**1451-1457.
51. **Kaufman DD, Kearney PC, Sheets TJ.** 1963. Simazine: degradation by soil microorganisms. Science **142:**405-406.
52. **Kaufman DD, Blake J.** 1970. Degradation of atrazine by soil fungi. Soil Biology and Biochemistry **2:**73-80.
53. **Giardina MC, Giardi MT, Filacchioni G.** 1982. Atrazine metabolism by *Nocardia* - elucidation of initial pathway and synthesis of potential metabolites. Agr Biol Chem **46:**1439-1445.
54. **Behki RM, Khan SU.** 1986. Degradation of atrazine by *Pseudomonas* - N-dealkylation and dehalogenation of atrazine and its metabolites. J Agr Food Chem **34:**746-749.
55. **Mirgain I, Green GA, Monteil H.** 1993. Degradation of atrazine in laboratory microcosms - isolation and identification of the biodegrading bacteria. Environ Toxicol Chem **12:**1627-1634.
56. **Behki RM, Khan SU.** 1994. Degradation of atrazine, propazine, and simazine by *Rhodococcus* strain B-30. J Agr Food Chem **42:**1237-1241.
57. **Bouquard C, Ouazzani J, Prome J, Michel-Briand Y, Plesiat P.** 1997. Dechlorination of atrazine by a *Rhizobium* sp. isolate. Appl Environ Microbiol **63:**862-866.
58. **Shelton DR, Karns JS, McCarty GW, Durham DR.** 1997. Metabolism of melamine by *Klebsiella terragena*. Appl Environ Microbiol **63:**2832-2835.
59. **Karns JS.** 1999. Gene sequence and properties of an *s*-triazine ring-cleavage enzyme from *Pseudomonas* sp strain NRRLB-12227. Appl Environ Microbiol **65:**3512-3517.
60. **Topp E, Zhu H, Nour SM, Houot S, Lewis M, Cuppels D.** 2000. Characterization of an atrazine-degrading *Pseudaminobacter* sp. isolated from Canadian and French agricultural soils. Appl Environ Microbiol **66:**2773-2782.
61. **Topp E, Mulbry WM, Zhu H, Nour SM, Cuppels D.** 2000. Characterization of *s*-triazine herbicide metabolism by a *Nocardioides* sp. isolated from agricultural soils. Appl Environ Microbiol **66:**3134-3141.
62. **Takats Z, Vargha M, Vekey K.** 2001. Investigation of atrazine metabolism in river sediment by high-performance liquid chromatography/mass spectrometry. Rapid Commun Mass Sp **15:**1735-1742.
63. **Kodama T, Ding LX, Yoshida M, Yajima M.** 2001. Biodegradation of an *s*-triazine herbicide, simazine. J Mol Catal B-Enzym **11:**1073-1078.
64. **Rousseaux S, Hartmann A, Soulas G.** 2001. Isolation and characterisation of new Gram-negative and Gram-positive atrazine degrading bacteria from different French soils. Fems Microbiol Ecol **36:**211-222.
65. **Strong LC, Rosendahl C, Johnson G, Sadowsky MJ, Wackett LP.** 2002. *Arthrobacter aurescens* TC1 metabolizes diverse *s*-triazine ring compounds. Appl Environ Microbiol **68:**5973-5980.

66. **Piutti S, Semon E, Landry D, Hartmann A, Dousset S, Lichtfouse E, Topp E, Soulas G, Martin-Laurent F.** 2003. Isolation and characterisation of *Nocardioides* sp. SP12, an atrazine-degrading bacterial strain possessing the gene *trzN* from bulk- and maize rhizosphere soil. FEMS Microbiol Lett **221:**111-117.
67. **Cai B, Han Y, Liu B, Ren Y, Jiang S.** 2003. Isolation and characterization of an atrazine-degrading bacterium from industrial wastewater in China. Lett Appl Microbiol **36:**272-276.
68. **Singh P, Suri CR, Cameotra SS.** 2004. Isolation of a member of *Acinetobacter* species involved in atrazine degradation. Biochem Bioph Res Co **317:**697-702.
69. **Ozawa T, Yoshida R, Wakashiro Y, Hase H.** 2004. Improvement of simazine degradation by inoculation of corn and soybean plants with rhizobacteria. Soil Sci Plant Nutr **50:**1295-1299.
70. **Aislabie J, Bej AK, Ryburn J, Lloyd N, Wilkins A.** 2005. Characterization of *Arthrobacter nicotinovorans* HIM, an atrazine-degrading bacterium, from agricultural soil New Zealand. FEMS Microbiol Ecol **52:**279-286.
71. **Satsuma K.** 2006. Characterisation of new strains of atrazine-degrading *Nocardioides* sp. isolated from Japanese riverbed sediment using naturally derived river ecosystem. Pest Manag Sci **62:**340-349.
72. **El-Sayed WS, El-Baz AF, Othman AM.** 2006. Biodegradation of melamine formaldehyde by *Micrococcus* sp. strain MF-1 isolated from aminoplastic wastewater effluent. Int Biodeter Biodegr **57:**75-81.
73. **Vaishampayan PA, Kanekar PP, Dhakephalkar PK.** 2007. Isolation and characterization of *Arthrobacter* sp. strain MCM B-436, an atrazine-degrading bacterium, from rhizospheric soil. Int Biodeter Biodegr **60:**273-278.
74. **Vibber LL, Pressler MJ, Colores GM.** 2007. Isolation and characterization of novel atrazine-degrading microorganisms from an agricultural soil. Appl Microbiol Biotechnol **75:**921-928.
75. **Devers M, El Azhari N, Kolic NU, Martin-Laurent F.** 2007. Detection and organization of atrazine-degrading genetic potential of seventeen bacterial isolates belonging to divergent taxa indicate a recent common origin of their catabolic functions. FEMS Microbiol Lett **273:**78-86.
76. **Kolic NU, Hrsak D, Kolar AB, Petric I, Stipicevic S, Soulas G, Martin-Laurent F.** 2007. Combined metabolic activity within an atrazine-mineralizing community enriched from agrochemical factory soil. Int Biodeter Biodegr **60:**299-307.
77. **Devers M, Rouard N, Martin-Laurent F.** 2007. Genetic rearrangement of the *atzAB* atrazine-degrading gene cassette from pADP1::Tn5 to the chromosome of *Variovorax* sp. MD1 and MD2. Gene **392:**1-6.
78. **Iwasaki A, Takagi K, Yoshioka Y, Fujii K, Kojima Y, Harada N.** 2007. Isolation and characterization of a novel simazine-degrading beta-proteobacterium and detection of genes encoding *s*-triazine-degrading enzymes. Pest Manag Sci **63:**261-268.
79. **Hernandez M, Villalobos P, Morgante V, Gonzalez M, Reiff C, Moore E, Seeger M.** 2008. Isolation and characterization of a novel simazine-degrading bacterium from agricultural soil of central Chile, *Pseudomonas* sp. MHP41. FEMS Microbiol Lett **286:**184-190.

80. **Garcia MH, Morgante V, Perez MA, Biaggini PV, Noe PM, Vergara MG, Pfeiffer MS.** 2008. Novel *s*-triazine-degrading bacteria isolated from agricultural soils of central Chile for herbicide bioremediation. Electron J Biotechn **11**.
81. **Li QY, Li Y, Zhu XK, Cai BL.** 2008. Isolation and characterization of atrazine-degrading *Arthrobacter* sp. AD26 and use of this strain in bioremediation of contaminated soil. J Environ Sci **20:**1226-1230.
82. **Siripattanakul S, Wirojanagud W, McEvoy J, Limpiyakorn T, Khan E.** 2009. Atrazine degradation by stable mixed cultures enriched from agricultural soil and their characterization. J Appl Microbiol **106:**986-992.
83. **Getenga Z, Dorfler U, Iwobi A, Schmid M, Schroll R.** 2009. Atrazine and terbuthylazine mineralization by an *Arthrobacter* sp. isolated from a sugarcane-cultivated soil in Kenya. Chemosphere **77:**534-539.
84. **Arbeli Z, Fuentes C.** 2010. Prevalence of the gene *trzN* and biogeographic patterns among atrazine-degrading bacteria isolated from 13 Colombian agricultural soils. FEMS Microbiol Ecol **73:**611-623.
85. **Satsuma K.** 2010. Mineralization of *s*-triazine herbicides by a newly isolated *Nocardioides* species strain DN36. Appl Microbiol Biotechnol **86:**1585-1592.
86. **Sajjaphan K, Heepngoen P, Sadowsky MJ, Boonkerd N.** 2010. *Arthrobacter* sp. strain KU001 isolated from a Thai soil degrades atrazine in the presence of inorganic nitrogen sources. J Microbiol Biotechnol **20:**602-608.
87. **Yang CY, Li Y, Zhang K, Wang X, Ma CQ, Tang HZ, Xu P.** 2010. Atrazine degradation by a simple consortium of *Klebsiella* sp. A1 and *Comamonas* sp. A2 in nitrogen enriched medium. Biodegradation **21:**97-105.
88. **Zhang Y, Jiang Z, Cao B, Hu M, Wang ZG, Dong XN.** 2011. Metabolic ability and gene characteristics of Arthrobacter sp strain DNS10, the sole atrazine-degrading strain in a consortium isolated from black soil. International Biodeterioration & Biodegradation **65:**1140-1144.
89. **El Sebai T, Devers-Lamrani M, Changey F, Rouard N, Martin-Laurent F.** 2011. Evidence of atrazine mineralization in a soil from the Nile Delta: Isolation of *Arthrobacter* sp. TES6, an atrazine-degrading strain. Int Biodeter Biodegr **65:**1249-1255.
90. **Dodge AG, Wackett LP, Sadowsky MJ.** 2012. Plasmid localization and organization of melamine degradation genes in *Rhodococcus* sp. strain Mel. Appl Environ Microbiol **78:**1397-1403.
91. **Takagi K, Fujii K, Yamazaki K, Harada N, Iwasaki A.** 2012. Biodegradation of melamine and its hydroxy derivatives by a bacterial consortium containing a novel *Nocardioides* species. Appl Microbiol Biot **94:**1647-1656.
92. **Fajardo C, Sacca ML, Gibello A, Martinez-Inigo MJ, Nande M, Lobo C, Martin M.** 2012. Assessment of *s*-triazine catabolic potential in soil bacterial isolates applying *atz* genes as functional biomarkers. Water Air Soil Poll **223:**3385-3392.
93. **Ngigi AN, Getenga ZM, Boga HI, Ndalut PK.** 2012. Biodegradation of *s*-triazine herbicide atrazine by *Enterobacter cloacae* and *Burkholderia cepacia* sp. from long-term treated sugarcane-cultivated soils in Kenya. J Environ Sci Health B **47:**769-778.

94. **Wang QF, Xie SG.** 2012. Isolation and characterization of a high-efficiency soil atrazine-degrading Arthrobacter sp strain. International Biodeterioration & Biodegradation **71:**61-66.
95. **Fernandez LA, Valverde C, Gomez MA.** 2013. Isolation and characterization of atrazine-degrading Arthrobacter sp strains from Argentine agricultural soils. Annals of Microbiology **63:**207-214.
96. **Omotayo AE, Ilori MO, Radosevich M, Amund OO.** 2013. Metabolism of Atrazine in Liquid Cultures and Soil Microcosms by Nocardioides Strains Isolated from a Contaminated Nigerian Agricultural Soil. Soil & Sediment Contamination **22:**365-375.
97. **Sneha S, Pooja B, Yadav TC, Asifa Q, Anshuman K, Purohit HJ, Atya K.** 2014. Draft genome sequence of atrazine-utilizing bacteria isolated from Indian agricultural soil. Genome Announc **2:**e01149-01113.
98. **Wang J, Zhu L, Wang Q, Wang J, Xie H.** 2014. Isolation and characterization of atrazine mineralizing *Bacillus subtilis* strain HB-6. PLoS One **9:**e107270.
99. **Guo QW, Zhang JX, Wan R, Xie SG.** 2014. Impacts of carbon sources on simazine biodegradation by Arthrobacter strain SD3-25 in liquid culture and soil microcosm. International Biodeterioration & Biodegradation **89:**1-6.
100. **Mesquini JA, Sawaya AC, Lopez BG, Oliveira VM, Miyasaka NR.** 2015. Detoxification of Atrazine by Endophytic Streptomyces sp. Isolated from Sugarcane and Detection of Nontoxic Metabolite. Bull Environ Contam Toxicol **95:**803-809.
101. **Hatakeyama T, Takagi K, Yamazaki K, Sakakibara F, Ito K, Takasu E, Naokawa T, Fujii K.** 2015. Mineralization of melamine and cyanuric acid as sole nitrogen source by newly isolated Arthrobacter spp. using a soil-charcoal perfusion method. World J Microbiol Biotechnol **31:**785-793.
102. **Douglass JF, Radosevich M, Tuovinen OH.** 2016. Biomineralization of atrazine and analysis of 16S rRNA and catabolic genes of atrazine degraders in a former pesticide mixing site and a machinery washing area. Journal of Soils and Sediments **16:**2263-2274.
103. **Ye JY, Zhang JB, Gao JG, Li HT, Liang D, Liu RM.** 2016. Isolation and characterization of atrazine-degrading strain Shewanella sp. YJY4 from cornfield soil. Lett Appl Microbiol **63:**45-52.
104. **Ma LM, Chen SS, Yuan J, Yang PP, Liu Y, Stewart K.** 2017. Rapid biodegradation of atrazine by *Ensifer* sp. strain and its degradation genes. Int Biodeter Biodegr **116:**133-140.
105. **Zhao X, Wang L, Ma F, Bai S, Yang J, Qi S.** 2017. *Pseudomonas* sp. ZXY-1, a newly isolated and highly efficient atrazine-degrading bacterium, and optimization of biodegradation using response surface methodology. J Environ Sci **54:**152-159.
106. **Yang XY, Wei HY, Zhu CX, Geng B.** 2018. Biodegradation of atrazine by the novel *Citricoccus* sp. strain TT3. Ecotox Environ Safe **147:**144-150.
107. **deSouza ML, Sadowsky MJ, Wackett LP.** 1996. Atrazine chlorohydrolase from *Pseudomonas* sp. strain ADP: gene sequence, enzyme purification, and protein characterization. J Bacteriol **178:**4894-4900.
108. **BoundyMills KL, deSouza ML, Mandelbaum RT, Wackett LP, Sadowsky MJ.** 1997. The *atzB* gene of *Pseudomonas* sp. strain ADP

encodes the second enzyme of a novel atrazine degradation pathway. Appl Environ Microbiol **63:**916-923.

109. **Sadowsky MJ, Tong ZK, de Souza M, Wackett LP.** 1998. AtzC is a new member of the amidohydrolase protein superfamily and is homologous to other atrazine-metabolizing enzymes. J Bacteriol **180:**152-158.
110. **Martinez B, Tomkins J, Wackett LP, Wing R, Sadowsky MJ.** 2001. Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. J Bacteriol **183:**5684-5697.
111. **de Souza ML, Seffernick J, Martinez B, Sadowsky MJ, Wackett LP.** 1998. The atrazine catabolism genes *atzABC* are widespread and highly conserved. J Bacteriol **180:**1951-1954.
112. **Seffernick JL, Aleem A, Osborne JP, Johnson G, Sadowsky MJ, Wackett LP.** 2007. Hydroxyatrazine N-ethylaminohydrolase (AtzB): an amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. J Bacteriol **189:**6989-6997.
113. **Seffernick JL, Erickson JS, Cameron SM, Cho S, Dodge AG, Richman JE, Sadowsky MJ, Wackett LP.** 2012. Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. J Bacteriol **194:**4579-4588.
114. **Fruchey I, Shapir N, Sadowsky MJ, Wackett LP.** 2003. On the origins of cyanuric acid hydrolase: purification, substrates, and prevalence of AtzD from *Pseudomonas* sp. strain ADP. Appl Environ Microbiol **69:**3653-3657.
115. **Shapir N, Sadowsky MJ, Wackett LP.** 2005. Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp. strain ADP. J Bacteriol **187:**3731-3738.
116. **Cheng G, Shapir N, Sadowsky MJ, Wackett LP.** 2005. Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism. Appl Environ Microbiol **71:**4437-4445.
117. **Sagarkar S, Bhardwaj P, Storck V, Devers-Lamrani M, Martin-Laurent F, Kapley A.** 2016. *s*-Triazine degrading bacterial isolate *Arthrobacter* sp. AK-YN10, a candidate for bioaugmentation of atrazine contaminated soil. Appl Microbiol Biotechnol **100:**903-913.
118. **Cameron SM, Durchschein K, Richman JE, Sadowsky MJ, Wackett LP.** 2011. A new family of biuret hydrolases involved in *s*-triazine ring metabolism. ACS Catal **2011:**1075-1082.
119. **Desouza ML, Wackett LP, Boundymills KL, Mandelbaum RT, Sadowsky MJ.** 1995. Cloning, characterization, and expression of a gene region from *Pseudomonas* sp. strain ADP involved in the dechlorination of atrazine. Appl Environ Microbiol **61:**3373-3378.
120. **Shao ZQ, Seffens W, Mulbry W, Behki RM.** 1995. Cloning and expression of the *s*-triazine hydrolase gene (*trza*) from *Rhodococcus-corallinus* and development of *Rhodococcus* recombinant strains capable of dealkylating and dechlorinating the herbicide atrazine. J Bacteriol **177:**5748-5755.
121. **de Souza ML, Newcombe D, Alvey S, Crowley DE, Hay A, Sadowsky MJ, Wackett LP.** 1998. Molecular basis of a bacterial consortium: interspecies catabolism of atrazine. Appl Environ Microbiol **64:**178-184.

122. **Mulbry WW, Zhu H, Nour SM, Topp E.** 2002. The triazine hydrolase gene *trzN* from *Nocardioides* sp. strain C190: cloning and construction of gene-specific primers. FEMS Microbiol Lett **206:**75-79.
123. **Garcia-Gonzalez V, Govantes F, Shaw LJ, Burns RG, Santero E.** 2003. Nitrogen control of atrazine utilization in *Pseudomonas* sp. strain ADP. Appl Environ Microbiol **69:**6987-6993.
124. **Garcia-Gonzalez V, Govantes F, Porrua O, Santero E.** 2005. Regulation of the *Pseudomonas* sp. strain ADP cyanuric acid degradation operon. J Bacteriol **187:**155-167.
125. **Garcia-Gonzalez V, Govantes F, Hervas A, Canosa I, Porrua O, Santero E.** 2007. Regulation of the atrazine degradative pathway in *Pseudomonas*, p 31-39. *In* Heipieper HJ (ed), Bioremediation of soils contaminated with aromatic compounds, vol 76. Springer, Dordrecht.
126. **Garcia-Gonzalez V, Jimenez-Fernandez A, Hervas AB, Canosa I, Santero E, Govantes F.** 2009. Distinct roles for NtrC and GlnK in nitrogen regulation of the *Pseudomonas* sp. strain ADP cyanuric acid utilization operon. FEMS Microbiol Lett **300:**222-229.
127. **Govantes F, Porrua O, Garcia-Gonzalez V, Santero E.** 2009. Atrazine biodegradation in the lab and in the field: enzymatic activities and gene regulation. Microb Biotechnol **2:**178-185.
128. **Govantes F, Garcia-Gonzalez V, Porrua O, Platero AI, Jimenez-Fernandez A, Santero E.** 2010. Regulation of the atrazine-degradative genes in *Pseudomonas* sp. strain ADP. FEMS Microbiol Lett **310:**1-8.
129. **Sajjaphan K, Shapir N, Wackett LP, Palmer M, Blackmon B, Tomkins J, Sadowsky MJ.** 2004. *Arthrobacter aurescens* TC1 atrazine catabolism genes *trzN*, *atzB*, and *atzC* are linked on a 160-kilobase region and are functional in *Escherichia coli*. Appl Environ Microbiol **70:**4402-4407.
130. **Mongodin EF, Shapir N, Daugherty SC, DeBoy RT, Emerson JB, Shvartzbeyn A, Radune D, Vamathevan J, Riggs F, Grinberg V, Khouri H, Wackett LP, Nelson KE, Sadowsky MJ.** 2006. Secrets of soil survival revealed by the genome sequence of *Arthrobacter aurescens* TC1. PLoS Genet **2:**e214.
131. **Young JP, Crossman LC, Johnston AW, Thomson NR, Ghazoui ZF, Hull KH, Wexler M, Curson AR, Todd JD, Poole PS, Mauchline TH, East AK, Quail MA, Churcher C, Arrowsmith C, Cherevach I, Chillingworth T, Clarke K, Cronin A, Davis P, Fraser A, Hance Z, Hauser H, Jagels K, Moule S, Mungall K, Norbertczak H, Rabbinowitsch E, Sanders M, Simmonds M, Whitehead S, Parkhill J.** 2006. The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol **7:**R34.
132. **Kapley A, Sagarkar S, Tanksale H, Sharma N, Qureshi A, Khardenavis A, Purohit HJ.** 2013. Genome sequence of *Alcaligenes* sp. strain HPC1271. Genome Announc **1:**e00235-00212.
133. **Nakatsu CH, Barabote R, Thompson S, Bruce D, Detter C, Brettin T, Han C, Beasley F, Chen WM, Konopka A, Xie G.** 2013. Complete genome sequence of *Arthrobacter* sp. strain FB24. Stand Genomic Sci **9:**106-116.
134. **Devers-Lamrani M, Spor A, Mounier A, Martin-Laurent F.** 2016. Draft genome sequence of *Pseudomonas* sp. strain ADP, a bacterial model for studying the degradation of the herbicide atrazine. Genome Announc **4:**e01733-01715.

135. **Jutzi K, Cook AM, Hutter R.** 1982. The degradative pathway of the s-triazine melamine. The steps to ring cleavage. Biochem J **208:**679-684.
136. **Seffernick JL, McTavish H, Osborne JP, de Souza ML, Sadowsky MJ, Wackett LP.** 2002. Atrazine chlorohydrolase from *Pseudomonas* sp strain ADP is a metalloenzyme. Biochemistry **41:**14430-14437.
137. **Shapir N, Osborne JP, Johnson G, Sadowsky MJ, Wackett LP.** 2002. Purification, substrate range, and metal center of AtzC: the N-isopropylammelide aminohydrolase involved in bacterial atrazine metabolism. J Bacteriol **184:**5376-5384.
138. **Shapir N, Cheng G, Sadowsky MJ, Wackett LP.** 2006. Purification and characterization of TrzF: biuret hydrolysis by allophanate hydrolase supports growth. Appl Environ Microbiol **72:**2491-2495.
139. **Shapir N, Pedersen C, Gil O, Strong L, Seffernick J, Sadowsky MJ, Wackett LP.** 2006. TrzN from *Arthrobacter aurescens* TC1 is a zinc amidohydrolase. J Bacteriol **188:**5859-5864.
140. **Seffernick JL, Reynolds E, Fedorov AA, Fedorov E, Almo SC, Sadowsky MJ, Wackett LP.** 2010. X-ray structure and mutational analysis of the atrazine chlorohydrolase TrzN. J Biol Chem **285:**30606-30614.
141. **Balotra S, Newman J, French NG, Briggs LJ, Peat TS, Scott C.** 2014. Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp. strain ADP. Acta Crystallogr F **70:**310-315.
142. **Balotra S, Newman J, Cowieson NP, French NG, Campbell PM, Briggs LJ, Warden AC, Easton CJ, Peat TS, Scott C.** 2015. X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. Appl Environ Microbiol **81:**470-480.
143. **Cho S, Shi K, Wackett LP, Aihara H.** 2013. Crystallization and preliminary X-ray diffraction studies of cyanuric acid hydrolase from *Azorhizobium caulinodans*. Acta Crystallogr F **69:**880-883.
144. **Cho S, Shi K, Seffernick JL, Dodge AG, Wackett LP, Aihara H.** 2014. Cyanuric acid hydrolase from *Azorhizobium caulinodans* ORS 571: crystal structure and insights into a new class of Ser-Lys dyad proteins. PLoS One **9:**e99349.
145. **Peat TS, Balotra S, Wilding M, French NG, Briggs LJ, Panjikar S, Cowieson N, Newman J, Scott C.** 2013. Cyanuric acid hydrolase: evolutionary innovation by structural concatenation. Mol Microbiol **88:**1149-1163.
146. **Balotra S, Warden AC, Newman J, Briggs LJ, Scott C, Peat TS.** 2015. X-ray structure and mutagenesis studies of the *N*-isopropylammelide isopropylaminohydrolase, AtzC. PLoS One **10:**e0137700.
147. **Peat TS, Newman J, Balotra S, Lucent D, Warden AC, Scott C.** 2015. The structure of the hexameric atrazine chlorohydrolase AtzA. Acta Crystallogr D Biol Crystallogr **71:**710-720.
148. **Bera AK, Aukema KG, Elias M, Wackett LP.** 2017. Structure of the cyanuric acid hydrolase TrzD reveals product exit channel. Sci Rep **7:**45277.
149. **Peat TS, Balotra S, Wilding M, Hartley CJ, Newman J, Scott C.** 2017. High-resolution X-ray structures of two functionally distinct members of the

cyclic amide hydrolase family of toblerone fold enzymes. Appl Environ Microbiol **83:**13.

150. **Devers M, Rouard N, Martin-Laurent F.** 2008. Fitness drift of an atrazine-degrading population under atrazine selection pressure. Environ Microbiol **10:**676-684.
151. **Porrua O, Garcia-Jaramillo M, Santero E, Govantes F.** 2007. The LysR-type regulator AtzR binding site: DNA sequences involved in activation, repression and cyanuric acid-dependent repositioning. Mol Microbiol **66:**410-427.
152. **Shapir N, Rosendahl C, Johnson G, Andreina M, Sadowsky MJ, Wackett LP.** 2005. Substrate specificity and colorimetric assay for recombinant TrzN derived from *Arthrobacter aurescens* TC1. Appl Environ Microbiol **71:**2214-2220.
153. **Robinson SL, Badalamenti JP, Dodge AG, Tassoulas LJ, Wackett LP.** 2018. Microbial biodegradation of biuret: defining biuret hydrolases within the isochorismatase superfamily. Environ Microbiol doi:10.1111/1462-2920.14094.
154. **Mulbry WW.** 1994. Purification and characterization of an inducible *s*-triazine hydrolase from *Rhodococcus-corallinus* Nrrl B-15444r. Appl Environ Microbiol **60:**613-618.
155. **Seibert CM, Raushel FM.** 2005. Structural and catalytic diversity within the amidohydrolase superfamily. Biochemistry **44:**6383-6391.
156. **Jackson CJ, Coppin CW, Carr PD, Aleksandrov A, Wilding M, Sugrue E, Ubels J, Paks M, Newman J, Peat TS, Russell RJ, Field M, Weik M, Oakeshott JG, Scott C.** 2014. 300-Fold increase in production of the Zn2+-dependent dechlorinase TrzN in soluble form via apoenzyme stabilization. Appl Environ Microbiol **80:**4003-4011.
157. **Sugrue E, Carr PD, Scott C, Jackson CJ.** 2016. Active site desolvation and thermostability trade-offs in the evolution of catalytically diverse triazine hydrolases. Biochemistry **55:**6304-6313.
158. **Noor S, Changey F, Oakeshott JG, Scott C, Martin-Laurent F.** 2014. Ongoing functional evolution of the bacterial atrazine chlorohydrolase AtzA. Biodegradation **25:**21-30.
159. **Scott C, Jackson CJ, Coppin CW, Mourant RG, Hilton ME, Sutherland TD, Russell RJ, Oakeshott JG.** 2009. Catalytic improvement and evolution of atrazine chlorohydrolase. Appl Environ Microbiol **75:**2184-2191.
160. **Seffernick JL, Aleem A, Osborne JP, Johnson G, Sadowsky MJ, Wackett LP.** 2007. Hydroxyatrazine N-Ethylaminohydrolase (AtzB): an Amidohydrolase Superfamily Enzyme Catalyzing Deamination and Dechlorination. J Bacteriol **189:**6989-6997.
161. **Sugrue E, Scott C, Jackson CJ.** 2017. Constrained evolution of a bispecific enzyme: lessons for biocatalyst design. Org Biomol Chem **15:**937-946.
162. **Yamazaki K, Fujii K, Iwasaki A, Takagi K, Satsuma K, Harada N, Uchimura T.** 2008. Different substrate specificities of two triazine hydrolases (TrzNs) from *Nocardioides* species. FEMS Microbiol Lett **286:**171-177.
163. **Noor S, Taylor MC, Russell RJ, Jermiin LS, Jackson CJ, Oakeshott JG, Scott C.** 2012. Intramolecular epistasis and the evolution of a new enzymatic function. PLoS One **7:**e39822.

164. **O'Sullivan O, Suhre K, Abergel C, Higgins DG, Notredame C.** 2004. 3DCoffee: combining protein sequences and structures within multiple sequence alignments. J Mol Biol **340:**385-395.
165. **Notredame C, Higgins DG, Heringa J.** 2000. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol **302:**205-217.
166. **Poirot O, Suhre K, Abergel C, O'Toole E, Notredame C.** 2004. 3DCoffee@igs: a web server for combining sequences and structures into a multiple sequence alignment. Nucleic Acids Res **32:**W37-W40.
167. **Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaeli B, Keduas V, Notredame C.** 2006. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-coffee. Nucleic Acids Res **34:**W604-W608.
168. **Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang JM, Taly JF, Notredame C.** 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. Nucleic Acids Res **39:**W13-W17.
169. **Robert X, Gouet P.** 2014. Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res **42:**W320-324.
170. **Seffernick JL, Wackett LP.** 2016. Ancient evolution and recent evolution converge for the biodegradation of cyanuric acid and related triazines. Appl Environ Microbiol **82:**1638-1645.
171. **Soong CL, Ogawa J, Sakuradani E, Shimizu S.** 2002. Barbiturase, a novel zinc-containing amidohydrolase involved in oxidative pyrimidine metabolism. J Biol Chem **277:**7051-7058.
172. **Patricelli MP, Cravatt BF.** 2000. Clarifying the catalytic roles of conserved residues in the amidase signature family. J Biol Chem **275:**19177-19184.
173. **Balotra S, Newman J, Cowieson NP, French NG, Campbell PM, Briggs LJ, Warden AC, Easton CJ, Peat TS, Scott C.** 2015. X-Ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. Appl and Environ Microbiol **81:**470-480.
174. **McKinney MK, Cravatt BF.** 2003. Evidence for distinct roles in catalysis for residues of the serine-serine-lysine catalytic triad of fatty acid amide hydrolase. J Biol Chem **278:**37393-37399.
175. **Shin S, Lee TH, Ha NC, Koo HM, Kim SY, Lee HS, Kim YS, Oh BH.** 2002. Structure of malonamidase E2 reveals a novel Ser-*cis*Ser-Lys catalytic triad in a new serine hydrolase fold that is prevalent in nature. Embo Journal **21:**2509-2516.
176. **Labahn J, Neumann S, Buldt G, Kula MR, Granzin J.** 2002. An alternative mechanism for amidase signature enzymes. J Mol Biol **322:**1053-1064.
177. **Lin Y, St Maurice M.** 2013. The structure of allophanate hydrolase from *Granulibacter bethesdensis* provides insights into substrate specificity in the amidase signature family. Biochemistry **52:**690-700.
178. **Sumrada RA, Cooper TG.** 1982. Urea carboxylase and allophanate hydrolase are components of a multifunctional protein in yeast. J Biol Chem **257:**9119-9127.
179. **Fan C, Li Z, Yin HY, Xiang S.** 2013. Structure and function of allophanate hydrolase. J Biol Chem **288:**21422-21432.

180. **Robinson SL, Badalamenti JP, Dodge AG, Tassoulas LJ, Wackett LP.** 2018. Microbial biodegradation of biuret: defining biuret hydrolases within the isochorismatase superfamily. Environ Microbiol **20**.
181. **Esquirol L, Peat TS, Wilding M, Liu JW, French NG, Hartley CJ, Onagi H, Nebl T, Easton CJ, Newman J, Scott C.** 2018. An unexpected vestigial protein complex reveals the evolutionary origins of an s-triazine catabolic enzyme. J Biol Chem doi:10.1074/jbc.RA118.001996.
182. **Goodacre NF, Gerloff DL, Uetz P.** 2014. Protein domains of unknown function are essential in bacteria. Mbio **5**.
183. **Baek J, Lee J, Yoon K, Lee H.** 2017. Identification of unannotated small genes in *Salmonella*. G3-Genes Genom Genet **7:**983-989.
184. **Bateman A, Coggill P, Finn RD.** 2010. DUFs: families in search of function. Acta Crystallogr F **66:**1148-1152.
185. **Jaroszewski L, Li ZW, Krishna SS, Bakolitsa C, Wooley J, Deacon AM, Wilson IA, Godzik A.** 2009. Exploration of uncharted regions of the protein universe. PLoS Biol **7**.
186. **Lin Y, Boese CJ, St Maurice M.** 2016. The urea carboxylase and allophanate hydrolase activities of urea amidolyase are functionally independent. Protein Sci **25:**1812-1824.
187. **Ray JCJ, Igoshin OA.** 2012. Interplay of gene expression noise and ultrasensitive dynamics affects bacterial operon organization. PLoS Comput Biol **8**.
188. **de Lorenzo V, Sekowska A, Danchin A.** 2015. Chemical reactivity drives spatiotemporal organisation of bacterial metabolism. FEMS Microbiol Rev **39:**96-119.
189. **Yoshikuni Y, Dietrich JA, Nowroozi FF, Babbitt PC, Keasling JD.** 2008. Redesigning enzymes based on adaptive evolution for optimal function in synthetic metabolic pathways. Chem Biol **15:**607-618.

Appendix

DNA and amino acid sequences used in this project. In the following DNA sequences, start codons are highlighted in green and stop codons in red.

AtzD from *Pseudomonas* sp. strain ADP

Reference NCBI of the ADP-1 plasmid complete sequence: NC_004956.1. Loci of *AtzD*: 101053-102144

Gene sequence:

ATGTATCACATCGACGTTTTCCGAATCCCTTGCCACAGCCCTGGTGATACATCGGGTCT
CGAGGATTTGATTGAAACAGGCCGCGTTGCCCCCGCCGACATCGTCGCGGTAATGGGCA
AGACCGAGGGCAATGGCTGCGTCAACGATTACACGCGTGAATACGCCACCGCCATGCTT
GCTGCGTGCCTTGGGCGTCATTTGCAACTCCCACCCCATGAGGTGGAAAAGCGGGTCGC
GTTTGTGATGTCAGGTGGGACGGAAGGCGTGCTGTCCCCCCACCACACGGTATTCGCAA
GACGTCCGGCAATCGACGCGCATCGTCCCGCTGGCAAACGTCTCACGCTTGGAATCGCC
TTCACGCGTGATTTTCTGCCGGAGGAAATTGGCCGCCACGCTCAGATAACGGAGACAGC
CGGCGCCGTCAAACGCGCAATGCGAGATGCCGGGATCGCTTCGATTGACGATCTGCATT
TTGTGCAGGTGAAGTGTCCGCTGCTGACACCAGCAAAGATCGCCTCGGCGCGATCACGC
GGATGCGCTCCAGTCACGACGGATACGTATGAATCGATGGGCTATTCGCGCGGCGCTTC
GGCCCTGGGCATCGCTCTCGCTACAGAAGAGGTGCCCTCCTCGATGCTCGTAGACGAAT
CAGTGCTGAATGACTGGAGTCTCTCATCGTCACTGGCGTCGGCGTCTGCAGGCATCGAA
CTGGAGCACAACGTGGTGATCGCTATTGGCATGAGCGAGCAGGCCACCAGTGAACTGGT
CATTGCCCACGGCGTGATGAGCGACGCGATCGACGCGGCCTCGGTGCGGCGAACGATTG
AATCGCTGGGCATACGTAGCGATGACGAGATGGATCGCATCGTCAACGTATTCGCCAAA
GCGGAGGCGAGCCCGGACGGGGTTGTACGAGGTATGCGGCACACGATGCTAAGTGACTC
CGACATTAATTCGACCCGCCATGCGCGGGCGGTCACCGGCGCGGCCATTGCCTCGGTAG
TTGGGCATGGCATGGTGTATGTGTCCGGTGGCGCCGAGCATCAGGGACCTGCCGGCGGC
GGCCCTTTTGCAGTCATTGCCCGCGCTTAA

Protein sequence:

MYHIDVFRIPCHSPGDTSGLEDLIETGRVAPADIVAVMGKTEGNGCVNDYTREYATAML
AACLGRHLQLPPHEVEKRVAFVMSGGTEGVLSPHHTVFARRPAIDAHRPAGKRLTLGIA
FTRDFLPEEIGRHAQITETAGAVKRAMRDAGIASIDDLHFVQVKCPLLTPAKIASARSR
GCAPVTTDTYESMGYSRGASALGIALATEEVPSSMLVDESVLNDWSLSSSLASASAGIE
LEHNVVIAIGMSEQATSELVIAHGVMSDAIDAASVRRTIESLGIRSDDEMDRIVNVFAK
AEASPDGVVRGMRHTMLSDSDINSTRHARAVTGAAIASVVGHGMVYVSGGAEHQGPAGG
GPFAVIARA

AtzE from *Pseudomonas* sp. strain ADP

Reference NCBI of the ADP-1 plasmid complete sequence: NC_004956.1. Loci of *AtzE*: 102427-103800

Gene sequence:

ATGAAGACAGTAGAAATTATTGAAGGTATCGCCTCTGGCAGAACCAGTGCGCGCGACGT
GTGCGAAGAGGCGCTCGCAACCATCGGCGCGACCGATGGACTCATCAATGCCTTTACAT
GCCGTACGGTTGAACGAGCCCGCGCAGAGGCGGATGCCATCGATGTTCGACGGGCGCGC
GGCGAGGTACTTCCGCCTCTTGCCGGCCTCCCCTACGCGGTAAAGAATCTGTTCGACAT
CGAAGGCGTGACGACGCTTGCCGGCTCGAAGATCAACCGTACTCTCCCGCCTGCGCGCG
CAGACGCCGTGCTGGTGCAACGGCTGAAAGCTGCCGGCGCCGTGCTCCTGGGCGGCCTC
AATATGGACGAGTTTGCCTATGGATTTACGACCGAAAATACGCACTATGGGCCGACCCG
GAACCCGCATGACACCGGGCGTATCGCTGGTGGTTCGTCAGGGGGGTCTGGAGCGGCAA
TCGCTGCGGGGCAGGTACCACTATCGCTCGGATCGGACACCAACGGTTCCATACGCGTG
CCAGCATCATTGTGTGGCGTGTGGGGGCTGAAGCCTACCTTCGGCCGCCTGTCCCGGCG
AGGGACATACCCGTTTGTTCACAGCATTGATCACCTCGGGCCATTGGCCGATAGCGTGG
AAGGCTTGGCGTTGGCCTACGATGCAATGCAGGGCCCGGATCCGCTCGACCCCGGATGC
AGCGCATCGCGCATCCAACCCTCGGTACCGGTCCTCAGTCAGGGTATCGCTGGGCTCCG
GATCGGCGTGCTGGGTGGCTGGTTTCGGGACAATGCCGGCCCGGCCGCGCGAGCCGCGG
TCGATGTTGCCGCGCTTACGCTCGGCGCCAGCGAAGTCGTCATGTGGCCCGACGCGGAG
ATCGGGCGCGCAGCCGCCTTCGTTATCACTGCCAGCGAGGGAGGCTGTCTGCATCTCGA
TGATCTTCGCATCCGTCCGCAAGACTTCGAGCCTCTGTCCGTAGATCGCTTTATCTCGG
GGGTTTTACAACCGGTCGCGTGGTACTTGCGTGCACAGCGGTTTCGACGTGTCTATCGA
GATAAGGTGAATGCTCTTTTCCGTGACTGGGACATATTAATCGCTCCCGCAACGCCAAT
AAGTGCTCCCGCAATCGGCACCGAATGGATCGAGGTAAACGGTACACGCCATCCGTGCC
GCCCGGCTATGGGACTTCTCACTCAGCCGGTCTCCTTCGCAGGCTGTCCGGTGGTCGCC
GCTCCAACGTGGCCTGGAGAAAACGATGGCATGCCGATCGGGGTACAGCTCATCGCGGC
GCCCTGGAACGAATCTCTATGCCTGCGCGCAGGCAAGGTATTACAAGACACCGGTATCG
CCCGACTGAAATGTTAA

Protein sequence:

MKTVEIIEGIASGRTSARDVCEEALATIGATDGLINAFTCRTVERARAEADAIDVRRAR
GEVLPPLAGLPYAVKNLFDIEGVTTLAGSKINRTLPPARADAVLVQRLKAAGAVLLGGL
NMDEFAYGFTTENTHYGPTRNPHDTGRIAGGSSGGSGAAIAAGQVPLSLGSDTNGSIRV
PASLCGVWGLKPTFGRLSRRGTYPFVHSIDHLGPLADSVEGLALAYDAMQGPDPLDPGC
SASRIQPSVPVLSQGIAGLRIGVLGGWFRDNAGPAARAAVDVAALTLGASEVVMWPDAE
IGRAAAFVITASEGGCLHLDDLRIRPQDFEPLSVDRFISGVLQPVAWYLRAQRFRRVYR
DKVNALFRDWDILIAPATPISAPAIGTEWIEVNGTRHPCRPAMGLLTQPVSFAGCPVVA
APTWPGENDGMPIGVQLIAAPWNESLCLRAGKVLQDTGIARLKC

AtzF from *Pseudomonas* sp. strain ADP

Reference NCBI of the ADP-1 plasmid complete sequence: NC_004956.1. Loci of *AtzF*: 104283-106100

Gene sequence:

ATGAATGACCGCGCGCCCCACCCTGAAAGATCTGGTCGAGTCACGCCGGATCACCTGAC
CGATCTGGCTTCCTATCAGGCTGCCTATGCCGCCGGTACAGACGCCGCCGACGTCATTT
CGGACCTGTATGCCCGTATCAAAGAAGACGGCGAAAATCCGATCTGGATTAGCCTGTTG
CCCTTGGAAAGCGCATTGGCGATGCTGGCCGACGCGCAGCAACGCAAGGACAAGGGAGA

AGCGTTGCCGCTCTTTGGCATCCCCTTCGGCGTCAAGGACAACATCGACGTCGCAGGCC
TTCCGACGACTGCCGGGTGTACGGGGTTCGCGCGTACGCCCCGACAGCACGCCTTCGTC
GTACAGCGCCTGGTGGACGCTGGCGCGATCCCGATCGGAAAAACGAACCTCGATCAATT
CGCGACCGGGTTGAACGGCACTCGCACGCCGTTTGGCATTCCGCGCTGCGTGTTCAACG
AGAACTACGTATCCGGCGGCTCCAGCAGTGGCTCCGCAGTGGCCGTCGCCAACGGCACG
GTACCGTTCTCGCTCGGGACGGACACTGCCGGTTCCGGCCGCATTCCTGCTGCGTTCAA
CAATCTGGTGGGCTTGAAACCGACCAAAGGCCTGTTCTCGGGCAGTGGACTGGTTCCCG
CGGCGCGAAGCCTTGACTGCATCAGCGTCCTCGCCCATACCGTAGATGACGCCCTTGCG
GTCGCACGCGTCGCCGCCGGCTACGATGCTGATGACGCTTTTTCGCGCAAGGCGGGCGC
CGCCGCACTGACAGAAAAGAGTTGGCCTCGTCGCTTCAATTTCGGGGTCCCAGCGGCGG
AACATCGCCAGTTTTTCGGTGACGCGGAAGCCGAGGCGCTTTTCAATAAAGCGGTTCGC
AAGCTTGAAGAGATGGGTGGCACCTGCATCTCGTTTGACTATACCCCCTTCAGGCAGGC
TGCTGAACTGCTCTACGCCGGCCCTTGGGTTGCGGAGCGCCTGGCGGCCATCGAGAGCC
TTGCGGACGAGCATCCCGAGGTGCTCCACCCGGTCGTTCGTGACATCATCTTGTCCGCG
AAGCGAATGAGCGCAGTCGACACGTTCAACGGTATCTATCGCCTGGCCGACCTTGTCAG
GGCTGCAGAGAGCACTTGGGAAAAGATCGATGTGATGCTGCTGCCGACGGCGCCGACCA
TCTACACTGTAGAAGACATGCTCGCCGATCCGGTACGCCTCAACAGCAATCTGGGCTTC
TACACGAACTTCGTGAACTTGATGGATTTGTCCGCGATTGCTGTTCCCGCAGGCTTCCG
AACCAATGGCCTGCCATTTGGCGTCACTTTCATCGGTCGGGCGTTCGAAGATGGGGCGA
TCGCAAGCTTGGGAAAAGCTTTCGTGGAGCACGACCTCGCCAAGGGCAACGCGGCCACG
GCGGCGCCACCCAAGGATACCGTCGCAATCGCCGTGGTAGGTGCACATCTCTCCGACCA
GCCCTTGAATCATCAGCTCACGGAGAGCGGCGGAAAGCTACGGGCAACAACGCGTACTG
CGCCGGGATATGCCTTGTACGCACTCCGTGATGCGACGCCGGCTAAGCCTGGAATGTTG
CGCGACCAGAATGCGGTCGGGAGCATCGAAGTGGAAATCTGGGATCTGCCGGTCGCCGG
GTTCGGTGCGTTTGTAAGTGAAATTCCGGCGCCGTTGGGTATCGGGACAATAACACTCG
AAGACGGCAGCCATGTGAAAGGCTTTCTGTGCGAGCCACATGCCATCGAGACGGCGCTC
GACATCACTCACTACGGCGGCTGGCGAGCATACCTCGCGGCTCAATAG

Protein sequence:

MNDRAPHPERSGRVTPDHLTDLASYQAAYAAGTDAADVISDLYARIKEDGENPIWISLL
PLESALAMLADAQQRKDKGEALPLFGIPFGVKDNIDVAGLPTTAGCTGFARTPRQHAFV
VQRLVDAGAIPIGKTNLDQFATGLNGTRTPFGIPRCVFNENYVSGGSSSGSAVAVANGT
VPFSLGTDTAGSGRIPAAFNNLVGLKPTKGLFSGSGLVPAARSLDCISVLAHTVDDALA
VARVAAGYDADDAFSRKAGAAALTEKSWPRRFNFGVPAAEHRQFFGDAEAEALFNKAVR
KLEEMGGTCISFDYTPFRQAAELLYAGPWVAERLAAIESLADEHPEVLHPVVRDIILSA
KRMSAVDTFNGIYRLADLVRAAESTWEKIDVMLLPTAPTIYTVEDMLADPVRLNSNLGF
YTNFVNLMDLSAIAVPAGFRTNGLPFGVTFIGRAFEDGAIASLGKAFVEHDLAKGNAAT
AAPPKDTVAIAVVGAHLSDQPLNHQLTESGGKLRATTRTAPGYALYALRDATPAKPGML
RDQNAVGSIEVEIWDLPVAGFGAFVSEIPAPLGIGTITLEDGSHVKGFLCEPHAIETAL
DITHYGGWRAYLAAQ

AtzG from *Pseudomonas* sp. strain ADP

Reference NCBI of the ADP-1 plasmid complete sequence: NC_004956.1. Loci of *AtzG*: 100645-100851

Gene sequence:

ATGACGGAAACTGAAATCTTCGCCTATATCGAGGCGGCTTCAATAGCGATCGGAATCCC
GCTTGAGCCGGCTCGTGCACGCGCTGTCGCACACCATTTTTCGCGTACAGCACTCCTTG
CAGAAATGCTTGAAAGCGTCCCTCTCTCGCCTGAGTCAGAACTCGCGGAGATTTATCGC
CCGGCGCCCTTCCCCGCAGAAGATATCTGA
Protein sequence:

MTETEIFAYIEAASIAIGIPLEPARARAVAHHFSRTALLAEMLESVPLSPESELAEIYR
PAPFPAEDI

AtzH from *Pseudomonas* sp. strain ADP

Reference NCBI of the ADP-1 plasmid complete sequence: NC_004956.1. Loci

of *AtzH*: 98874-99263

Gene sequence:

ATGCAAATTAATCTACCTGAAGTTCATGCGGAAGTGACCGCCCAATTCGTTCGCTACGA
GAAAGCGCTTACCTCGAATGACACGGCTGTTTTGAATGAGCTGTTTTGGAACAGCCCCC
AAACCCTGCGCTACGGTGCGACCGAGAACCTATACGGCTATGAGGCCATTGCAGGCTTT
CGCGCCACACGCTCTCCAAACAACTTGGAACGCGAGATCGTCCGGACGGTGATTACAAC
TTATGGACACGACTTCGCTACCGCCAACATCGAATTCCGTCGGCTCAGTCATAGTCAGC
TTACGGGCCGGCAGAGCCAAACCTGGATGCGCACGTCGCAGGGTTGGCGCGTGGTAGCC
GCGCACGTAAGTCTTATTGCCCTGCCCGTTTCCTGA
Protein sequence:

MQINLPEVHAEVTAQFVRYEKALTSNDTAVLNELFWNSPQTLRYGATENLYGYEAIAGF
RATRSPNNLEREIVRTVITTYGHDFATANIEFRRLSHSQLTGRQSQTWMRTSQGWRVVA
AHVSLIALPVS

Biuret hydrolase from *Rhizobium leguminasorum* bv. *viciae* 3841

Reference NCBI of the plasmid 10 complete sequence: AM236084.1. Loci of

biuret hydrolase: 363092-363793

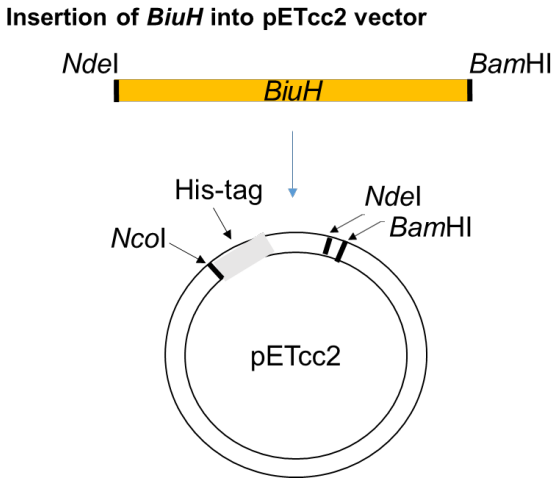
Gene sequence:

ATGGACGCGATGGTCGAAACCAACCGGCATTTTATCGACGCCGATCCGTATCCGTGGCC
CTATAACGGAGCTCTGAGGCCTGACAATACCGCCCTCATCATCATCGACATGCAGACGG
ATTTCTGCGGCAAGGGCGGTTATGTCGACCACATGGGCTACGACCTGTCGCTGGTGCAG
GCGCCGATCGAACCCATCAAACGCGTGCTTGCCGCCATGCGGGCCAAGGGTTATCACAT
CATCCACACCCGCGAGGGCCACCGCCCCGACCTCGCCGATCTGCCAGCAAACAAACGCT
GGCGCTCGCAACGGATCGGGGCCGGCATCGGTGATCCCGGCCCCTGCGGCCGAATCCTG
ACGCGTGGCGAACCCGGCTGGGACATCATCCCCGAACTCTACCCGATCGAAGGCGAGAC
GATCATCGACAAGCCCGGCAAGGGTTCGTTCTGCGCCACCGACCTCGAACTCGTCCTCA
ACCAGAAACGCATCGAGAACATTATCCTCACCGGGATCACCACCGATGTCTGCGTCTCG
ACGACGATGCGCGAGGCGAACGACCGCGGCTACGAATGCCTGCTGCTGGAGGACTGCTG
TGGTGCGACCGACTACGGAAACCACCTCGCCGCCATCAAGATGGTGAAGATGCAGGGCG
GCGTCTTCGGCTCGGTCTCCAATTCCGCGGCTCTAGTCGAGGCGCTGCCCTGA

Protein sequence:

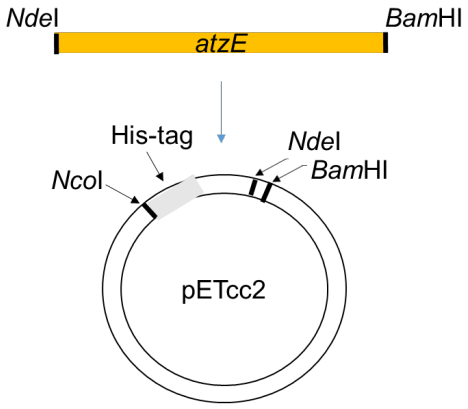
MDAMVETNRHFIDADPYPWPYNGALRPDNTALIIIDMQTDFCGKGGYVDHMGYDLSLVQ
APIEPIKRVLAAMRAKGYHIIHTREGHRPDLADLPANKRWRSQRIGAGIGDPGPCGRIL
TRGEPGWDIIPELYPIEGETIIDKPGKGSFCATDLELVLNQKRIENIILTGITTDVCVS
TTMREANDRGYECLLLEDCCGATDYGNHLAAIKMVKMQGGVFGSVSNSAALVEALP

Cloning strategy used in Chapter 2:



Cloning strategy used in Chapter 3:

1. Insertion of *atzE* into pETcc2 vector



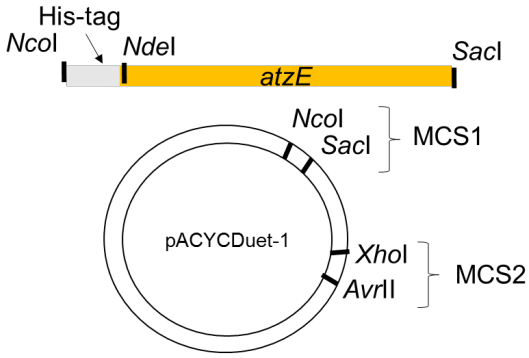
2. Cloning of *atzE* and insertion into the MCS1 using NcoI SacI

NcoI

GTACACCCATGGGAATGAAGACAGTAGAAATTATTGAAGG

SacI Stop codon of *atzE*

TTTTTTGAGCTCCCACATTTCAGTCGGGCGATAC



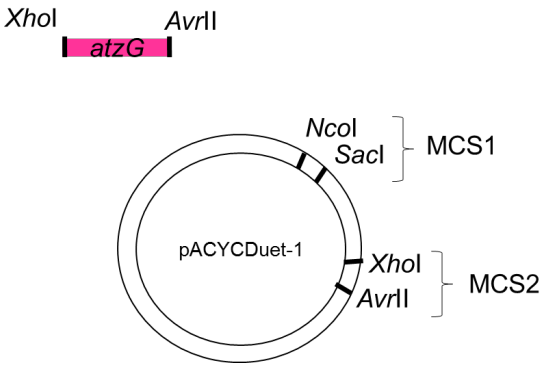
3. Amplification of *atzG* and cloning into MSC2 using *XhoI* *AvrII*

NdeI XhoI

CGACGACATATGCTCGAGATGACGGAAACTG

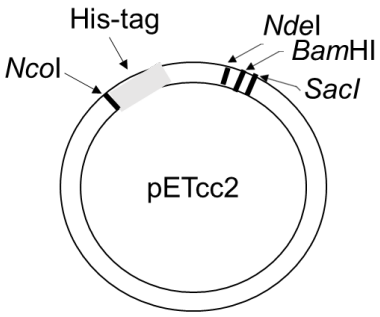
SacI AvrII Stop codon of *atzG*

TGCTGCGAGCTCCCTAGGTCAGATATCTTCTGC



Cloning strategy used in Chapter 4:

1. Insertion of *SacI* into pETcc2 vector



2. Amplification of *atzH* and cloning into pETcc2 vector

NdeI *XhoI*

CGACGACATATGCTCGAGATGCAAATTAATCTACC

SacI *AvrII* Stop codon of *atzH*

TGCTGCGAGCTCCCTAGGTCAGGAAACGGGC