

**A structure-function study of the enzymes of the atrazine
catabolic pathway**

Ph.D. Thesis

SAHIL BALOTRA



THESIS DEDICATION

I dedicate this thesis to my parents,
Mr. Madan Mohan Balotra and Mrs. Rama Balotra,
without whom none of my success would be possible.



**Australian
National
University**



The thesis submitted for the degree of Doctor of Philosophy of

The Australian National University

A Structure-Function Study of the Enzymes of the Atrazine Catabolic Pathway

Sahil Balotra

Research School of Chemistry

The Australian National University

Canberra, ACT, Australia

and

Land and Water Flagship

Commonwealth Scientific and Industrial Research Organisation

Canberra, ACT, Australia

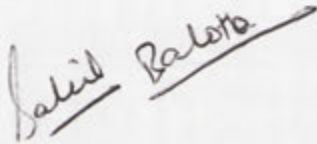
May, 2015

Declaration

This thesis consists of original research work performed by me and to the best of my knowledge it does not have any material or data, except where due acknowledgment is made in the thesis, previously reported by another person for the award of any degree or diploma at the Australian National University or any other university / institution worldwide. I sincerely and fully acknowledge the contributions of my supervisors and colleagues for this completed study.

Name: SAHIL BALOTRA

Signature:

Handwritten signature of Sahil Balotra in black ink, written in a cursive style with a horizontal line underneath.

Date:

30/09/2015
(dd/mm/yyyy)

Acknowledgements

It is a pleasure for me to express my gratitude to my supervisor Dr Colin Scott for his guidance and encouragement throughout my PhD studies. Without his guidance and support this dissertation would not have been possible. I also thank Professor Chris Easton, the chair of my supervisory panel for his support throughout my PhD.

I am also most grateful to Drs Janet Newman and Tom Peat (CSIRO) for screening and obtaining protein crystals and for their painstaking efforts to solve structure from the diffraction data obtained from these crystals.

I am also grateful to Drs Carol Hartley, Robyn Russell, Andrew Warden and especially Dr Matt Wilding for the precious time they spent in discussion with me on various fronts and for their critical thoughts. I also thank Chris Blake at ANU for training me on NMR instrument.

The support and encouragement of my many friends and colleagues at CSIRO has been instrumental to my successful completion of this thesis. In no particular order they are: Peter Hart, Sue Dorrian, Dr Chris Coppin, Dr Nikki Dellas, Dr Chris Greening, Dr Matt Taylor, Nigel French, Dr John Oakeshott, Drs Gunjan and Rinku Pandey, Michelle Williams, Lyndall Briggs, Faisal Younus, Shahana Afroze, Dr Sameer Sharma, Norma Blunden, Dr Jian-Wei Liu, Jane Green, Dr Peter Campbell, Claire Farnsworth, Fiza Mohd Pushri, and Dr Muhammad Tehseen.

I am also very thankful to the ANU and CSIRO for the scholarship and fee sponsorship which they provided me. In particular, I thank Alison Scott at the Research School of Chemistry, ANU and Eva Zinkovsky and Ree Murnane at CSIRO for their help in administrative matters concerning my studies.

Last but not least, indeed I am indebted for ever to my beloved parents (Mr Madan Mohan Balotra and Mrs Rama Balotra), loving sister Mrs Rainy Bhardwaj, devoted and caring wife Mrs Madhura Shettigar, for their love, care and prayers for my success throughout my life. No doubt, my success is actually theirs.

Sahil Balotra

List of Abbreviations

AH	=	Allophanate hydrolase
AH _{Gb}	=	Allophanate hydrolase from <i>Granulibacter bethesdensis</i>
AH _{KL}	=	Allophanate hydrolase from <i>Kluyveromyces Lactis</i>
ATP	=	Adenosine triphosphate
AtzA	=	Atrazine chlorohydrolase
AtzB	=	Hydroxyatrazine <i>N</i> -ethylaminohydrolase
AtzC	=	<i>N</i> -ethylaminohydrolase
AtzD	=	Cyanuric acid hydrolase
AtzE	=	Biuret hydrolase
AtzF	=	Allophanate hydrolase
BA	=	Barbituric acid
BAEE	=	<i>N</i> _α -Benzoyl-L-arginine ethyl ester hydrochloride
BH	=	Barbituric acid hydrolase
bp	=	Base pair
BSA	=	Bovine serum albumin
CA	=	Cyanuric acid
CH	=	Cyanuric acid hydrolase
CHES	=	<i>N</i> -Cyclohexyl-2-aminoethanesulphonic acid
DFT	=	Differential functional theory
DLS	=	Dynamic light scattering
DMSO	=	Dimethylsulfoxide
DNA	=	Deoxyribonucleic acid
dNTP	=	Deoxy nucleotide triphosphate
DOPA	=	3,4-Dihydroxyphenylalanine

DTT	=	Dithiothreitol
EDTA	=	Ethylenediaminetetraacetic acid
FPLC	=	Fast protein liquid chromatography
GDH	=	Glutamate dehydrogenase
HEPES	=	2-[4-(2-hydroxyethyl)piperazin-1-yl] ethanesulfonic acid
IPA	=	Isopropylamine
IPTG	=	Isopropyl β -D-1-thiogalactopyranoside
k_{cat}	=	Catalysis constant
K_d	=	Dissociation constant
kDa	=	Kilo Dalton
K_M	=	Michaelis constant
LB	=	Luria Bertani
LC/MS	=	Liquid chromatography/mass spectrometry
MES	=	2-(N-morpholino) ethanesulphonic acid
MOPS	=	[3-(N-morpholino) propanesulphonic acid]
NAD	=	Nicotinamide adenine dinucleotide
NAD(P)	=	Nicotinamide adenine dinucleotide phosphate
NEB	=	New England Biolabs
NMR	=	Nuclear magnetic resonance
OD₆₀₀	=	Optical density at 600 nM
PCR	=	Polymerase chain reaction
PDB	=	Protein data bank
PEG	=	Poly ethylene glycol
PISA	=	Protein Interfaces, surfaces and assemblies
PMSF	=	Phenyl methane sulfonyl fluoride
ppb	=	Parts per billion

RDX	=	Royal demolition explosive
RMSD	=	Root mean square deviation
RNA	=	Ribonucleic acid
RT-PCR	=	Real time polymerase reaction
RU	=	Repeating units
SAXS	=	Small angle X-ray scattering
SDS-PAGE	=	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	=	Size exclusion chromatography
TAPS	=	(N-Tris(hydroxymethyl)methyl-3-aminopropanesulphonic acid)
TriA	=	Melamine deaminase

List of publications generated from the research presented herein:

1. Peat, T.S., Balotra, S., Wilding, M., French, N.G., Briggs, L.J., Panjikar, S., Cowieson, N., Newman, J., and Scott, C., Cyanuric acid hydrolase: Evolutionary innovation by structural concatenation. *Molecular Microbiology*, 2013. 88: 1149-1163.
2. Balotra, S., Newman, J., French, N.G., Briggs, L.J., Peat, T.S., and Scott, C., Crystallization and preliminary x-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp strain ADP. *Acta Crystallographica Section F-Structural Biology Communications*, 2014. 70: 310-315.
3. 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., X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. *Applied and Environmental Microbiology*, 2015. 81: 470-480.
4. Peat, T.S., Newman, J., Balotra, S., Lucent, D., Warden, A.C., and Scott, C., The structure of the hexameric atrazine chlorohydrolase AtzA. *Acta Crystallographica Section D-Biological Crystallography*, 2015. 71: 710-720.
5. Balotra, S., Newman, J., French, N.G., Briggs, L. J., Christopher J. E., Peat, T. S., and Scott, C. The reaction mechanism of the *N*-isopropylammelide isopropylaminohydrolase AtzC: insights from structural and mutagenesis studies. Submitted to PLoS One for review.

List of publications not part of this thesis

1. Shettigar, M., Pearce, S., Pandey, R., Khan, F., Dorrian, S.J., Balotra, S., Russell, R.J., Oakshott, J.G., and Pandey, G., Cloning of a novel 6-chloronicotinic acid chlorohydrolase from the newly isolated 6-chloronicotinic acid mineralizing *Bradyrhizobiaceae* strain SG-6C. *PLoS One*, 2012. 7: e51162

List of publications in preparation from research not presented herein:

1. Expanding the physiological roles of the ser-lys cyclic amidases of the Toblerone fold family
2. Molecular dynamics of active site variants of the Cyanuric acid hydrolase, AtzD.
3. Random walks through sequence space with intermediate forms of the atrazine chlorohydrolase, AtzA and melamine deaminase, TriA.
4. Structural insights into the evolution of the triazine-catabolism pathway (review)

Table of Contents

Abstract	1
1 Introduction	
1.1 The evolution of new enzyme function.....	2
1.2 New metabolic functions have evolved in response to the use of herbicidal <i>s</i> -triazines.....	5
1.3 Bioremediation of <i>s</i> -triazines.....	17
1.4 Aim of the thesis.....	19
2 The structure of the hexameric atrazine chlorohydrolase AtzA	31-49
3 Purification and crystallographic studies with AtzB (hydroxyatrazine hydrolase).	
3.1 Introduction.....	52
3.2 Materials and methods.....	54
3.2.1 Cloning	54
3.2.2 Expression and purification of AtzB.....	54
3.2.3 Stability analysis.....	56
3.2.4 Crystallisation.....	56
3.3 Results and discussion.....	56
3.3.1 Expression and purification of AtzB	56
3.3.2 Stability analysis.....	57
3.3.3 Crystallographic trials	57
3.4 Conclusion and future work	58
4 X-ray structure and mutagenesis studies of the <i>N</i>-isopropylammelide isopropylaminohydrolase, AtzC.....	61-87
5 Cyanuric acid hydrolase: evolutionary innovation by structural concatenation.....	88-115

6	X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex.....	116-139
7	Discussion	
7.1	Studies with the enzymes of atrazine mineralisation pathway.....	140
7.2	AtzA and the limits of evolutionary innovation.....	141
7.3	Partner swapping and domain recruitment as engines of innovation	143
7.4	The role of metal-specificity in the evolution of new enzyme function.....	144
8.	Appendices	
8.1	Sequences of genes used in this study.....	149
8.2	Amino acid sequences of the enzymes investigated in this study.....	155
8.3	List of PDBe entries from this study.....	158
8.4	List of known inhibitors and substrates for the enzymes of the atrazine catabolic pathway.....	160

Abstract

The aim of this study was to develop an understanding of the structure-function relationship of the enzymes from atrazine mineralization pathway from *Pseudomonas* sp. strain ADP. This pathway has evolved recently in response to the selection pressure exerted by the extensive use of herbicide atrazine over the past six decades. The pathway is comprised of six hydrolytic enzymes, some of which have recently evolved to possess novel functionalities making them excellent models for studying enzyme evolution.

A literature review is presented in chapter 1 with the focus on how the enzyme evolution works in general and how the atrazine mineralizing enzymes have evolved and assembled into a catabolic pathway that achieves the complete mineralization of atrazine.

In chapter 2, the X-ray crystal structure of AtzA, the 'archetypal' atrazine chlorohydrolase is presented. The active site of AtzA is studied by structural and *in silico* docking studies and the comparisons were drawn with the alternative and superior metal-dependent atrazine chlorohydrolase, TrzN. On the basis of structural and previously obtained biochemical data, a plausible reaction mechanism is also presented.

Chapter 3 is about the over-expression, purification and attempts made so far to obtain the X-ray structure of AtzB (hydroxyatrazine hydrolase). This enzyme performs the second step in atrazine mineralisation pathway by removing the alkylamine chain from hydroxyatrazine.

In chapter 4, the X-ray structure of AtzC with bound inhibitor is presented. This enzyme perform the third step in this pathway by removing the isopropylamine chain from the *N*-isopropylammelide. The presence of bound inhibitor in the active site guided the *in silico* substrate docking experiments. The mutagenesis work was performed in the light of substrate interactions information obtained from the docking experiments which allowed us to propose a plausible reaction mechanism for this enzyme.

Chapter 5 describes the structure of AtzD, a cyanuric acid hydrolase. The structure represents a novel fold and is termed as the 'toblerone' fold. The structure of AtzD suggests that this novel fold has evolved by the structural concatenation of monomers of the trimeric YjgF superfamily and acquisition of metal binding site. Based on the crystallographic and biochemical data, a plausible reaction mechanism is also proposed. The residues imparting substrate specificity were also identified by making comparisons with a homology model of related barbiturase and the phylogeny of the 68 AtzD-like enzymes in the database were analysed based on this structure-function relationship.

In chapter 6, the X-ray structure of the amidase domain of allophanate hydrolase, AtzF, from *Pseudomonas* sp ADP was presented. The structure was compared with the structures of allophanate hydrolases from other organisms. We have also provided the preliminary evidence to propose that AtzF forms a large, *ca.* 660 kDa multi-enzyme complex with other two enzymes of cyanuric acid mineralisation pathway.

Chapter 1: Introduction

1.1 The evolution of new enzyme function

The theory of evolution, described by Charles Darwin in 1859 [1], teaches that the characteristics that are best suited to their current environmental conditions tend to out-survive less-well suited characteristics and, over the course of generations, will tend to dominate within a population. However, while this theory convincingly explains the changes in frequency of genetic characters within a population, it does not adequately explain how entirely new characteristics (such as new enzyme function) are acquired in the first place.

Functional changes in proteins are caused by amino acid substitutions, insertions or deletions [2, 3]. Although significant quantitative or qualitative functional changes in an enzyme can be conferred by as little as a single amino acid substitution, it often requires multiple amino acid substitutions for a new catalytic activity to become efficient enough to be physiologically relevant [4-9]. In addition, mutations that confer new catalytic functions upon an enzyme generally reduce its capacity to perform its original physiological role [4, 10-14]. This trade-off between novel and pre-existing functions will most likely result in a decreased fitness of the organism, and so mutations that lead to new enzymatic functions will be out-competed [9, 15-17]. The situation is apparently paradoxical; new enzymatic functions have arisen, but the process of gaining new functionality results in a reduced fitness that should ensure that mutations that confer new enzymatic functions are lost. The

question is not how are new enzymatic function acquired; but rather, how are new enzymatic functions retained?

Susumo Ohno proposed one potential solution [18]; gene duplication can result in two functional copies of a gene, one of which can retain the original function while the second copy acquires random mutations at no cost to the fitness of the organism [18, 19]. Comparative genomic analysis has revealed that gene duplication is indeed widespread, accounting for the presence of > 50% of bacterial genes and > 90% eukaryotic genes [20-24]. However, there is a fitness cost in maintaining two copies of a gene, because of which duplicate genes are often lost *via* a process known as purifying selection [25-27]. Given that the frequency of acquiring a beneficial mutation that confers a new enzymatic function is low ($< 2 \times 10^{-3}$) [27], most duplicate genes would be lost to purifying selection before they acquire a mutation that confers a new catalytic activity that provides a fitness advantage. Therefore, the acquisition of new enzyme function cannot be explained by Ohno's solution alone.

Enzymes frequently possess 'promiscuous' activities, i.e., low-level catalytic activities towards physiologically innocuous substrates [28, 29]. Many examples are available to support that the functional promiscuity is a common feature of proteins [28, 30-33] (Table 1).

Table 1. Examples of functional promiscuity in enzymes

Enzyme	Native activity	Promiscuous activity	Source
Dihydroorotase	Dihydroorotic acid hydrolysis	Phosphotriesterase	[34, 35]
Malonate semialdehyde decarboxylase	Decarboxylation	Hydration	[36]
<i>O</i> -succinylbenzoate synthase	Dehydration	<i>N</i> -acyl amino acid racemase	[37, 38]
3' keto L-gluconate 6-phosphate decarboxylase	Decarboxylation	Aldol condensation	[39]
C-C hydrolase MbpC	C-C bond cleavage	Esterase	[40]
Tetrachlorohydroquinone dehalogenase	Dehalogenation	Isomerase	[41]
DOPA decarboxylase	L-aromatic acids decarboxylation	Oxidative deamination of 5-hydroxytryptamine	[42]
Melamine deaminase	Deamination	Dechlorination	[8]
Phosphotriesterase-like lactonase	Lactone hydrolysis	Organophosphates hydrolysis	[43]
Hydroxyatrazine hydrolase	Deamination	Dechlorination	[44]
L-asparaginase	Amidohydrolase	Nitrilase	[28]
Chymotrypsin	Amidase	Phosphotriesterase	[28]

Promiscuity offers a source of pre-existing biochemical diversity upon which selection can act immediately following a gene duplication event and where those promiscuous functions provide a selective advantage, the duplicate genes are retained [28, 45]. Further mutations can then accrue and those that provide an enhancement of the promiscuous function are retained through natural selection, even if it results in the loss of the original physiological function which is retained by the organism *via* the other copy of the duplicated gene [28, 45-48].

The process by which enzymes gain new functions, or neofunctionalization, has been well studied in laboratory experiments, under non-natural selection [49-53]. There have been fewer studies with neofunctionalization under natural selection, as there are a limited number of examples that can be used as case studies or model systems [54-57]. Commonly, examples of naturally occurring neofunctionalization are drawn from the acquisition of novel enzymatic and metabolic functions that evolve in bacteria in response to anthropogenic selection pressures (e.g. antibiotic resistance and pesticide catabolism) [58, 59].

1.2 New metabolic functions have evolved in response to the use of herbicidal s-triazines

The introduction of anthropogenic compounds, such as antibiotics, insecticides, herbicides *etc.*, into the environment has provided new selection pressures. Many of these compounds are toxic and all can act as nutrient and/or energy sources, introducing selection pressures that derive from improved survival or enhanced growth [59, 60]. In many cases, these new selection pressures have resulted in the evolution of new enzymatic activities and metabolic pathways for the degradation of these compounds [60].

The response of microbiota to the introduction of the symmetrical triazines (*s*-triazines) has been studied in considerable detail [61-63], and provides rich insight into the process of evolving new metabolic function. First synthesized in the 1800s, the symmetrical triazines (*s*-triazines) are a family of anthropogenic compounds that share the same core structure, a ring (generally aromatic) of three carbon atoms and three nitrogen atoms arranged alternately to provide three-fold rotational symmetry (Fig. 1). The *s*-triazines are used in agriculture as herbicides (e.g. atrazine; Fig. 1), as monomers in plastics and polymers (e.g. melamine; Fig. 1), in explosives (e.g. RDX; Fig. 1) and as carriers of disinfectants (e.g. the trichloro salt of cyanuric acid; Fig. 1) [64-67]. Although most of the *s*-triazines are synthetic and their introduction in environment is relatively recent, related compounds, such as pyrimidines (e.g. uracil and thymine), are abundant and have been so since the early stages of life on Earth, forming the basis of genetic material (DNA and RNA) and other cellular components such as ATP and NAD [68]. Cyanuric acid (Fig. 2) is also thought to occur in nature as a product of oxidative and ionizing radiation damage of purines [69, 70], albeit the natural source of cyanuric acid is likely to produce relative low concentrations.

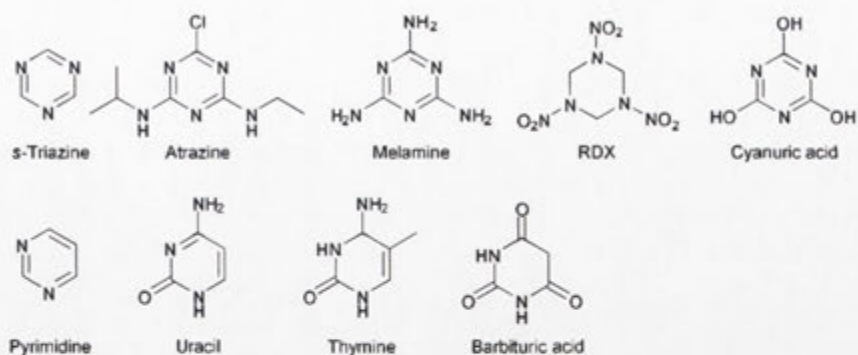


Figure 1. Examples of the *s*-triazines and related compounds. Structures of the *s*-triazine and pyrimidine core structures along with some common derivatives.

In the mid-twentieth century, *s*-triazine compounds became increasingly abundant and at high localized concentrations due to the dissemination of anthropogenic *s*-triazines in the environment [71]. The presence of an abundant source of carbon and nitrogen, in the form of anthropogenic *s*-triazines, has promoted the evolution of novel pathways for *s*-triazine catabolism in microbes [59]. On the basis of the arguments presented below, it is anticipated that these pathways have evolved recently over a relatively short period of time, therefore, they are excellent model systems for studying the evolution of new metabolic pathways and the acquisition of new enzyme functions [59, 72].

When first introduced, atrazine was relatively resistant to biodegradation, with half-lives ranging from months to years depending on soil type and environmental conditions [73-77]. Chemical degradation products were prevalent, produced by processes such as chlorohydrolysis under acidic or alkaline conditions [78-81]. Biodegradation of atrazine was often incomplete, resulting in the accumulation of metabolites, such as deethylatrazine and deisopropylatrazine [59]. Fungi and bacteria were isolated that could biological transform atrazine *via* oxidative dealkylation (Table 2) [82, 83]. Later, these dealkylation reactions were

shown to be cytochrome P450-mediated and non-specific (acting on broad range of structurally diverse herbicides) [84, 85].

Eventually, soil microbes were isolated that could rapidly degrade atrazine for use as a source of carbon and nitrogen [86-88]. The accelerated rate of atrazine degradation by these microbes was probably a result of the acquisition of new enzymatic activities and metabolic pathways that were more efficient and specific than the previously characterized dealkylation pathways [89-91]. Consortia of soil bacteria were isolated that could mineralise s-triazines, but there was no report of complete mineralization by any single bacterial strain until the mid-1990s.

In 1993, a *Pseudomonas* strain was isolated from an atrazine spillage site in Minnesota [92]. This strain (*Pseudomonas* sp. strain ADP) mineralised atrazine to CO₂ and grew on atrazine as a sole nitrogen source [86, 92]. Other bacterial strains, such as *Agrobacterium* and *Pseudomonas*, which could mineralise atrazine were subsequently isolated [93-95]. This suggested that, in response to direct selection pressure exerted by ubiquitous presence of atrazine in the environment, bacteria had evolved a new, efficient atrazine-mineralisation pathway.

Table 2. Time-line for the isolation of s-triazine metabolizing microbes.

Organism	Year reported	Source	Compounds tested	Moiety removed
<i>Aspergillus fumigatus</i>	1965	[82, 96]	Simazine, atrazine	Side chains
<i>Penicillium luteum</i>	1970	[82]	Atrazine	Chloride
<i>Penicillium janthinellum</i>	1970	[82]	Atrazine	Ethylamine
<i>Rhizopus stolonifera</i>	1970	[82]	Atrazine	Isopropylamine
<i>Stachybotrys chartarum</i>	1975	[97]	Atrazine, cyanuric acid	Cyanuric acid ring cleavage
<i>Hendersonula toruloidea</i>	1975	[97]	Atrazine, cyanuric acid	Cyanuric acid ring cleavage
<i>Pseudomonas sp. strain N.R.R.L. B-12227</i>	1982	[98-104]	Ammeline, ammelide, melamine, cyanuric acid, biuret, urea	Ammonia, ring cleavage of cyanuric acid
<i>Pseudomonas sp. strain N.R.R.L. B-12228</i>	1982	[99-101]	Cyanuric acid, biuret, urea	Ammonia, ring cleavage of cyanuric acid
<i>Klebsiella pneumoniae strain 99</i>	1982	[99-101]	Cyanuric acid, biuret, urea	Ammonia, ring cleavage of cyanuric acid
<i>Rhodococcus</i> NI86/21	1984	[105, 106]	Atrazine, deethylsimazine	Side chains
<i>Pleurotus pulmonarius</i>	1993	[107]	Atrazine	Side chains
<i>Rhodococcus strain TE1</i>	1993	[108]	Atrazine, propazine, simazine, cyanazine	Side chains
<i>Pseudomonas sp. strain ADP</i>	1993	[92, 109-115]	Atrazine	Halide, side chains, ring cleavage, ammonia
<i>Arthrobacter aurescens</i> TC1	2002	[116, 117]	Atrazine, simazine, and different atrazine analogues where chlorine is replaced by groups like mercaptan, fluorine and cyano.	Dehalogenation, mineralisation of side chains.

The atrazine-mineralizing pathway of *Pseudomonas sp.* strain ADP has been well studied and is composed of six hydrolytic enzymes: AtzA (E.C. 3.8.1.8; Atrazine chlorohydrolase; [118]), AtzB (EC 3.5.99.3; hydroxyatrazine hydrolase; [68]), AtzC (EC 3.5.99.4; N-isopropylammelide Amidohydrolase; [68]), AtzD (E.C. 3.5.2.15; cyanuric acid hydrolase; [113]), AtzE (EC 3.5.1.84; biuret hydrolase; [119]), and AtzF (EC 3.5.1.54; allophanate hydrolase; [119]). The enzymes of this pathway can be broadly categorized into an 'upper pathway' (AtzA, AtzB and AtzC) and a 'lower pathway' (AtzD, AtzE and AtzF).

The upper pathway enzymes are involved in dechlorination and deamination of side chains on atrazine (Fig. 2). These enzymes share a common α/β fold and are metal dependent hydrolases of the large amidohydrolase family [44, 111, 118, 120-122] (Pfam- PF01979). AtzE and AtzF are members of the Ser-*cis*Ser-Lys family of amidases (Pfam- PF02682) and catalyze the final two hydrolytic steps of the pathway [114, 119] (Fig. 2), sequentially deaminating biuret to allophanate and dicarboxyammonia. Allophanate hydrolase (AtzF) has been shown to also possess low-level biuret hydrolase activity [123]. Cyanuric acid hydrolase (AtzD) belongs to a novel family of ring opening hydrolases (Pfam PF09663). AtzD hydrolytically ring-opens cyanuric acid to give 1-carboxybiuret, which spontaneously decays to give biuret as a product (Fig. 2).

To date, the only other enzyme known AtzD-like enzyme is barbiturase [113, 124, 125]. Barbiturase ring-opens barbituric acid, a structural analogue of cyanuric acid that differs only by one less nitrogen in the heterocyclic ring (Fig. 1). The reaction product (ureidomalonic acid) of this ring-opening reaction is stable, unlike cyanuric acid. Barbiturase, which was purified from *Rhodococcus erythropolis* JCM 3132 in 2002, was reported to be a zinc

metalloenzyme [124]; however, subsequent metal chelation experiments with cyanuric acid suggested that a cation is not required for activity [125]. From these reports it was unclear if cyanuric acid hydrolase and barbiturase had different metal requirements, or if one (or both) of the reports had reached erroneous conclusions.

The genes encoding these six hydrolases are found on a single, self-transmissible plasmid (pADP-1) (Fig. 3), which encodes 104 putative open reading frames (ORFs), including all of the genes of required for atrazine mineralisation [114]. The genes of upper pathway (*atzA*, *atzB* and *atzC*) are somewhat dispersed on pADP-1, they are flanked by insertion sequences and transposases and are expressed constitutively [114]. The lower pathway genes are arranged in an operon [114], the expression of which has been shown to be controlled by nitrogen availability, *via* the action of RpoN [126], and cyanuric acid concentration, *via* a cyanuric-acid responsive LysR-like transcriptional activator (AtzR), encoded by a gene adjacent to, but divergent from, the *atzDEF* operon [127]. The lack of genetic and regulatory co-ordination in the upper pathway genes and their co-location with mobile genetic elements suggests that they may have been acquired more recently the lower pathway genes [114]. Cyanuric acid formation has been reported to occur during the repair of radiation damaged DNA, suggesting that it may have been produced in cells for billions of years (albeit in low abundance) [69, 70]. These observations suggest that the genes encoding the lower pathway enzymes may predate the introduction of synthetic s-triazine compounds [70].

Although the atrazine mineralisation pathway is conserved in atrazine-mineralizing bacteria, there are some variations in the identity of the enzymes involved. Isofunctional homologues

of AtzD and AtzF have been described (TrzD and TrzF, respectively) from *Pseudomonas* sp. strain NRRLB-12227 and *Enterobacter cloacae* strain 99, with identities of ~50% with AtzD and AtzF [112, 123, 128]. An alternative atrazine chlorohydrolase, TrzN, has also been identified [117, 129]. TrzN has been identified in different Gram-positive bacteria, including *Arthrobacter* and *Nocardioiodes*, and has relatively low sequence identity (< 30%) with its counterpart from *Pseudomonas* sp. strain ADP [130]. The first TrzN-containing isolate, *Nocardioiodes* sp. (strain C190), was isolated from a Canadian farm [129]. Although it was unable to mineralise *s*-triazines, it was found to hydrolyse a broad range of *s*-triazines including those with -OCH₃ and -SCH₃ substituents at the 2 position (such as ametryn and atraton, respectively) rather than a chloride [129]. The broader substrate range of *Nocardioiodes* sp. strain C190 compared with that of *Pseudomonas* sp. strain ADP was attributed to the presence of TrzN [129]. Although *Nocardioiodes* sp. strain C190 was found to not possess the genes encoding the lower pathway enzymes, another Gram-positive bacterium, *Arthrobacter aurescens* strain TC1, [116] was isolated that could mineralise a wide range of *s*-triazine compounds (Table. 2). Like *Nocardioiodes* sp. strain C190, *A. aurescens* strain TC1 was found to use TrzN, rather than AtzA, to catalyze the first step of herbicidal *s*-triazine mineralisation [117].

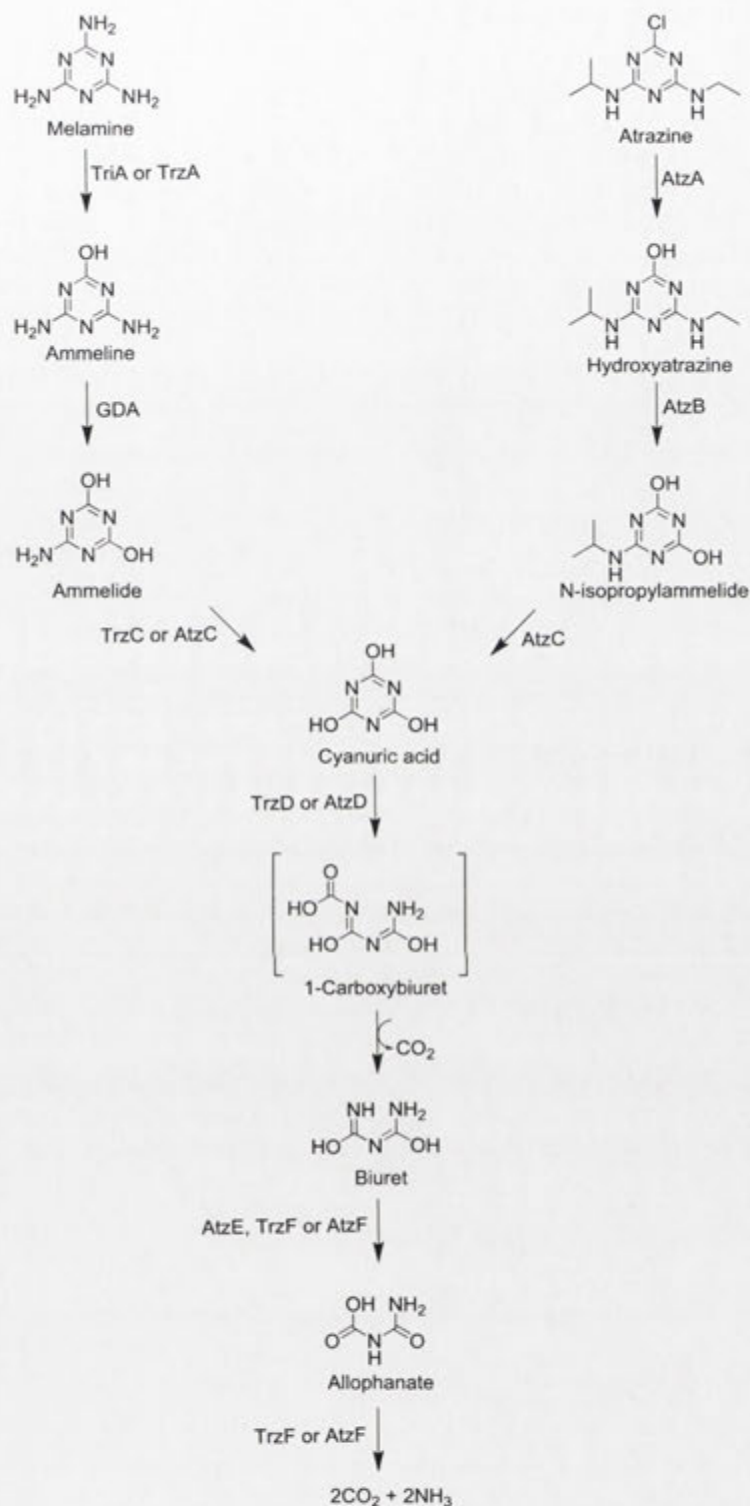


Figure 2. The atrazine and melamine mineralisation pathways. The groups at positions 2, 4 and 6 are removed by hydrolysis in both melamine and atrazine mineralizing pathways to yield cyanuric acid, which is then further hydrolyzed to ammonia and carbon dioxide.

It is of note that the lower pathway in atrazine degradation is also found in bacteria that degrade the non-herbicidal *s*-triazine melamine (2,4,6-diamino, 1,3,5-triazine), including *Pseudomonas* sp. strain NRRLB-12227, *Rhodococcus corallinus* strain NRRLB-15444R, *Rhodococcus* strain Mel and *Klebsiella terrigena* [102, 131-133]. This suggests that the lower pathway has been co-opted to provide metabolic capacity for cyanuric acid degradation in at least two independently evolved catabolic pathways. The upper pathways differ for melamine compared with the herbicidal *s*-triazines; however, both proceed *via* remarkably similar biochemical steps, with three successive hydrolyses leading to the formation of cyanuric acid (*via* ammeline and ammelide for melamine; Fig. 2) [102, 131-133]. Two related melamine deaminases have been identified, TriA and TrzA [102, 103, 131]; TrzA is *ca.* 43% identical to AtzA and 45% to TriA [103, 133] and TriA shares 98% sequence identity with AtzA, differing at just nine amino acids. Both TriA and TrzA possess low-level, promiscuous chlorohydrolase activity [8, 131, 134, 135] (notably, AtzA possesses no detectable deaminase activity) [8, 103, 135]. This suggests that TriA, AtzA and TrzA share a relatively recent common ancestor that had physiological deaminase activity and promiscuous dehalogenase activity.

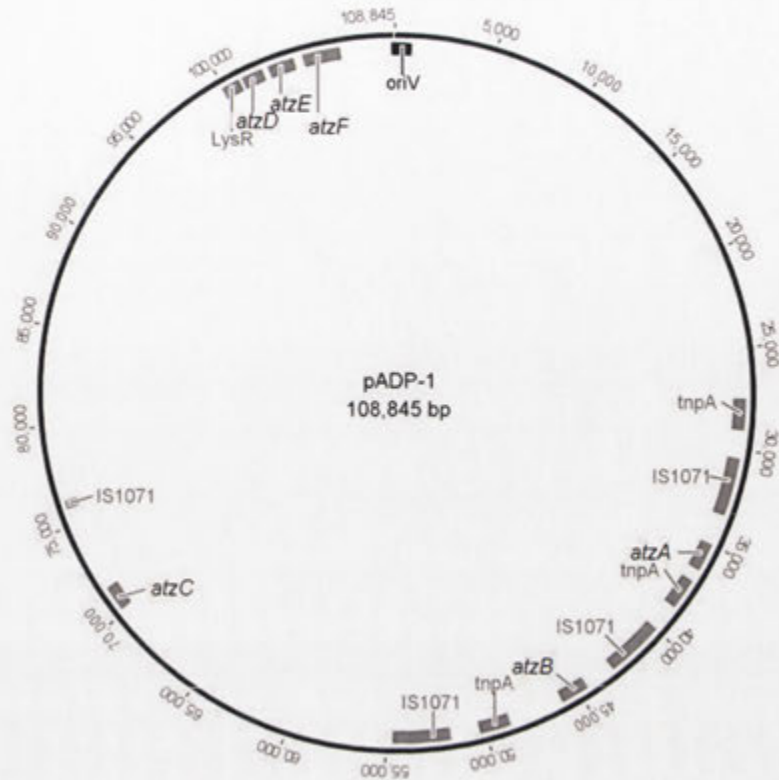


Figure 4. Map of pADP-1 plasmid from *Pseudomonas* sp. strain ADP. pADP-1 is a self-transmissible plasmid that encodes all the six genes of atrazine mineralisation pathway. The *atzA*, *atzB* and *atzC* genes are flanked with transposons (*tnpA*) and insertion elements (IS1071). The *atzD*, *atzE* and *atzF* genes are arranged in an operon and are regulated by a LysR-like regulator (AtzR).

Despite extensive studies of the environmental fate of atrazine, a hydrolytic pathway for atrazine mineralisation was not reported before 1993. It is therefore reasonable to assume that the catabolic pathways for atrazine were assembled recently [136]. Moreover, some of the new functionalities, such as the dechlorination of atrazine by AtzA, have likely been acquired recently and in response to the introduction a novel and highly-abundant nutrient (i.e., atrazine) [136, 137]. These 'new' enzyme functions are potentially excellent naturally occurring models for studying the acquisition of new enzymatic function.

The upper pathway enzymes (TrzN, AtzA, AtzB and AtzC) have evolved from members of the broader amidohydrolase family [59], albeit their low sequence identity may indicate that they do not share a common ancestral amidohydrolase origin. The evolution of isofunctional enzymes with low sequence identity (i.e., TrzN and AtzA) also suggests that similar enzymatic functions have been acquired *via* multiple evolutionary events and may therefore provide examples of convergent evolution. It is likely that AtzA and TriA diverged recently (98% identical), but are functionally distinct. This has provided the opportunity to explore a relatively small evolutionary landscape (i.e. nine amino acids) in which there are large quantitative and qualitative shifts in enzyme function.

Raillard and co-workers studied the functional plasticity of AtzA and TriA by shuffling the *atzA* and *triA* genes and testing the permutants for the parental activities and a range of novel activities. A high level of functional diversity was obtained, indicating that a relatively small sequence space defined by simple binary amino acid diversity at nine positions could contain considerable functional richness [135]. In a more recent study, potential evolutionary trajectories were reconstructed between AtzA and TriA [72], in which catalytic and biophysical characteristics of intermediates along the path of trajectories were studied [72]. In this study it was found that the order of substitutions when travelling from AtzA to TriA was not simply the reverse of the order from the TriA to AtzA trajectory, and that strong epistatic interactions between amino acids predicted to be in the active site governed the order of amino acid substitutions that separate the functionalities of these two enzymes [72]. Moreover, the strength of the trade-off between the two catalytic activities were quantitatively different in the two direction, with the transition from AtzA to TriA following a trade-off trajectory that permitted the retention of substantial AtzA activity during the

acquisition of moderate TriA activity, while the transition between TriA and AtzA resulted in the rapid loss of TriA-activity [72].

1.3 Bioremediation of *s*-triazines

Biotic atrazine degradation has considerable potential for applied outcomes in the environmental bioremediation of the *s*-triazines. Atrazine is used to control grassy and broadleaf weeds in corn, sorghum, and sugarcane [138]. The consumption of atrazine in United States for the year 2012 was reported to be 34.6 million kg according to the United States Environmental Protection Agency and often atrazine has been detected in ground water at concentrations 21 ppb and 42 ppb in surface water; these are concentrations well above the regulatory limits (0.1 ppb in Europe and 3 ppb in the USA) [139-143].

Despite their broad utility, there are a number of concerns about the environmental and human-health effects of the *s*-triazines. Hayes *et al.* reported that atrazine induced hermaphroditism and demasculinization in male frogs at ecologically relevant concentrations (≥ 0.1 ppb) [141]. Since then, a number of animal studies have supported and contradicted the hypothesis concurrently that atrazine could be an endocrine disruptor [144-151]. There have also been several studies that suggest that atrazine may be a carcinogen [152, 153]. In response to these potential issues the European Union banned the use of atrazine in 2004 [59].

There are number of different bioremediation techniques that have been used for the bioremediation of atrazine. Bioaugmentation and biostimulation are the techniques most

commonly used for atrazine bioremediation. Bioaugmentation is the introduction of microorganisms that have the metabolic capacity to degrade target pollutants to soil that does not otherwise have an endemic microbial population capable of metabolizing that pollutant [154]. Biostimulation is the addition of nutrients to the soil to enhance the metabolic activity of the soil microbiota and accelerate pollutant degradation [154]. There have been a number of successful field trials using atrazine-degrading inocula [155-158] and biostimulants [155, 158, 159].

Plant-mediated bioremediation (phytoremediation) of atrazine has also been explored using transgenic plants expressing bacterial atrazine-degrading enzymes. In one of such studies, transgenic rice plants expressing human cytochrome P450 genes were used to degrade atrazine and metolachlor [160]. Similarly, Wang *et al* studied the potential of transgenic tobacco plants for phytoremediation of atrazine [161]. The *atzA* genes from *Pseudomonas* sp. strain ADP and *Arthrobacter* strain AD1 were transferred into tobacco plant and the resulting transgenic lines were studied for ability to transform atrazine. The transgenic plants were able to achieve 100 % transformation of up to 2 mg/kg of atrazine in soil [161].

Cell-free enzymes have also been used to achieve atrazine bioremediation. Cell-free enzyme bioremediation is generally more rapid than the other techniques discussed herein and, unlike other bioremediation techniques, the effectiveness of cell-free enzyme bioremediation is not compromised by conditions unsupportive of the growth of microbes (e.g., nutrient availability, high/low pH, oxygen availability, etc.) [162]. However, this technique is unsuitable for reactions dependent upon diffusible cofactors (e.g., ATP,

NAD(P)H), which limits the range of enzymes that can be used [162]. The feasibility of cell-free bioremediation of atrazine has been demonstrated in two field trials to date [163, 164].

1.4 Aim of the thesis

Despite the isolation of *Pseudomonas* sp. strain ADP nearly 20 years ago, little has information has emerged regarding the structures of the atrazine-degrading enzymes or the structure-function relationships that underpin the catalytic mechanisms of these enzymes. Indeed, only the structure of the atrazine chlorohydrolase, TrzN, had been reported when the work reported before the work presented in this thesis had started. Structure-function information will be invaluable in future efforts in engineering these enzymes for improved kinetic and physical properties (such as reducing K_M or increasing enzyme stability). Moreover, such information improves our understanding of the evolutionary history of these model enzymes and metabolic pathways.

The aim of this thesis was to obtain structures for the enzymes of the newly evolved atrazine catabolic pathway from *Pseudomonas* sp. strain ADP and to use structure-function relationships to form a better understanding of metabolic pathway assembly and enzyme neofunctionalisation in this rare model system for natural enzyme and metabolic evolution.

References

1. Darwin, C., *On the origin of species*. 1871, New York :: D. Appleton and Co.
2. D'Argenio, D.A., Vetting, M.W., Ohlendorf, D.H., and Ornston, L.N., Substitution, insertion, deletion, suppression, and altered substrate specificity in functional protocatechuate 3,4-dioxygenases. *Journal of Bacteriology*, 1999. **181**: 6478-6487.
3. Rockah-Shmuel, L., Toth-Petroczy, A., Sela, A., Wurtzel, O., Sorek, R., and Tawfik, D.S., Correlated occurrence and bypass of frame-shifting insertion-deletions (indels) to give functional proteins. *PLoS Genetics*, 2013. **9**: e1003882.
4. Seebeck, F.P. and Hilvert, D., Conversion of a PLP-dependent racemase into an aldolase by a single active site mutation. *Journal of the American Chemical Society*, 2003. **125**: 10158-10159.
5. Tsugeno, Y. and Ito, A., A key amino acid responsible for substrate selectivity of monoamine oxidase A and B. *Journal of Biological Chemistry*, 1997. **272**: 14033-14036.
6. Tokuriki, N., Stricher, F., Serrano, L., and Tawfik, D.S., How protein stability and new functions trade off. *PLoS Computational Biology*, 2008. **4**: e1000002
7. Porter, J.L., Boon, P.L.S., Murray, T.P., Huber, T., Collyer, C.A., and Ollis, D.L., Directed evolution of new and improved enzyme functions using an evolutionary intermediate and multidirectional search. *ACS Chemical Biology*, 2015. **10**: 611-621.
8. Noor, S., Taylor, M.C., Russell, R.J., Jermiin, L.S., Jackson, C.J., Oakeshott, J.G., and Scott, C., Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS One*, 2012. **7**: e39822.
9. Khersonsky, O. and Tawfik, D.S., Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annual Review of Biochemistry*, Vol 79, 2010. **79**: 471-505.
10. Copley, S.D., Toward a systems biology perspective on enzyme evolution. *The Journal of Biological Chemistry*, 2012. **287**: 3-10.
11. Soskine, M. and Tawfik, D.S., Mutational effects and the evolution of new protein functions. *Nature Reviews Genetics*, 2010. **11**: 572-582.
12. Wang, X.J., Minasov, G., and Shoichet, B.K., Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *Journal of Molecular Biology*, 2002. **320**: 85-95.
13. Tokuriki, N., Stricher, F., Serrano, L., and Tawfik, D.S., How protein stability and new functions trade off. *PLoS Computational Biology*, 2008. **4**: p. e1000002.
14. Jackson, C.J., Oakeshott, J.G., Sanchez-Hernandez, J.C., and Wheelock, C.E., Carboxylesterases in the metabolism and toxicity of pesticides, in *Anticholinesterase pesticides*. 2011, John Wiley & Sons, Inc. 57-75.
15. Vincent, B.M., Lancaster, A.K., Scherz-Shouval, R., Whitesell, L., and Lindquist, S., Fitness trade-off's restrict the evolution of resistance to amphotericin B. *PLoS Biology*, 2013. **11**: e1001692.
16. Kaltenbach, M. and Tokuriki, N., Dynamics and constraints of enzyme evolution. *Journal of Experimental Zoology Part B-Molecular and Developmental Evolution*, 2014. **322**: 468-487.
17. Tokuriki, N., Jackson, C.J., Afriat-Jurnou, L., Wyganowski, K.T., Tang, R.M., and Tawfik, D.S., Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nature Communications*, 2012. **3**.

18. Ohno, S., Wolf, U., and Atkin, N.B., Evolution from fish to mammals by gene duplication. *Hereditas*, 1968. **59**: 169-187.
19. Ohno, S., Evolution by gene duplication. 1970, New York: Springer-Verlag.
20. Zhang, J., Evolution by gene duplication: An update. *Trends in Ecology & Evolution*, 2003. **18**: 292-298.
21. Teichmann, S.A. and Babu, M.M., Gene regulatory network growth by duplication. *Nature genetics*, 2004. **36**: 492-496.
22. Labedan, B. and Riley, M., Widespread protein sequence similarities: Origins of *Escherichia coli* genes. *Journal of Bacteriology*, 1995. **177**: 1585-1588.
23. Gough, J., Karplus, K., Hughey, R., and Chothia, C., Assignment of homology to genome sequences using a library of hidden markov models that represent all proteins of known structure1. *Journal of Molecular Biology*, 2001. **313**: 903-919.
24. de Rosa, R. and Labedan, B., The evolutionary relationships between the two bacteria *Escherichia coli* and *Haemophilus influenzae* and their putative last common ancestor. *Molecular Biology and Evolution*, 1998. **15**: 17-27.
25. Lynch, M. and Conery, J.S., The evolutionary demography of duplicate genes. *Journal of Structural and Functional Genomics*, 2003. **3**: 35-44.
26. Guo, H., Lee, T.H., Wang, X.Y., and Paterson, A.H., Function relaxation followed by diversifying selection after whole-genome duplication in flowering plants. *Plant Physiology*, 2013. **162**: 769-778.
27. Bershtein, S. and Tawfik, D.S., Ohno's model revisited: Measuring the frequency of potentially adaptive mutations under various mutational drifts. *Molecular Biology and Evolution*, 2008. **25**: 2311-2318.
28. O'Brien, P.J. and Herschlag, D., Catalytic promiscuity and the evolution of new enzymatic activities. *Chemistry & Biology*, 1999. **6**: R91-R105.
29. Bornscheuer, U.T. and Kazlauskas, R.J., Catalytic promiscuity in biocatalysis: Using old enzymes to form new bonds and follow new pathways. *Angewandte Chemie-International Edition*, 2004. **43**: 6032-6040.
30. Copley, S.D., An evolutionary biochemist's perspective on promiscuity. *Trends in Biochemical Sciences*, 2015. **40**: 72-78.
31. Khersonsky, O. and Tawfik, D.S., Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annual Review of Biochemistry*, 2010. **79**: 471-505.
32. Copley, S.D., Enzymes with extra talents: Moonlighting functions and catalytic promiscuity. *Current Opinion in Chemical Biology*, 2003. **7**: 265-272.
33. Arora, B., Mukherjee, J., and Gupta, M., Enzyme promiscuity: Using the dark side of enzyme specificity in white biotechnology. *Sustainable Chemical Processes*, 2014. **2**: 25.
34. Roodveldt, C. and Tawfik, D.S., Shared promiscuous activities and evolutionary features in various members of the amidohydrolase superfamily. *Biochemistry*, 2005. **44**: 12728-12736.
35. Porter, T.N., Li, Y.C., and Raushel, F.M., Mechanism of the dihydroorotase reaction. *Biochemistry*, 2004. **43**: 16285-16292.
36. Poelarends, G.J., Serrano, H., Johnson, W.H., Hoffman, D.W., and Whitman, C.P., The hydratase activity of malonate semialdehyde decarboxylase: Mechanistic and evolutionary implications. *Journal of the American Chemical Society*, 2004. **126**: 15658-15659.

37. Thoden, J.B., Ringia, E.A.T., Garrett, J.B., Gerlt, J.A., Holden, H.M., and Rayment, I., Evolution of enzymatic activity in the enolase superfamily: Structural studies of the promiscuous *O*-succinylbenzoate synthase from amycolatopsis. *Biochemistry*, 2004. **43**: 5716-5727.
38. Palmer, D.R.J., Garrett, J.B., Sharma, V., Meganathan, R., Babbitt, P.C., and Gerlt, J.A., Unexpected divergence of enzyme function and sequence: "N-acylamino acid racemase" is *O*-succinylbenzoate synthase. *Biochemistry*, 1999. **38**: 4252-4258.
39. Yew, W.S., Akana, J., Wise, E.L., Rayment, I., and Gerlt, J.A., Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: Enhancing the promiscuous d-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-l-gulonate 6-phosphate decarboxylase. *Biochemistry*, 2005. **44**: 1807-1815.
40. Li, C., Hassler, M., and Bugg, T.D.H., Catalytic promiscuity in the alpha/beta-hydrolase superfamily: Hydroxamic acid formation, C-C bond formation, ester and thioester hydrolysis in the C-C hydrolase family. *Chembiochem*, 2008. **9**: 71-76.
41. Anandarajah, K., Kiefer, P.M., Donohoe, B.S., and Copley, S.D., Recruitment of a double bond isomerase to serve as a reductive dehalogenase during biodegradation of pentachlorophenol. *Biochemistry*, 2000. **39**: 5303-5311.
42. Nath, A. and Atkins, W.M., A quantitative index of substrate promiscuity. *Biochemistry*, 2008. **47**: 157-166.
43. Zhang, Y., An, J., Ye, W., Yang, G.Y., Qian, Z.G., Chen, H.F., Cui, L., and Feng, Y., Enhancing the promiscuous phosphotriesterase activity of a thermostable lactonase (gkap) for the efficient degradation of organophosphate pesticides. *Applied and Environmental Microbiology*, 2012. **78**: 6647-6655.
44. Seffernick, J.L., Aleem, A., Osborne, J.P., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Hydroxyatrazine *N*-ethylaminohydrolase (AtzB): An amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *Journal of Bacteriology*, 2007. **189**: 6989-6997.
45. Copley, S.D., Toward a systems biology perspective on enzyme evolution. *Journal of Biological Chemistry*, 2012. **287**: 3-10.
46. Nam, H., Lewis, N.E., Lerman, J.A., Lee, D.H., Chang, R.L., Kim, D., and Palsson, B.O., Network context and selection in the evolution to enzyme specificity. *Science*, 2012. **337**: 1101-1104.
47. Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., and Postlethwait, J., Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 1999. **151**: 1531-1545.
48. Rodriguez-Trelles, F., Tarrío, R., and Ayala, F.J., Convergent neofunctionalization by positive darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proceedings of the National Academy of Sciences of the United States of America*, 2003. **100**: 13413-13417.
49. Assis, R. and Bachtrog, D., Neofunctionalization of young duplicate genes in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 2013. **110**: 17409-17414.
50. Voordeckers, K., Brown, C.A., Vanneste, K., van der Zande, E., Voet, A., Maere, S., and Verstrepen, K.J., Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. *PLoS Biology*, 2012. **10**: e1001446

51. Rastogi, S. and Liberles, D., Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evolutionary Biology*, 2005. **5**
52. Assis, R. and Bachtrog, D., Neofunctionalization of young duplicate genes in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 2013. **110**: 17409-17414.
53. Cancela, M.L., Laize, V., and Conceicao, N., Matrix Gla protein and osteocalcin: From gene duplication to neofunctionalization. *Archives of Biochemistry and Biophysics*, 2014. **561**: 56-63.
54. Rodriguez-Trelles, F., Tarrío, R., and Ayala, F.J., Convergent neofunctionalization by positive darwinian selection after ancient recurrent duplications of the xanthine dehydrogenase gene. *Proceedings of the National Academy of Sciences of the United States of America*, 2003. **100**: 13413-13417.
55. Dun, X., Shen, W., Hu, K., Zhou, Z., Xia, S., Wen, J., Yi, B., Shen, J., Ma, C., Tu, J., Fu, T., and Lagercrantz, U., Neofunctionalization of duplicated tic40 genes caused a gain-of-function variation related to male fertility in *Brassica oleracea* lineages. *Plant Physiology*, 2014. **166**: 1403-1419.
56. Teshima, K.M. and Innan, H., Neofunctionalization of duplicated genes under the pressure of gene conversion. *Genetics*, 2008. **178**: 1385-1398.
57. Gibson, T.A. and Goldberg, D.S., Questioning the ubiquity of neofunctionalization. *PLoS Computational Biology*, 2009. **5**: e1000252.
58. Kolar, M., Urbanek, K., and Latal, T., Antibiotic selective pressure and development of bacterial resistance. *International Journal of Antimicrobial Agents*, 2001. **17**: 357-363.
59. Udikovic-Kolic, N., Scott, C., and Martin-Laurent, F., Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology*, 2012. **96**: 1175-1189.
60. Russell, R.J., Scott, C., Jackson, C.J., Pandey, R., Pandey, G., Taylor, M.C., Coppin, C.W., Liu, J.W., and Oakeshott, J.G., The evolution of new enzyme function: Lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. *Evolutionary Applications*, 2011. **4**: 225-248.
61. Udikovic-Kolic, N., Scott, C., and Martin-Laurent, F., Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology*, 2012. **96**: 1175-1189.
62. Ralebits, T.K., Senior, E., and van Verseveld, H.W., Microbial aspects of atrazine degradation in natural environments. *Biodegradation*, 2002. **13**: 11-19.
63. Shapir, N., Mongodin, E.F., Sadowsky, M.J., Daugherty, S.C., Nelson, K.E., and Wackett, L.P., Evolution of catabolic pathways: Genomic insights into microbial S-triazine metabolism. *Journal of Bacteriology*, 2007. **189**: 674-682.
64. Cox, J.R., Triazine derivatives as non-selective herbicides. *Journal of the Science of Food and Agriculture*, 1962. **13**: 99-103.
65. Wirpsza, Z., Achievements in the use of melamine as a chemical raw material . Part 1. General review and melamine-based coating resins and coatings. *Polimery*, 1996. **41**: 265-276.
66. Bachmann, W.E. and Sheehan, J.C., A new method of preparing the high explosive RDX. *Journal of the American Chemical Society*, 1949. **71**: 1842-1845.

67. Canelli, E., Chemical, bacteriological, and toxicological properties of cyanuric acid and chlorinated isocyanurates as applied to swimming pool disinfection - review. *American Journal of Public Health*, 1974. **64**: 155-162.
68. Wackett, L.P., Sadowsky, M.J., Martinez, B., and Shapir, N., Biodegradation of atrazine and related s-triazine compounds: From enzymes to field studies. *Applied Microbiology and Biotechnology*, 2002. **58**: 39-45.
69. Dherin, C., Gasparutto, D., O'connor, T.R., Cadet, J., and Boiteux, S., Excision by the human methylpurine DNA N-glycosylase of cyanuric acid, a stable and mutagenic oxidation product of 8-oxo-7,8-dihydroguanine. *International Journal of Radiation Biology*, 2004. **80**: 21-27.
70. Wackett, L.P., Questioning our perceptions about evolution of biodegradative enzymes. *Current Opinion in Microbiology*, 2009. **12**: 244-251.
71. Noor, S., Changey, F., Oakeshott, J.G., Scott, C., and Martin-Laurent, F., Ongoing functional evolution of the bacterial atrazine chlorohydrolase AtzA. *Biodegradation*, 2014. **25**: 21-30.
72. Noor, S., Taylor, M.C., Russell, R.J., Jermiin, L.S., Jackson, C.J., Oakeshott, J.G., and Scott, C., Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS One*, 2012. **7**: e39822.
73. Harris, C.I., Decomposition in soil fate of 2-chloro-s-triazine herbicides in soil. *Journal of Agricultural and Food Chemistry*, 1967. **15**: 157-162.
74. Sheets, T.J., Persistence of triazine herbicides in soils, in *Single pesticide volume: The triazine herbicides*, F. Gunther and J. Gunther, Editors. 1970, Springer New York. 287-310.
75. Jones, T.W., Kemp, W.M., Stevenson, J.C., and Means, J.C., Degradation of atrazine in Estuarine water/sediment systems and soils. *Journal of Environmental Quality*, 1982. **11**: 632-638.
76. Frank, R. and Sirons, G., Dissipation of atrazine residues from soils. *Bulletin of Environmental Contamination and Toxicology*, 1985. **34**: 541-548.
77. Kruger, E.L., Somasundaram, L., Coats, J.R., and Kanwar, R.S., Persistence and degradation of [14C]atrazine and [14C]deisopropylatrazine as affected by soil depth and moisture conditions. *Environmental Toxicology and Chemistry*, 1993. **12**: 1959-1967.
78. Skipper, H.D., Gilmour, C.M., and Furtick, W.R., Microbial versus chemical degradation of atrazine in soils. *Soil Science Society of America Journal.*, 1967. **31**: 653-656.
79. Dao, T.H., Lavy, T.L., and Sorensen, R.C., Atrazine degradation and residue distribution in soil1. *Soil Science Society of America Journal.*, 1979. **43**: 1129-1134.
80. Capriel, P. and Haisch, A., Persistence of atrazine and its metabolites in soil after a single herbicide application. *Zeitschrift für Pflanzenernährung und Bodenkunde*, 1983. **146**: 474-480.
81. Armstrong, De, Chesters, G., and Harris, R.F., Atrazine hydrolysis in soil. *Soil Science Society of America Journal.*, 1967. **31**: 61-66.
82. Kaufman, D.D. and Blake, J., Degradation of atrazine by soil fungi. *Soil Biology and Biochemistry*, 1970. **2**: 73-80.
83. Giardina, M.C., Giardi, M.T., and Filacchioni, G., Atrazine metabolism by nocardia: Elucidation of initial pathway and synthesis of potential metabolites. *Agricultural and Biological Chemistry*, 1982. **46**: 1439-1445.

84. Behki, R.M. and Khan, S.U., Degradation of atrazine by pseudomonas: N-dealkylation and dehalogenation of atrazine and its metabolites. *Journal of Agricultural and Food Chemistry*, 1986. **34**: 746-749.
85. Behki, R.M. and Khan, S.U., Degradation of atrazine, propazine, and simazine by *Rhodococcus* strain B-30. *Journal of Agricultural and Food Chemistry*, 1994. **42**: 1237-1241.
86. Mandelbaum, R.T., Allan, D.L., and Wackett, L.P., Isolation and characterization of a *Pseudomonas* sp. that mineralises the s-triazine herbicide atrazine. *Applied and Environmental Microbiology*, 1995. **61**: 1451-1457.
87. Vargha, M., Takats, Z., and Marialigeti, K., Degradation of atrazine in a laboratory scale model system with Danube river sediment. *Water Research*, 2005. **39**: 1560-1568.
88. Singh, P., Suri, C.R., and Cameotra, S.S., Isolation of a member of *Acinetobacter* species involved in atrazine degradation. *Biochemical and Biophysical Research Communications*, 2004. **317**: 697-702.
89. Barriuso, E. and Houot, S., Rapid mineralization of the s-triazine ring of atrazine in soils in relation to soil management. *Soil Biology and Biochemistry*, 1996. **28**: 1341-1348.
90. Ostrofsky, E.B., Traina, S.J., and Tuovinen, O.H., Variation in atrazine mineralization rates in relation to agricultural management practice. *Journal of Environmental Quality*, 1997. **26**: 647-657.
91. Vanderheyden, V., Debongnie, P., and Pussemier, L., Accelerated degradation and mineralization of atrazine in surface and subsurface soil materials. *Pesticide Science*, 1997. **49**: 237-242.
92. Mandelbaum, R.T., Wackett, L.P., and Allan, D.L., Mineralization of the s-triazine ring of atrazine by stable bacterial mixed cultures. *Applied and Environmental Microbiology*, 1993. **59**: 1695-1701.
93. Yanzekontchou, C. and Gschwind, N., Mineralization of the herbicide atrazine as a carbon source by a *Pseudomonas* strain. *Applied and Environmental Microbiology*, 1994. **60**: 4297-4302.
94. Radosevich, M., Traina, S.J., Hao, Y.L., and Tuovinen, O.H., Degradation and mineralization of atrazine by a soil bacterial isolate. *Applied and Environmental Microbiology*, 1995. **61**: 297-302.
95. Struthers, J.K., Jayachandran, K., and Moorman, T.B., Biodegradation of atrazine by *Agrobacterium radiobacter* J14A and use of this strain in bioremediation of contaminated soil. *Applied and Environmental Microbiology*, 1998. **64**: 3368-3375.
96. Kearney, P.C., Kaufman, D.D., and Sheets, T.J., Fungal metabolism of herbicides, metabolites of simazine by *Aspergillus fumigatus*. *Journal of Agricultural and Food Chemistry*, 1965. **13**: 369-372.
97. Wolf, D.C. and Martin, J.P., Microbial decomposition of ring C-14 atrazine, ayanuric acid, and 2-chloro-4,6-diamino-s-triazine. *Journal of Environmental Quality*, 1975. **4**: 134-139.
98. Jutzi, K., Cook, A.M., and Hutter, R., The degradative pathway of the s-triazine melamine - the steps to ring cleavage. *Biochemical Journal*, 1982. **208**: 679-684.
99. Cook, A.M. and Hutter, R., s-triazines as nitrogen-sources for bacteria. *Journal of Agricultural and Food Chemistry*, 1981. **29**: 1135-1143.

100. Cook, A.M., Beilstein, P., Grossenbacher, H., and Hutter, R., Ring cleavage and degradative pathway of cyanuric acid in bacteria. *Biochemical Journal*, 1985. **231**: 25-30.
101. Eaton, R.W. and Karns, J.S., Cloning and comparison of the DNA encoding ammelide aminohydrolase and cyanuric acid amidohydrolase from 3 s-triazine-degrading bacterial strains. *Journal of Bacteriology*, 1991. **173**: 1363-1366.
102. Eaton, R.W. and Karns, J.S., Cloning and analysis of s-triazine catabolic genes from *Pseudomonas* sp. Strain NRRLB-12227. *Journal of Bacteriology*, 1991. **173**: 1215-1222.
103. Seffernick, J.L., de Souza, M.L., Sadowsky, M.J., and Wackett, L.P., Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *Journal of Bacteriology*, 2001. **183**: 2405-2410.
104. Karns, J.S., Gene sequence and properties of an s-triazine ring-cleavage enzyme from *Pseudomonas* sp strain NRRLB-12227. *Applied and Environmental Microbiology*, 1999. **65**: 3512-3517.
105. Nagy, I., Compennolle, F., Ghys, K., Vanderleyden, J., and Demot, R., 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. *Applied and Environmental Microbiology*, 1995. **61**: 2056-2060.
106. Cook, A.M. and Hutter, R., Deethylsimazine - bacterial dechlorination, deamination, and complete degradation. *Journal of Agricultural and Food Chemistry*, 1984. **32**: 581-585.
107. Masaphy, S., Levanon, D., Vaya, J., and Henis, Y., Isolation and characterization of a novel atrazine metabolite produced by the fungus *Pleurotus-pulmonarius*, 2-chloro-4-ethylamino-6-(1-hydroxyisopropyl)amino-1,3,5-triazine. *Applied and Environmental Microbiology*, 1993. **59**: 4342-4346.
108. Behki, R., Topp, E., Dick, W., and Germon, P., Metabolism of the herbicide atrazine by *Rhodococcus* strains. *Applied and Environmental Microbiology*, 1993. **59**: 1955-1959.
109. deSouza, M.L., Sadowsky, M.J., and Wackett, L.P., Atrazine chlorohydrolase from *Pseudomonas* sp strain ADP: Gene sequence, enzyme purification, and protein characterization. *Journal of Bacteriology*, 1996. **178**: 4894-4900.
110. BoundyMills, K.L., deSouza, M.L., Mandelbaum, R.T., Wackett, L.P., and Sadowsky, M.J., The *atzB* gene of *Pseudomonas* sp. strain ADP encodes the second enzyme of a novel atrazine degradation pathway. *Applied and Environmental Microbiology*, 1997. **63**: 916-923.
111. Sadowsky, M.J., Tong, Z.K., de Souza, M., and Wackett, L.P., AtzC is a new member of the amidohydrolase protein superfamily and is homologous to other atrazine-metabolizing enzymes. *Journal of Bacteriology*, 1998. **180**: 152-158.
112. Fruchey, I., Shapir, N., Sadowsky, M.J., and Wackett, L.P., On the origins of cyanuric acid hydrolase: Purification, substrates, and prevalence of AtzD from *Pseudomonas* sp strain ADP. *Applied and Environmental Microbiology*, 2003. **69**: 3653-3657.
113. Peat, T.S., Balotra, S., Wilding, M., French, N.G., Briggs, L.J., Panjikar, S., Cowieson, N., Newman, J., and Scott, C., Cyanuric acid hydrolase: Evolutionary innovation by structural concatenation. *Molecular Microbiology*, 2013. **88**: 1149-1163.

114. Martinez, B., Tomkins, J., Wackett, L.P., Wing, R., and Sadowsky, M.J., Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. Strain ADP. *Journal of Bacteriology*, 2001. **183**: 5684-5697.
115. Shapir, N., Sadowsky, M.J., and Wackett, L.P., Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp strain ADP. *Journal of Bacteriology*, 2005. **187**: 3731-3738.
116. Strong, L.C., Rosendahl, C., Johnson, G., Sadowsky, M.J., and Wackett, L.P., *Arthrobacter aurescens* TC1 metabolizes diverse s-triazine ring compounds. *Applied and Environmental Microbiology*, 2002. **68**: 5973-5980.
117. Sajjaphan, K., Shapir, N., Wackett, L.P., Palmer, M., Blackmon, B., Tomkins, J., and Sadowsky, M.J., *Arthrobacter aurescens* TC1 atrazine catabolism genes *trzN*, *atzB*, and *atzC* are linked on a 160-kilobase region and are functional in *Escherichia coli*. *Applied and Environmental Microbiology*, 2004. **70**: 4402-4407.
118. Wang, Y., Li, X., Chen, X., and Chen, D., Directed evolution and characterization of atrazine chlorohydrolase variants with enhanced activity. *Biochemistry (Moscow)*, 2013. **78**: 1104-1111.
119. Balotra, S., Newman, J., French, N.G., Briggs, L.J., Peat, T.S., and Scott, C., Crystallization and preliminary x-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp strain ADP. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*, 2014. **70**: 310-315.
120. Seffernick, J.L., McTavish, H., Osborne, J.P., de Souza, M.L., Sadowsky, M.J., and Wackett, L.P., Atrazine chlorohydrolase from *Pseudomonas* sp strain ADP is a metalloenzyme. *Biochemistry*, 2002. **41**: 14430-14437.
121. Wackett, L.P., Evolution of new enzymes and pathways: Soil microbes adapt to s-triazine herbicides. *Pesticide Decontamination and Detoxification*, 2004. **863**: 37-48.
122. Wang, Y., Li, X., Chen, X., and Chen, D., Directed evolution and characterization of atrazine chlorohydrolase variants with enhanced activity. *Biochemistry-Moscow*, 2013. **78**: 1104-1111.
123. Shapir, N., Cheng, G., Sadowsky, M.J., and Wackett, L.P., Purification and characterization of TrzF: Biuret hydrolysis by allophanate hydrolase supports growth. *Applied and Environmental Microbiology*, 2006. **72**: 2491-2495.
124. Soong, C.L., Ogawa, J., Sakuradani, E., and Shimizu, S., Barbiturase, a novel zinc-containing amidohydrolase involved in oxidative pyrimidine metabolism. *Journal of Biological Chemistry*, 2002. **277**: 7051-7058.
125. Seffernick, J.L., Erickson, J.S., Cameron, S.M., Cho, S., Dodge, A.G., Richman, J.E., Sadowsky, M.J., and Wackett, L.P., Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. *Journal of Bacteriology*, 2012. **194**: 4579-4588.
126. Garcia-Gonzalez, V., Govantes, F., Porrúa, O., and Santero, E., Regulation of the *Pseudomonas* sp. Strain ADP cyanuric acid degradation operon. *Journal of bacteriology*, 2005. **187**: 155-167.
127. Govantes, F., Garcia-Gonzalez, V., Porrúa, O., Platero, A.I., Jimenez-Fernandez, A., and Santero, E., Regulation of the atrazine-degradative genes in *Pseudomonas* sp strain ADP. *FEMS Microbiology Letters*, 2010. **310**: 1-8.
128. Cheng, G., Shapir, N., Sadowsky, M.J., and Wackett, L.P., Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism. *Applied and Environmental Microbiology*, 2005. **71**: 4437-4445.

129. Mulbry, W.W., Zhu, H., Nour, S.M., and Topp, E., The triazine hydrolase gene *trzN* from *Nocardioides* sp strain C190: Cloning and construction of gene-specific primers. *FEMS Microbiology Letters*, 2002. **206**: 75-79.
130. Seffernick, J.L., Reynolds, E., Fedorov, A.A., Fedorov, E., Almo, S.C., Sadowsky, M.J., and Wackett, L.P., X-ray structure and mutational analysis of the atrazine chlorohydrolase TrzN. *Journal of Biological Chemistry*, 2010. **285**: 30606-30614.
131. Mulbry, W.W., Purification and characterization of an inducible *s*-triazine hydrolase from *Rhodococcus-coralinus* NRLLB-15444R. *Applied and Environmental Microbiology*, 1994. **60**: 613-618.
132. Shelton, D.R., Karns, J.S., McCarty, G.W., and Durham, D.R., Metabolism of melamine by *Klebsiella terrigena*. *Applied and Environmental Microbiology*, 1997. **63**: 2832-2835.
133. Dodge, A.G., Wackett, L.P., and Sadowsky, M.J., Plasmid localization and organization of melamine degradation genes in *Rhodococcus* sp strain mel. *Applied and Environmental Microbiology*, 2012. **78**: 1397-1403.
134. Seffernick, J.L., Shapir, N., Schoeb, M., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Enzymatic degradation of chlorodiamino-*s*-triazine. *Applied and Environmental Microbiology*, 2002. **68**: 4672-4675.
135. Raillard, S., Krebber, A., Chen, Y.C., Ness, J.E., Bermudez, E., Trinidad, R., Fullem, R., Davis, C., Welch, M., Seffernick, J., Wackett, L.P., Stemmer, W.P.C., and Minshull, J., Novel enzyme activities and functional plasticity revealed by recombining highly homologous enzymes. *Chemistry & Biology*, 2001. **8**: 891-898.
136. Udikovic-Kolic, N., Scott, C., and Martin-Laurent, F., Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology*, 2012. **96**: 1175-1189.
137. Renata, H., Wang, Z.J., and Arnold, F.H., Expanding the enzyme universe: Accessing non-natural reactions by mechanism-guided directed evolution. *Angewandte Chem International Edition in English*, 2015. **54**: 3351-3367.
138. Fan, X.X. and Song, F.Q., Bioremediation of atrazine: Recent advances and promises. *Journal of Soils and Sediments*, 2014. **14**: 1727-1737.
139. Ta, N., Hong, J., Liu, T.F., and Sun, C., Degradation of atrazine by microwave-assisted electrodeless discharge mercury lamp in aqueous solution. *Journal of Hazardous Materials*, 2006. **138**: 187-194.
140. Mahia, J., Martin, A., Carballas, T., and Diaz-Ravina, M., Atrazine degradation and enzyme activities in an agricultural soil under two tillage systems. *Science of the Total Environment*, 2007. **378**: 187-194.
141. Hayes, T.B., Collins, A., Lee, M., Mendoza, M., Noriega, N., Stuart, A.A., and Vonk, A., Hermaphroditic, demasculinized frogs after exposure to the herbicide atrazine at low ecologically relevant doses. *Proceedings of the National Academy of Sciences of the United States of America*, 2002. **99**: 5476-5480.
142. Steinberg, C.E.W., Lorenz, R., and Spieser, O.H., Effects of atrazine on swimming behavior of zebrafish, *Brachydanio-rerio*. *Water Research*, 1995. **29**: 981-985.
143. *Atrazine background*. 2012 [cited 2015 07052015]; Available from: http://www.epa.gov/pesticides/factsheets/atrazine_background.htm.
144. Kucka, M., Pogrmic-Majkic, K., Fa, S., Stojilkovic, S.S., and Kovacevic, R., Atrazine acts as an endocrine disrupter by inhibiting camp-specific phosphodiesterase-4. *Toxicology and Applied Pharmacology*, 2012. **265**: 19-26.

145. Fan, W., Yanase, T., Morinaga, H., Gondo, S., Okabe, T., Nomura, M., Hayes, T.B., Takayanagi, R., and Nawata, H., Herbicide atrazine activates sf-1 by direct affinity and concomitant co-activators recruitments to induce aromatase expression via promoter ii. *Biochemical and Biophysical Research Communications*, 2007. **355**: 1012-1018.
146. Carr, J.A., Gentles, A., Smith, E.E., Goleman, W.L., Urquidi, L.J., Thuett, K., Kendall, R.J., Giesy, J.P., Gross, T.S., Solomon, K.R., and Van der Kraak, G., Response of larval *Xenopus laevis* to atrazine: Assessment of growth, metamorphosis, and gonadal and laryngeal morphology. *Environmental Toxicology and Chemistry*, 2003. **22**: 396-405.
147. Shenoy, K., Environmentally realistic exposure to the herbicide atrazine alters some sexually selected traits in male guppies. *PLoS One*, 2012. **7**: e30611
148. Coady, K.K., Murphy, M.B., Villeneuve, D.L., Hecker, M., Jones, P.D., Carr, J.A., Solomon, K.R., Smith, E.E., Van der Kraak, G., Kendall, R.J., and Giesy, J.P., Effects of atrazine on metamorphosis, growth, and gonadal development in the green frog (*Rana clamitans*). *Journal of Toxicology and Environmental Health. Part A-Current Issues*, 2004. **67**: 941-957.
149. Hayes, T.B., Stuart, A.A., Mendoza, M., Collins, A., Noriega, N., Vonk, A., Johnston, G., Liu, R., and Kpodzo, D., Characterization of atrazine-induced gonadal malformations in African clawed frogs (*Xenopus laevis*) and comparisons with effects of an androgen antagonist (cyproterone acetate) and exogenous estrogen (17 β -estradiol): Support for the demasculinization/feminization hypothesis. *Environmental Health Perspectives*, 2006. **114**: 134-141.
150. Coady, K.K., Murphy, M.B., Villeneuve, D.L., Hecker, M., Jones, P.D., Carr, J.A., Solomon, K.R., Smith, E.E., Van Der Kraak, G., Kendall, R.J., and Giesy, J.P., Effects of atrazine on metamorphosis, growth, laryngeal and gonadal development, aromatase activity, and sex steroid concentrations in *Xenopus laevis*. *Ecotoxicology and Environmental Safety*, 2005. **62**: 160-173.
151. Solomon, K.R., Carr, J.A., Du Preez, L.H., Giesy, J.P., Kendall, R.J., Smith, E.E., and Van Der Kraak, G.J., Effects of atrazine on fish, amphibians, and aquatic reptiles: A critical review. *Critical Reviews in Toxicology*, 2008. **38**: 721-772.
152. Jowa, L. and Howd, R., Should atrazine and related chlorotriazines be considered carcinogenic for human health risk assessment? *Journal of Environmental Science and Health Part C-Environmental Carcinogenesis & Ecotoxicology Reviews*, 2011. **29**: 91-144.
153. Hauswirth, J.W., Farber, T.M., Burnam, W., Quest, J.A., Rinde, E., Levy, R., Barnes, D., Engler, R., Beliles, R., Barton, A., and Beal, D., Peer review of atrazine, USEPA, Editor. 1988, USEPA: Washington, D.C.
154. Ramdas, K.G. and Gerald, S.K., Biostimulation for the enhanced degradation of herbicides in soil. *Applied and Environmental Soil Science*, 2011.
155. Strong, L.C., McTavish, H., Sadowsky, M.J., and Wackett, L.P., Field-scale remediation of atrazine-contaminated soil using recombinant *Escherichia coli* expressing atrazine chlorohydrolase. *Environmental Microbiology*, 2000. **2**: 91-98.
156. Morgante, V., López-López, A., Flores, C., González, M., González, B., Vásquez, M., Rosselló-Mora, R., and Seeger, M., Bioaugmentation with *Pseudomonas* sp. Strain MHP41 promotes simazine attenuation and bacterial community changes in agricultural soils, 2010. **71**: 114-126.

157. Runes, H.B., Jenkins, J.J., and Bottomley, P.J., Atrazine degradation by bioaugmented sediment from constructed wetlands. *Applied Microbiology and Biotechnology*, 2001. **57**: 427-432.
158. Lima, D., Viana, P., André, S., Chelinho, S., Costa, C., Ribeiro, R., Sousa, J.P., Fialho, A.M., and Viegas, C.A., Evaluating a bioremediation tool for atrazine contaminated soils in open soil microcosms: The effectiveness of bioaugmentation and biostimulation approaches. *Chemosphere*, 2009. **74**: 187-192.
159. Gan, J., Becker, R.L., Koskinen, W.C., and Buhler, D.D., Degradation of atrazine in two soils as a function of concentration. *Journal of Environmental Quality*, 1996. **25**: 1064-1072.
160. Kawahigashi, H., Hirose, S., Ohkawa, H., and Ohkawa, Y., Phytoremediation of the herbicides atrazine and metolachlor by transgenic rice plants expressing human CYP1A1, CYP2B6, and CYP2C19. *Journal of Agricultural and Food Chemistry*, 2006. **54**: 2985-2991.
161. Wang, H.Z., Chen, X.W., Xing, X.G., Hao, X.H., and Chen, D.F., Transgenic tobacco plants expressing *atzA* exhibit resistance and strong ability to degrade atrazine. *Plant Cell Reports*, 2010. **29**: 1391-1399.
162. Scott, C., Begley, C., Taylor, M.J., Pandey, G., Momiroski, V., French, N., Brearley, C., Kotsonis, S.E., Selleck, M.J., Carino, F.A., Bajet, C.M., Clarke, C., Oakeshott, J.G., and Russell, R.J., Free-enzyme bioremediation of pesticides: A case study for the enzymatic remediation of organophosphorous insecticide residues. *Pesticide Mitigation Strategies for Surface Water Quality*, 2011. **1075**: 155-174.
163. Scott, C., Lewis, S.E., Milla, R., Taylor, M.C., Rodgers, A.J.W., Dumsday, G., Brodie, J.E., Oakeshott, J.G., and Russell, R.J., A free-enzyme catalyst for the bioremediation of environmental atrazine contamination. *Journal of Environmental Management*, 2010. **91**: 2075-2078.
164. Mutlu, B.R., Yeom, S., Tong, H.W., Wackett, L.P., and Aksan, A., Silicon alkoxide cross-linked silica nanoparticle gels for encapsulation of bacterial biocatalysts. *Journal of Materials Chemistry A*, 2013. **1**: 11051-11060.

Chapter 2. The structure of the hexameric atrazine chlorohydrolase AtzA

Pages 31-49

Overview

In the following paper, we present the X-ray crystal structure of AtzA, the 'archetypal' atrazine chlorohydrolase. This enzyme catalyzes the first step (dechlorination) in the catabolism of atrazine by *Pseudomonas* sp. strain ADP. Phylogenetic and biochemical evidence suggest that AtzA evolved recently from an ancestral deaminase, which makes AtzA a useful model system for studying the evolution of new enzymatic functions. Structural and *in silico* docking studies with the AtzA structure suggest that the active site of AtzA is less well adapted to its physiological role than the alternative metal-dependent atrazine chlorohydrolase, TrzN, which is catalytically superior to AtzA. A plausible catalytic mechanism, consistent with the structure presented herein and previously obtained biochemical and mutagenesis analyses, is also presented.

Contributions

Sahil Balotra: DNA cloning, mutagenesis, enzyme expression and purification, and enzyme characterization. Contributed to writing and figure preparation.

Drs Janet Newman and Tom Peat: Crystallography and structural determination

Dr Del Lucent: Substrate docking

Dr Andrew C. Warden: Metal ligands bond lengths analysis

Dr Colin Scott: Planning and supervision

This paper has been published: Peat *et al* (2015) *Acta Crystallographica D* **71**: 710-720.



The structure of the hexameric atrazine chlorohydrolase AtzA

T. S. Peat,^a J. Newman,^a S. Balotra,^{b,c} D. Lucent,^d A. C. Warden^c and C. Scott^{a*}

^aCSIRO Biomedical Manufacturing, Parkville, Australia, ^bResearch School of Chemistry, Australian National University, Canberra, Australia, ^cCSIRO Land and Water Flagship, Black Mountain, Canberra, Australia, and ^dDivision of Engineering and Physics, Wilkes University, Wilkes-Barr, Pennsylvania, USA. *Correspondence e-mail: colin.scott@csiro.au

Received 14 November 2014

Accepted 12 January 2015

Keywords: atrazine chlorohydrolase.

PDB references: AtzA, 4v1x; 4v1y

Supporting information: this article has supporting information at journals.iucr.org/d

Atrazine chlorohydrolase (AtzA) was discovered and purified in the early 1990s from soil that had been exposed to the widely used herbicide atrazine. It was subsequently found that this enzyme catalyzes the first and necessary step in the breakdown of atrazine by the soil organism *Pseudomonas* sp. strain ADP. Although it has taken 20 years, a crystal structure of the full hexameric form of AtzA has now been obtained. AtzA is less well adapted to its physiological role (*i.e.* atrazine dechlorination) than the alternative metal-dependent atrazine chlorohydrolase (TrzN), with a substrate-binding pocket that is under considerable strain and for which the substrate is a poor fit.

1. Introduction

The last century saw the introduction of a host of anthropogenic compounds, such as pesticides and antibiotics, into the environment (Russell *et al.*, 2011; Copley, 2000). This influx of new compounds has introduced a range of new selection pressures through factors such as toxicity and the availability of abundant potential nutrient sources (Copley, 2000; Russell *et al.*, 2011; Wackett, 2009). Nature has evolved mechanisms to cope with these new selection pressures, including the establishment of new metabolic pathways and enzymes. These new metabolic pathways and their associated enzymes are a tremendous resource for advancing our understanding of the mechanisms and constraints that underpin the evolutionary process, particularly with respect to the acquisition of new enzymatic functions.

The bacterial atrazine catabolism pathway from *Pseudomonas* sp. strain ADP is a particularly well studied example of a metabolic pathway that has evolved in response to human perturbations of the chemical composition of the environment. The pathway is comprised of six enzymatic steps (Fig. 1): the sequential hydrolysis of the chloride and two *N*-alkyl chains to produce cyanuric acid by AtzA (de Souza *et al.*, 1996; Scott *et al.*, 2009), AtzB (Boundy-Mills *et al.*, 1997; Seffernick *et al.*, 2007) and AtzC (Sadowsky *et al.*, 1998; Shapir *et al.*, 2002), followed by a ring-opening hydrolysis (AtzD; Fruchey *et al.*, 2003; Seffernick *et al.*, 2012; Peat *et al.*, 2013) and two deamination steps (AtzE and AtzF; Balotra *et al.*, 2014, 2015; Shapir, Sadowsky *et al.*, 2005; Cameron *et al.*, 2011; Martinez *et al.*, 2001). The catabolism of cyanuric acid possibly predates the introduction of anthropogenic herbicides, as cyanuric acid is a naturally occurring compound (albeit not an abundant one; Wackett, 2009). However, the enzymes that catalyze the first three steps of the pathway (AtzA, AtzB and AtzC) are likely to have evolved recently in response to the presence of atrazine in the environment (Wackett, 2009; Udiković-Kolić *et al.*, 2013).



OPEN ACCESS

The atrazine chlorohydrolyase AtzA has attracted a great deal of attention as a model for studying the evolutionary process, not least because of its relationship to melamine deaminase (TriA). AtzA and TriA share 98% identity, differing by just nine amino acids over their ~470-amino-acid length, and each difference is conferred by one of just nine point mutations (Seffernick *et al.*, 2001; Noor *et al.*, 2012). However, AtzA and TriA have quite distinct enzymatic functions, with AtzA only capable of hydrolytic dehalogenation (Fig. 1) and the deaminase TriA possessing only low levels of promiscuous dehalogenase activity (Seffernick *et al.*, 2001).

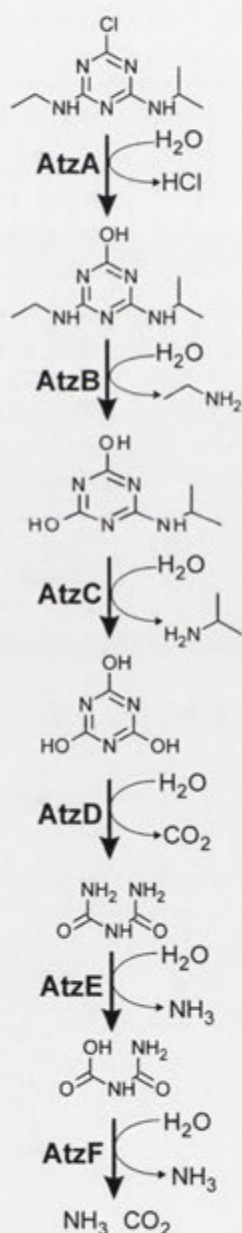


Figure 1
Atrazine catabolic pathway. The six hydrolytic steps of the atrazine catabolic pathway of *Pseudomonas* sp. strain ADP as catalysed by AtzA, AtzB, AtzC, AtzD, AtzE and AtzF are shown.

There have been studies that have used DNA shuffling (Raillard *et al.*, 2001) and evolutionary trajectory reconstruction (Noor *et al.*, 2012) to understand how these large functional differences could have evolved with so few genetic changes.

In other bacterial genera, particularly Gram-positive bacteria (*e.g.* *Arthrobacter* and *Nocardioideis*), the role of AtzA is occupied by the alternative chlorohydrolyase TrzN (Mulbry *et al.*, 2002; Sajjaphan *et al.*, 2004; Shapir *et al.*, 2006; Seffernick *et al.*, 2010). Both AtzA and TrzN belong to the same large family of amidohydrolyases, although they are so different physically and phylogenetically that it is likely that atrazine chlorohydrolyase activity evolved independently in each enzyme. While AtzA is a hexamer that contains one essential Fe²⁺ per monomer (de Souza *et al.*, 1996; Scott *et al.*, 2009), TrzN (which is ~25% identical to AtzA) is a dimer that has a single Zn²⁺ bound in each active site (Shapir *et al.*, 2006). Moreover, while AtzA can only hydrolyze triazine halides, TrzN can hydrolyze a broad range of substituents (including halide, -OCH₃ and -SCH₃ groups) from both triazines and pyrimidines (Shapir, Rosendahl *et al.*, 2005; de Souza *et al.*, 1996). TrzN is also an order of magnitude more efficient than AtzA, with a k_{cat}/K_m value of $\sim 10^5 \text{ s}^{-1} \text{ M}^{-1}$ compared with $\sim 1.5 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$. This difference is in large part owing to the relatively high K_m of AtzA, which is greater than the water solubility of atrazine (153 μM ; Scott *et al.*, 2009), compared with that of TrzN for atrazine ($\sim 20 \mu\text{M}$; Jackson *et al.*, 2014). In contrast to most other known hydrolytic dehalogenases, which use an active-site carboxylic acid (Asp) to displace the halide ion (Verschuere *et al.*, 1993; Newman *et al.*, 1999), the metal-dependent reaction mechanisms of AtzA and TrzN make these two enzyme lineages somewhat unusual in nature.

Despite intensive genetic and biochemical study of AtzA, the lack of an experimentally derived structural model has hampered efforts to understand the details of the sequence–function relationship. Here, we present the first X-ray structure of AtzA, upon which we have based a reinterpretation of the genetic and biochemical analyses of this model chlorohydrolyase.

2. Materials and methods

2.1. Protein expression and purification

The construction of the pCS150 expression vector containing the wild-type *atzA* gene has been described elsewhere (Scott *et al.*, 2009). The mutant *atzA* gene encoding the AtzA Ala170Thr, Met256Ile, Pro258Thr, Tyr261Ser variant was obtained from Life Technologies (Australia) and provided in pMK-RQ with NdeI and BamHI sites placed for subcloning into the pCS150 vector. pCS150 plasmids containing either the wild-type or mutant *atzA* genes were used to transform electrocompetent *Escherichia coli* BL21 λDE3 cells (Invitrogen) and were grown at 310 K on Luria–Bertani (LB; Lennox, 1955) agar [1.5% (w/v)] plates supplemented with 100 $\mu\text{g ml}^{-1}$ ampicillin. Overnight starter cultures of 50 ml were inoculated with a single colony. The overnight cultures were diluted 1:20 into 950 ml LB medium and were shaken at 310 K and

200 rev min⁻¹ until an OD₆₀₀ of 0.6–0.8 was obtained. Protein expression was initiated by the addition of 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). The induced cultures were kept at 310 K overnight whilst shaking at 200 rev min⁻¹.

The cells were then harvested by centrifugation at 4000g for 10 min in an Avanti J-E centrifuge (Beckman Coulter, Indianapolis, USA), resuspended in lysis buffer (50 mM HEPES, pH 7.5, 100 mM NaCl) and lysed by passage through an Avestin C3 homogenizer three times at 124 MPa. Insoluble cellular debris was removed by centrifugation at 21 000g using an Avanti J-E centrifuge.

AtzA and the AtzA variant were purified from lysed cells in three steps: His-tag affinity chromatography using an Ni-NTA Superflow Cartridge (Qiagen, Maryland, USA) with a gradient of 0–300 mM imidazole in 50 mM HEPES pH 7.5, 100 mM NaCl and size-exclusion chromatography using a 130 ml column packed with Superdex 200 prep-grade resin (GE Healthcare Life Sciences, Australia) with a buffer composed of 50 mM HEPES pH 7.5, 100 mM NaCl. The His tag was removed by thrombin proteolysis and the protein was then purified by size-exclusion chromatography with a Superdex 200 column. The protein was concentrated in an Amicon Ultra-15 Centrifugal Filter Unit with an Ultracel-30 membrane (Millipore, Carrigtwohill, Ireland) to 11.6 mg ml⁻¹ and snap-frozen in liquid nitrogen in 100 μl aliquots. The final purity was estimated to be 98% from a Coomassie-stained gel and typical yields were 5–7 mg purified protein from 1 l of LB medium.

2.2. Crystallization and structure solution

All crystallization experiments were performed in 96-well SD-2 plates (IDEX, USA) against a 50 μl reservoir. Initial crystals were obtained using droplets consisting of 150 nl concentrated protein solution (~11 mg ml⁻¹) combined with 150 nl reservoir solution and these were used in seeding experiments. The final crystals were obtained using droplets consisting of 150 nl concentrated protein solution combined with 120 nl reservoir solution and 30 nl seed stock. Either a Phoenix (ARI, USA) or a Mosquito (TTP Labtech, UK) robot was used to place the crystallization drops. Crystals (Supplementary Fig. S1) grew unreliably from a reservoir consisting of 5.5% (w/v) PEG 8000, 2.7% (v/v) diethylene glycol, 50 mM HEPES pH 7.1 at 281 K and were used to collect X-ray data on the MX-2 beamline of the Australian Synchrotron at a wavelength of 0.9537 Å (13 000 eV). The crystals were cryo-protected by the addition of reservoir solution supplemented with a further 20% diethylene glycol prior to cryocooling in liquid N₂. Data were indexed with XDS (Kabsch, 2010) and scaled using AIMLESS (Evans, 2011). The structure was solved using molecular replacement (*Phaser*; McCoy *et al.*, 2007) with PDB entry 3hpa (32% sequence identity, 15 gaps; Hall *et al.*, 2010). A clear solution was only found when a dimer of 3hpa was used, with six dimers placed in the asymmetric unit. The space group was found to be *P*2₁2₁ and the resolution of the data extended to 2.8 Å (see Table 1). The model was initially built using *Buccaneer* (Cowtan, 2006) and

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

PDB code	4vly	4vix
Space group	<i>P</i> 2 ₁ 2 ₁	<i>P</i> 2 ₁ 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 117.5, <i>b</i> = 195.6, <i>c</i> = 283.9, $\alpha = \beta = \gamma = 90.0$	<i>a</i> = 117.3, <i>b</i> = 146.1, <i>c</i> = 196.4, $\alpha = \beta = \gamma = 90.0$
Resolution (Å)	49.5–2.80 (2.85–2.80)	48.7–2.20 (2.24–2.20)
<i>R</i> _{merge}	0.229 (0.906)	0.235 (1.151)
<i>R</i> _{meas}	0.266 (1.042)	0.249 (1.223)
<i>R</i> _{p,i,m}	0.135 (0.531)	0.090 (0.438)
CC _{1/2}	0.989 (0.795)	0.995 (0.840)
<i>I</i> / σ (<i>I</i>)	8.8 (2.6)	11.6 (3.2)
Completeness (%)	100 (100)	99.6 (99.2)
Multiplicity	7.5 (7.6)	15.1 (15.3)
Refinement		
Resolution (Å)	49.5–2.8	48.7–2.2
Unique reflections	153180	170140
<i>R</i> _{work} / <i>R</i> _{free} (%)	19.3/22.2	16.6/19.6
No. of atoms		
Total	44810	23156
Protein	44565	22221
Other	105	42
Ions	15	6
Water	126	887
<i>B</i> factors (Å ²)		
Overall	29.2	20.1
Protein	29.4	19.7
Other	32.4	36.7
Ions	29.2	24.3
Water	11.8	22.8
R.m.s. deviations		
Bond lengths (Å)	0.014	0.017
Bond angles (°)	1.613	1.675

was then rebuilt by hand using *Coot* (Emsley *et al.*, 2010) and refined using *REFMAC* (Murshudov *et al.*, 2011). Two full hexamers were found to be present in the asymmetric unit, each being a trimer of dimers. The model refined to *R*_{work} and *R*_{free} values of 19.3 and 22.2%, respectively (Table 1). The Ramachandran plot (from *Coot*) shows 94.9% of the residues in the most favourable region, 3.9% in the allowed region and 1.1% in the outlier region. After this original structure was built, a second set of crystals were obtained under similar conditions [50 mM HEPES pH 7.3, 4.6% (w/v) PEG 10 000] and another data set was obtained on the MX-2 beamline with data that extended to 2.2 Å resolution. This crystal was determined to belong to space group *P*2₁2₁ and a single hexamer was found in the asymmetric unit. The model and structure factors for both structures have been deposited in the PDB as entries 4vix (2.2 Å resolution) and 4vly (2.8 Å resolution).

2.3. Molecular modelling

The gas-phase equilibrium conformation of the AtzA metal-coordination site was calculated using density functional theory with the B3LYP functional and the 6-31+G* basis set (with the LANL2DZ pseudopotential used for iron core electrons) as implemented in *Gaussian 09* (Frisch *et al.*, 2009).

Atrazine was docked into a flexible AtzA binding pocket using the *RosettaLigand* docking protocol with the coordi-

nates from PDB entry 4v1x. Partial atomic charges were calculated for the ligands using the AM1BCC Hamiltonian as implemented in *QuacPac* (v.1.6.3.1; OpenEye Scientific Software, Santa Fe, New Mexico, USA) and conformers were enumerated using *Omega* (v.2.5.1.4; OpenEye Scientific Software, Santa Fe, New Mexico, USA). 10 000 docking trajectories were performed using the *RosettaLigand* docking protocol (Davis *et al.*, 2009). The output ensembles were culled to retain the top 5% of structures according to the total energy (computed from the *Rosetta* energy function). From this smaller ensemble, the top 20 protein–ligand complexes

(according to protein–ligand interaction energy) were considered for final examination.

2.4. Dynamic light scattering

Concentrated AtzA (11.6 mg ml⁻¹) was diluted by 50% in a serial fashion with 50 mM HEPES pH 7.5, 100 mM NaCl to obtain a range of concentrations from 11.6 to 0.09 mg ml⁻¹. 20 µl volumes of each sample, in duplicate, were placed in a 384-well plate with blanks consisting of just buffer alone. The plate was placed in a DynaPro plate reader DLS machine (Wyatt Technology, Santa Barbara, California, USA) and 50

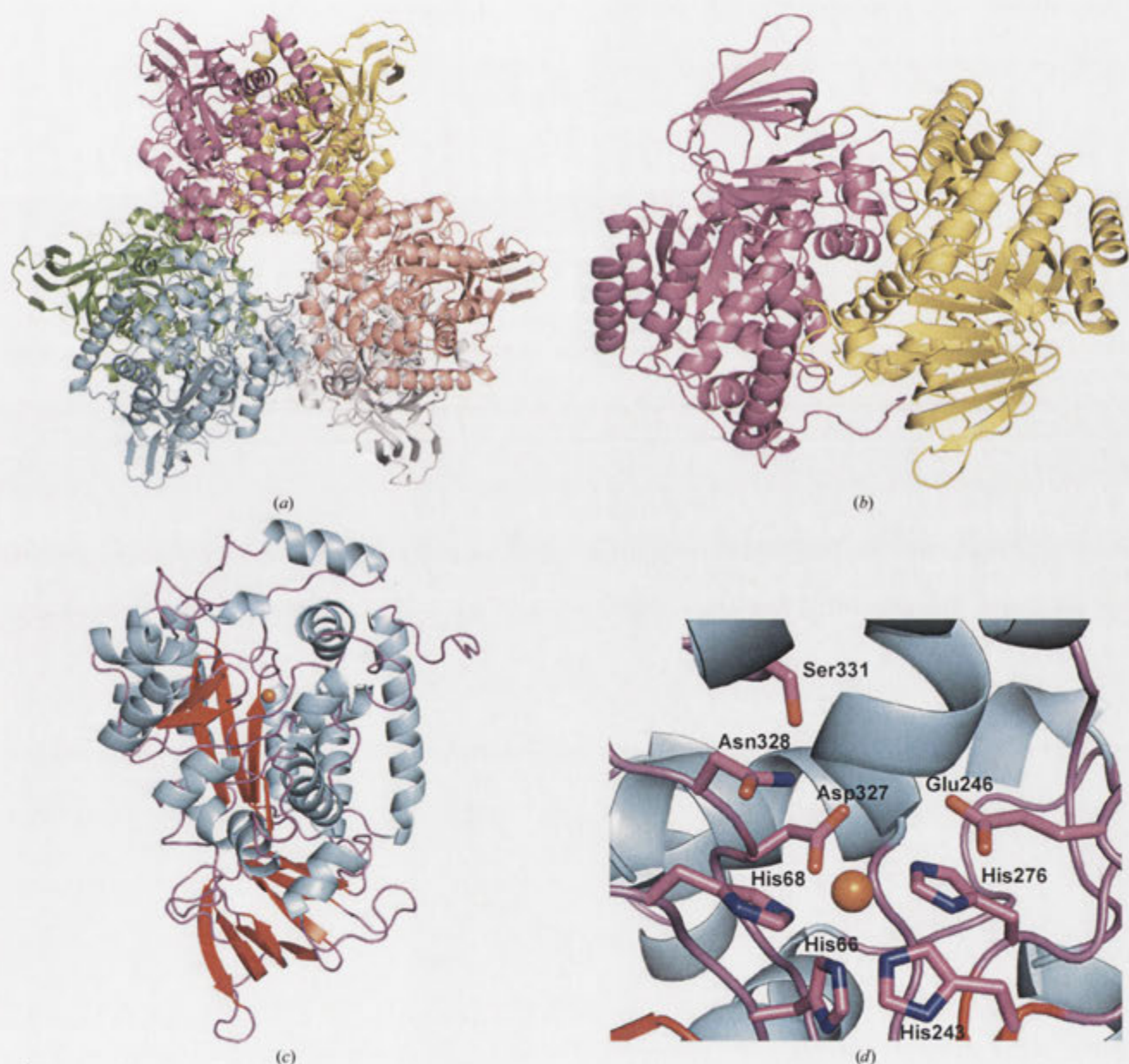


Figure 2
Structure of AtzA. (a) The X-ray structure of the AtzA hexamer is a trimer of dimers and is coloured by monomer. The structures of the dimer (b) (coloured by monomer), monomer (c) (coloured by secondary structure) and active site (d) are shown. The active-site metal (Fe²⁺) is shown as an orange sphere.

spectra were obtained at 297 K for each sample and averaged. The top three concentrations gave a radius of gyration of 5.9 nm, whereas the lowest concentration (0.09 mg ml^{-1}) gave an average radius of 5.2–5.3 nm. Measuring the hexamer always gave a diameter of greater than 10 nm (10.5 to 13.0 nm, depending on the points chosen), whereas the longest dimension seen for a dimer was 8.3 nm.

3. Results and discussion

3.1. The structure of hexameric AtzA

Analysis with *PDBFold* revealed that the closest structure in the PDB to AtzA was that used for molecular replacement, PDB entry 3hpa (r.m.s.d. of 1.6 Å; Hall *et al.*, 2010), which was also the structure with the greatest amino-acid sequence identity (32% over 411 residues). Another structural genomics target, PDB entry 3lnp (Kube *et al.*, 2013), is the next closest by structural homology, followed by two adenosine deaminases from *Xanthomonas campestris* (PDB entry 4dzh; New York Structural Genomics Research Consortium, unpublished work) and *Pseudomonas aeruginosa* (PDB entry 3pao; New York Structural Genomics Research Consortium, unpublished work). The TrzN structure (Seffernick *et al.*, 2010; PDB entry 3lsb), which has 26% sequence identity over 414 residues and an r.m.s.d. of 1.8 Å, is the fifth most similar structure to AtzA in the PDB. All of these molecules are classified as amidohydrolases, with 3hpa proposed to be an 8-oxoguanine deaminase. Although 3lnp has a highly homologous active site (four histidine residues and an aspartic acid residue), a calcium ion was modelled into the density instead of either Fe^{2+} or Zn^{2+} ; all of the other top hits on the list have either Fe^{2+} or (more commonly) Zn^{2+} in their active sites.

The AtzA hexamer (about 315 kDa in molecular weight) is a trimer of dimers (Fig. 2a), with the dimer interface being significantly larger than the individual protomer interfaces used to make up the hexamer (Figs. 2a and 2b): 3295 \AA^2 compared with 585 \AA^2 . The hexamer was confirmed by DLS and size-exclusion chromatography (SEC; Supplementary Fig. S2), consistent with previous reports for the solution state of the protein (de Souza *et al.*, 1996; Scott *et al.*, 2009). The dimer surface covers 17–18% of the monomer (a single monomer has a surface area of $\sim 18\,600 \text{ \AA}^2$; Fig. 2c). The dimer interface is similar to that found in the other amidohydrolases and the dimer is similar enough in overall structure that the search model used for molecular replacement (PDB entry 3hpa) gave a significantly more robust solution when the dimer was used instead of the monomer. Two full hexamers (over 600 kDa) were found in the asymmetric unit of the crystal structure in space group $P22_12_1$, whereas a single hexamer was found in space group $P2_12_12_1$. The two structures are essentially identical (r.m.s.d. of 0.3 Å), with the higher resolution data giving greater clarity in amino-acid positions and having significantly more water molecules added to the model. The most obvious interaction seen when looking down the threefold axis of the hexamer is the 11-amino-acid insertion (relative to PDB entry 3hpa) that forms the loop and helix encompassing residues 160–183 that has not been seen in the other amidohydrolase structures to date. The 'cavity' in the hexamer is approximately 14 Å across at the entrance, with one residue in particular, Arg163, forming much of the surface of this entrance (Fig. 3).

The *atzA* genes from *Aminobacter aminovorans* isolates recently collected from vineyards in France (Noor *et al.*, 2014) have accumulated a variety of mutations that have been

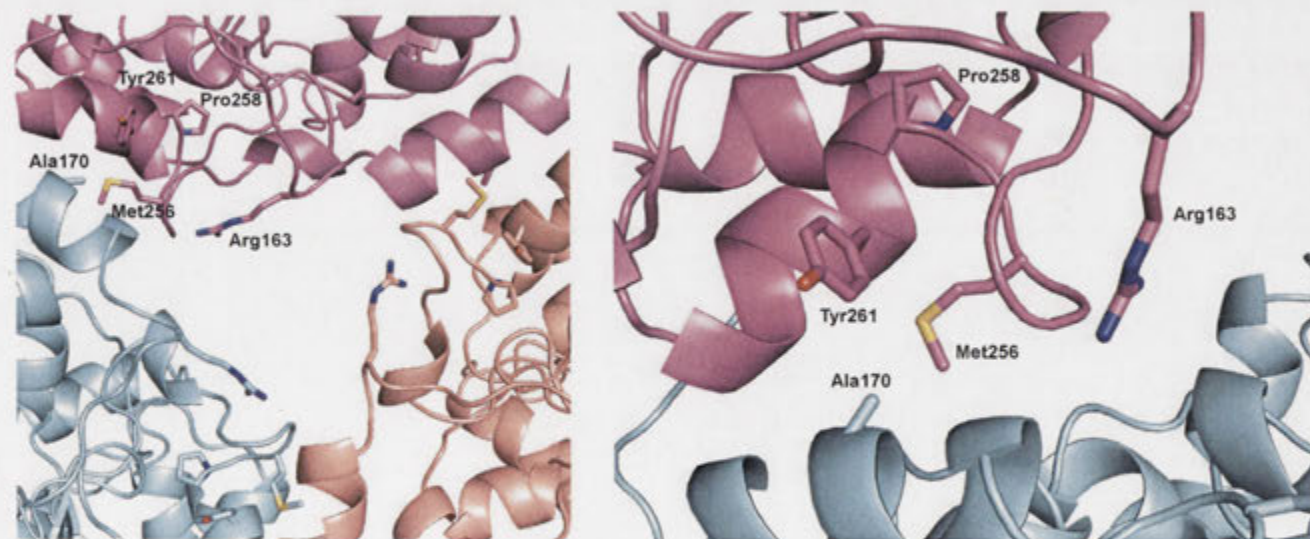


Figure 3

The entrance to the central cavity of the AtzA hexamer and the amino-acid residues of the hexamer-stabilizing interface. The entrance to the AtzA hexamer central cavity is largely formed by Arg163 contributed by each monomer. The hole formed in the hexamer is approximately 14 Å across at the entrance. Amino-acid substitutions at residues Ala170, Met256, Pro258 and Tyr261 destabilize the hexamer. The interactions between three monomers on one face of the hexamer (left) and at a single monomer-monomer interface (right) are shown.

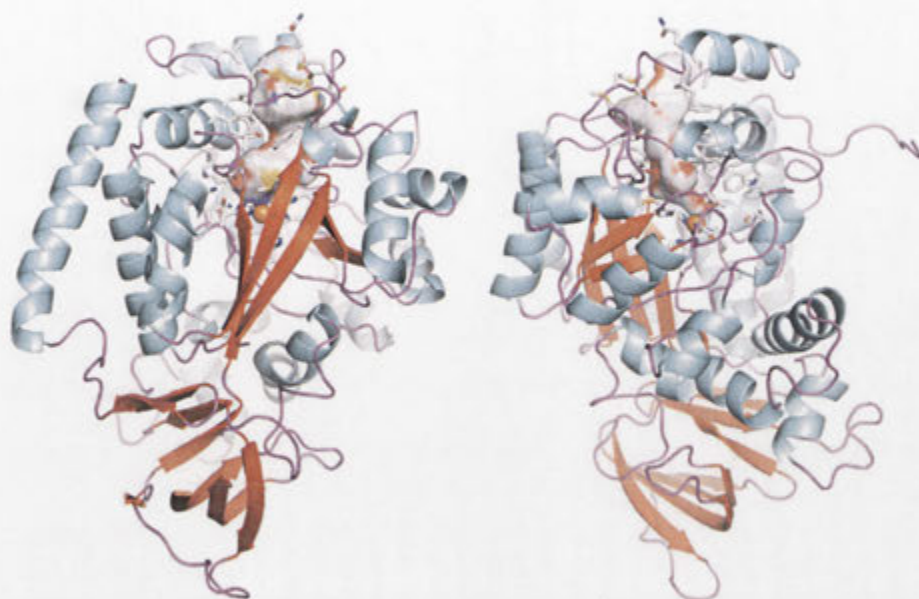


Figure 4

Access to the active site of AtzA. The surface of the active site and solvent-access channel (which form a single continuous channel from the active-site metal to the solvent) is shown in an AtzA monomer which has been coloured by secondary structure. The amino acids that form this surface (His66, His68, Gln71, Phe84, Tyr85, Trp87, Leu88, Phe89, Val92, Tyr93, Asp128, Met155, Phe157, Met160, Asp161, Ile164, Gln165, Val168, Leu180, Ser182, Ile183, Met184, Ala216, Thr217, Thr219, Ala220, His243, Glu246, Asp250, His276, Leu305, Asp327, Asn328 and Ser331) are shown as sticks, but have not been labelled for clarity. The amino-acid residues are labelled in Supplementary Fig. S5. The active-site metal (Fe^{2+}) is shown as an orange sphere. The images are rotated 180° around the vertical axis with respect to each other.

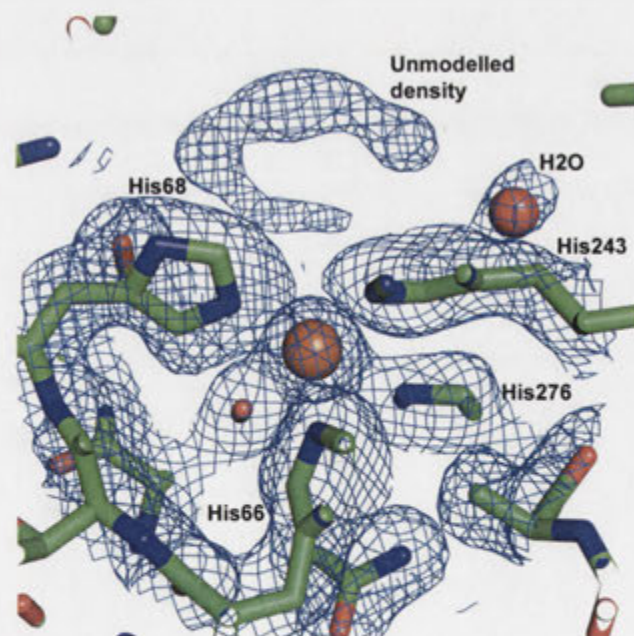


Figure 5

Density found in the active site of AtzA. A composite OMIT map was generated using the CCP4 package and the figure shows the 'extra' density seen in the active site which may be owing to atrazine (top of the figure, blue mesh). The active-site iron is shown as an orange sphere, water is shown as a red sphere and several active-site residues are shown with associated density. The OMIT map was set to 1σ .

demonstrated to confer changes in substrate specificity and alter the K_d of the enzyme for Fe^{2+} . Interestingly, four of the amino-acid changes conferred by these mutations (Ala170Thr, Met256Ile, Pro258Thr and Tyr261Ser) are located in the interface responsible for coordinating the AtzA hexamer (Fig. 3). AtzA variants carrying all four amino-acid substitutions were observed in the *A. aminovorans* isolates in the study by Noor *et al.* (2014). As these amino acids could effect a change in hexamer formation, the AtzA Ala170Thr, Met256Ile, Pro258Thr, Tyr261Ser variant was expressed and purified and its size was determined by SEC. Unlike the wild-type protein, which eluted from SEC with an estimated molecular weight of ~ 300 kDa, the variant eluted with an estimated molecular weight of ~ 100 kDa (*i.e.* a dimer; data not shown). This suggests that there is some factor that has selected dimeric AtzA in the more recent French isolates when compared with the AtzA from the original *Pseudomonas* isolate obtained from the USA, although it is uncertain whether the dimeric AtzA variants predate the hexameric form or *vice versa*.

3.2. Analysis of the AtzA active site

The substrate-binding pocket and catalytic centre of each monomer is accessible *via* a long hydrophobic channel. The channel, substrate-binding pocket and active site are comprised of His66, His68, Gln71, Phe84, Tyr85, Trp87, Leu88, Phe89, Val92, Tyr93, Asp128, Met155, Phe157, Met160, Asp161, Ile164, Gln165, Val168, Leu180, Ser182, Ile183, Met184, Ala216, Thr217, Thr219, Ala220, His243, Glu246, Asp250, His276, Leu305, Asp327, Asn328 and Ser331 (Fig. 4). The identities of these residues are in good agreement with a homology model used to successfully predict sites for mutagenesis to substantially reduce the K_m for atrazine (*i.e.* Ala216, Thr217, Thr218, Ala219 and Asp250; Scott *et al.*, 2009).

The mononuclear active-site metal is coordinated by His66, His68, His243, His276 and Asp327 (Fig. 2*d*); it is a subtype III amidohydrolase metal-binding motif similar to that found in *E. coli* adenine deaminase (Seibert & Rauschel, 2005). The bond lengths between the bound cation and ligating amino-acid side chains are longer than would be expected for a metal centre of this type, with lengths in the following ranges: Fe–His66, 2.68–2.88 Å; Fe–His68, 2.76–2.96 Å; Fe–His243, 3.01–

3.14 Å; Fe–His276, 2.76–2.99 Å; Fe–Asp327, 2.41–2.49 Å. Although the conformations of the metal-coordinating side chains found in the active site are indicative of an octahedral complex (as would be expected for Fe²⁺ binding), the geometry is distorted from that predicted by gas-phase electronic structure calculations (Supplementary Fig. S3). Iron-coordinating histidine residues at these distances are observed in other structures in the PDB, although they are far less frequent than those with shorter bond lengths (the mean His–Fe bond length is 2.2 Å, with a standard deviation of ~0.2 Å; Supplementary Fig. S4). Although no negative density was seen when the Fe atoms were set to 100% occupancy, the resulting *B* factors were over double those for the nearby residues. As a result, we have set the occupancies of the Fe atoms to 50%, which gives much more reasonable *B* factors. These results imply that the Fe atom is bound but not tightly coordinated in the active site, consistent with the unusually high observed *K_d* of AtzA for iron (~5 μM; Noor *et al.*, 2014). Interestingly, Fe²⁺ bound in related amidohydrolases (e.g. cytosine deaminases; PDB entries 1ra0, 1k70 and 4r88; Ireton *et al.*, 2001; Mahan *et al.*, 2004) is also found to be less tightly coordinated than Zn²⁺ cations, despite Fe²⁺ being the physiologically relevant cation.

There are a number of Ramachandran outliers associated with the active site, His66, His276, Glu298, Asp327 and Asn332, which all have good electron density (Fig. 5) and are unambiguously placed. His66, His276 and Asp327 bind the Fe²⁺ in the active site directly and introduce strain into the active site. Most of the other Ramachandran outliers are also in good electron density. The one exception is Glu251, which is in a region of weak electron density and is likely to be mobile in the protein. This mobile region sits over the substrate and

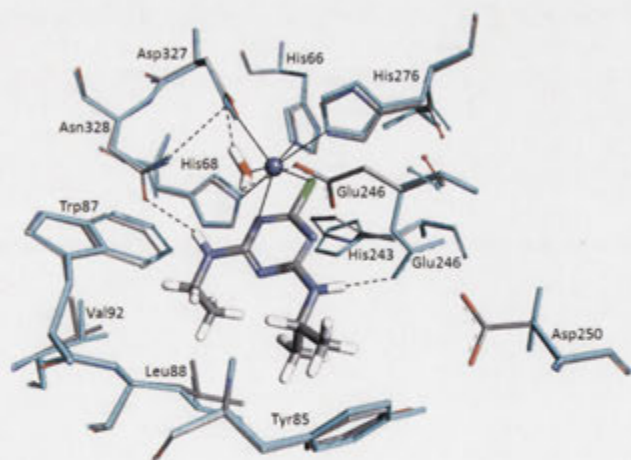


Figure 6
Simulated docking of atrazine in the AtzA active site: the highest scoring docked pose produced using the *RosettaLigand* docking protocol with the coordinates from PDB entry 4v1x that placed atrazine within a physically reasonable distance for nucleophilic substitution by the activated water. Although it is notionally possible to rotate the plane of the atrazine ring through 180°, the steric environment of the binding pocket appear to prohibit this alternative binding mode. The crystal structure (grey) and the modelled structure produced during docking simulations (cyan) are shown. The active-site metal (Fe²⁺) is shown as a blue sphere.

may be mobile in order to allow substrate entry or exit. One other Ramachandran outlier, Thr368, is interesting as it is in the middle of a helix; the preceding residue, Ala367, which is also part of the helical structure, has its carbonyl hydrogen-bonded to another residue (Ile63) instead of being in line with the other carbonyls in the helix.

Examination of the binding pocket indicated that it might be difficult for atrazine to bind in an orientation that would permit nucleophilic substitution by an activated water (as is the case in TrzN). Atrazine was both soaked into crystals and co-crystallized with the protein, with the co-crystallization experiment resulting in crystals that diffracted to 2.2 Å resolution. However, only a small amount of broken density was found in the active site in some of the monomers of the co-crystals (Fig. 5). The density in the active site was insufficient to unambiguously place atrazine. For this reason, atrazine docking was investigated using extensive molecular-docking calculations with the *RosettaLigand* docking protocol (Davis *et al.*, 2009). The results of these calculations suggest that nonproductive binding modes are common for atrazine.

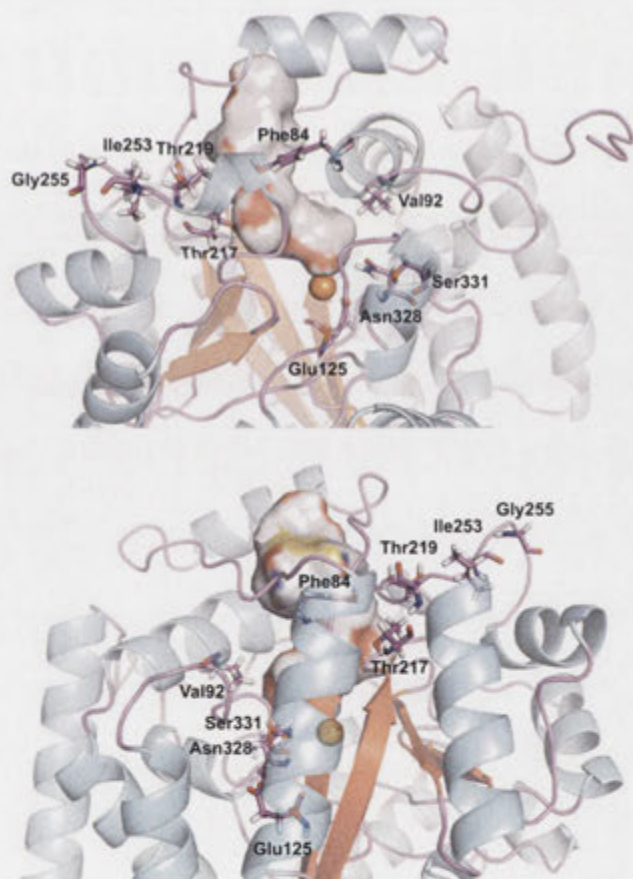


Figure 7
Differences between AtzA and the melamine deaminase TriA. The positions of Phe84, Val92, Asp125, Thr217, Thr219, Ile253, Gly255, Asn328 and Ser331, which distinguish AtzA from TriA, are shown in an AtzA monomer (coloured by secondary structure). The active-site metal (Fe²⁺) is shown as an orange sphere and the surface of the active site/solvent-access channel is shown. The images are rotated 180° around the vertical axis with respect to each other.

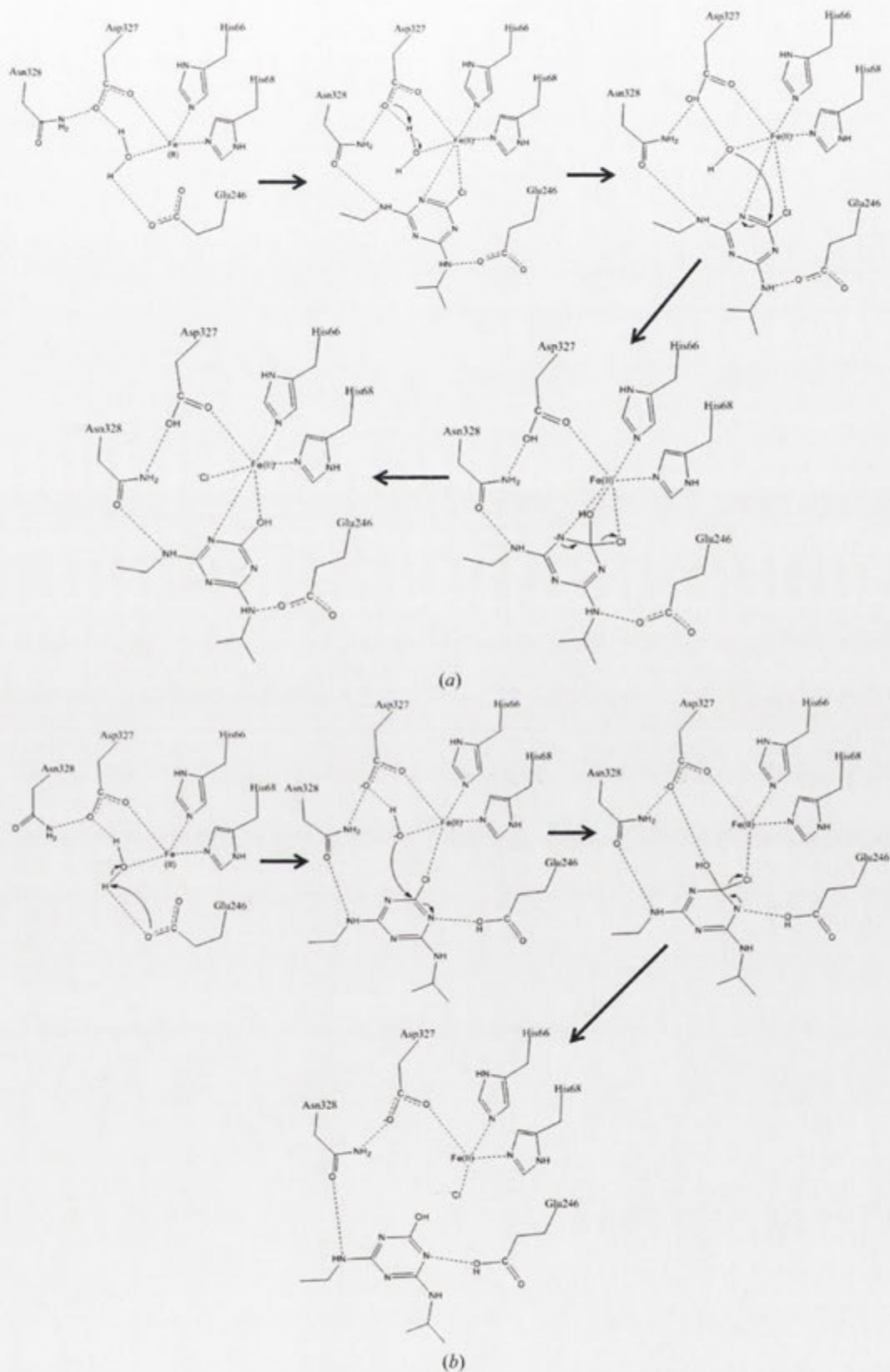


Figure 8

Plausible reaction mechanisms for AtzA. Two plausible reaction mechanisms are proposed involving either bidentate (a) or monodentate (b) coordination of the atrazine Cl atom to the Fe²⁺ centre of the AtzA active site.

A high-scoring pose was produced that positioned the appropriate C atom of atrazine for nucleophilic attack by the metal-bound water. In contrast to our previously proposed mechanism based upon homology modelling (Scott *et al.*, 2009), the *N*-ethyl and *N*-isopropyl moieties of the substrate are oriented away from the metal centre, making hydrophobic contacts with Val92, Trp87, Leu88, Tyr85 and Phe84 (Fig. 6). Phe84 presents a different rotamer to that found in the substrate-free crystal structure. The Cl atom and a ring N atom of atrazine are coordinated to the Fe²⁺, which positions the substrate ideally for nucleophilic attack by the coordinated nucleophile.

To determine whether the bidentate coordination mode is likely or whether there is monodentate coordination through the Cl atom, not requiring rotation of His243 but requiring a slight outwards movement of the hydrophobic pocket described above, would involve further investigation that is beyond the scope of the current work. Regardless, the lateral shift to form the monodentate coordination mode would be less than 1 Å and would be able to be accommodated without major reorganization of the active site.

There are two hydrogen-bonding interactions from the *N*-ethyl and *N*-isopropyl substituents to the side-chain O atom of Asn328 and the side chain of the new rotamer of Glu246, respectively. The latter of these residues may play a role in stabilizing the negative charge on a ring N atom during the formation of the tetrahedral intermediate. The symmetry of the triazine ring of the subunit could make it possible for the *N*-alkyl side chains of atrazine to bind in the opposite orientation to that presented here; however, the additional volume required to accommodate the *N*-isopropyl side chain compared with the *N*-ethyl side chain makes this alternate binding mode unlikely from a steric perspective.

The substrate-binding pocket of AtzA appears to be ill-adapted for binding atrazine, with strained conformers of key amino residues, a low-affinity metal-binding site and a high probability of unproductive substrate binding. These observations are consistent with the high K_m of wild-type AtzA for atrazine, which exceeds the aqueous solubility of the substrate (*i.e.* >159 μM; Scott *et al.*, 2009), and may explain why we have been unable to capture clear X-ray structures of the AtzA–atrazine complex.

3.3. Location of the nine amino acids that distinguish AtzA and the melamine deaminase TriA

The melamine deaminase TriA differs from AtzA in the identities of just nine amino-acid

residues: Leu84 (Phe), Leu92 (Val), Asp125 (Glu), Ile217 (Thr), Pro219 (Thr), Leu253 (Ile), Trp255 (Gly), Asp328 (Asn) and Cys331 (Ser) (AtzA residues shown in parentheses; Seffernick *et al.*, 2001). Of these nine amino-acid residues, six are part of the substrate-binding pocket/access channel in AtzA: Phe84, Val92, Thr217, Thr219, Asn328 and Ser331 (Fig. 7).

The presence of Phe84, Asn328 and Ser331 in the substrate-binding pocket is unsurprising, as these three amino acids had previously been shown to be the major determinants of atrazine/melamine specificity in AtzA and TriA, respectively (Noor *et al.*, 2012; Scott *et al.*, 2009; Raillard *et al.*, 2001). In particular, Ser331 and Asn328 are located within the vicinity of the active-site metal, which is consistent with their previously proposed role in catalysis in TriA-mediated melamine deamination and passive promotion of atrazine-mediated atrazine dechlorination (Noor *et al.*, 2012; Scott *et al.*, 2009).

In laboratory evolution experiments, altering Thr217 and Thr219 has been demonstrated to lead to substantial improvements in the K_m and k_{cat}/K_m values of AtzA for atrazine (Scott *et al.*, 2009). Thr217 and Thr219 were targeted for mutagenesis as homology modelling had predicted their presence in the substrate-binding pocket, and the AtzA structure presented here confirms this (Fig. 7). Ile253 and Gly255 are not located in the substrate-binding pocket. Interestingly, however, Ile253 and Gly255 do cluster close to Thr219 (Fig. 7), with Ile253 within 4 Å of Thr219. Changes to the residues at positions 253 and 255 may therefore influence the packing of the residues of the substrate-binding pocket (*i.e.* positions 217 and 219), which is consistent with their relatively minor role in 'tuning' the specificity of the enzyme for melamine or atrazine (Noor *et al.*, 2012; Raillard *et al.*,

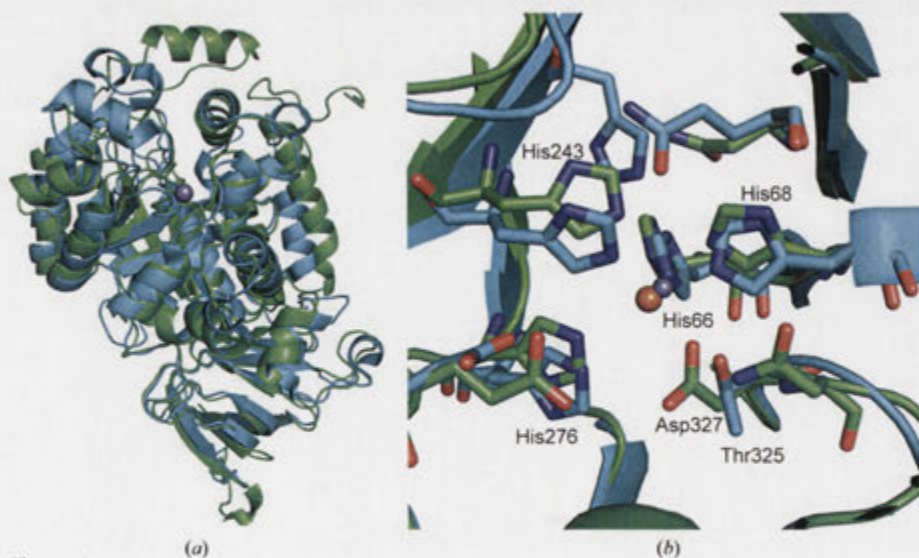


Figure 9
Comparison of AtzA with the alternative atrazine chlorohydrolase TrzN. (a) Superposition of an AtzA monomer (green) and a TrzN monomer (cyan). (b) Superposition of the active sites of AtzA (green) and a TrzN monomer (cyan). The active-site metals (Fe²⁺ and Zn²⁺; shown in orange and grey, respectively) are shown and the AtzA metal-binding ligands are labelled in green. The unique threonine (Thr325) of TrzN, which is equivalent to Asp327 in AtzA, is also labelled in cyan.

2001). The remaining difference between AtzA and TriA is at position 125 (Glu in AtzA and Asp in TriA) and has been shown to contribute substantially to the specificity of AtzA and TriA for dechlorination or deamination (Noor *et al.*, 2012). Although Glu125 is buried (Fig. 7), it does contribute to the active-site geometry *via* a hydrogen-bonding interaction between the Glu125 side chain and the backbone N atom of Val69, next to His68 which is bound to the Fe²⁺.

3.4. The catalytic mechanism of AtzA

We present two alternative potential catalytic mechanisms for the hydrolytic dechlorination of atrazine by AtzA that are constant with the *in silico* docking presented here, notwithstanding that empirical data will be required to test these and previously proposed reaction mechanisms. The first is based directly upon the interactions shown in the only docking pose that oriented atrazine appropriately for nucleophilic attack and the second is based upon the abovementioned monodentate coordination mode for atrazine (Figs. 8*a* and 8*b*, respectively) that better enables Glu246 to stabilize negative charge on another of the ring N atoms. In both of these mechanisms, coordination of the Cl atom to Fe²⁺ draws electron density from the C–Cl bond, increasing the susceptibility of the C atom to nucleophilic attack. In the first mechanism, the coordinated water is deprotonated by Asp327 followed by nucleophilic attack of the resulting activated hydroxide at the chlorine-bearing ring C atom. The negative charge of the tetrahedral intermediate is stabilized on the Fe²⁺-coordinated aromatic N atom, after which chloride is released and aromaticity is re-established in the hydroxylated product.

The second mechanism with monodentate atrazine coordination to Fe²⁺ involves deprotonation of the coordinated water by Glu246, which then rotates to donate a hydrogen bond to a ring N atom of atrazine upon substrate binding. This N atom then gains the negative charge of the tetrahedral intermediate, stabilized by the hydrogen bond from the protonated Glu246, followed by generation of the products as per the first mechanism.

It is possible that the 'loose' coordination geometry of the Fe²⁺ has evolved to facilitate the bidentate coordination of atrazine, which would not be possible with four ideally coordinating residues and a coordinated water. His243 in particular, at over 3 Å from the Fe²⁺, provides sufficient space for bidentate coordination, albeit by adopting a higher energy conformation. We note, however, that such a binding mode is unprecedented, and without a crystal structure of the AtzA–atrazine complex there is no empirical evidence to support it.

While the mechanisms for atrazine dechlorination that we have proposed, in light of the crystal structure and docking, differ from those that we have previously reported based upon homology modelling and manual placement of the substrate, the newly proposed mechanisms remain consistent with the empirically determined biochemical and mutagenesis data reported elsewhere (de Souza *et al.*, 1996; Raillard *et al.*, 2001; Scott *et al.*, 2009; Noor *et al.*, 2012). The sequential mutations reported earlier (Scott *et al.*, 2009) that reduce the atrazine dechlorination activity and increase the deaminase activity

may do so through combinations of adjusted positioning of the metal-bound water through shifted hydrogen-bond networks, altering its p*K*_a, and a changed steric factor in substrate binding and product release.

3.5. A comparison of the two metal-dependent atrazine chlorohydrolases AtzA and TrzN

The tertiary structure of AtzA is similar to that of TrzN (Seffernick *et al.*, 2010). The monomers of AtzA and TrzN superpose with a root-mean-square difference of 1.8 Å on C^α atoms over 414 residues (Fig. 9*a*). The metal-binding histidine side chains and metal overlay (Fig. 9*b*), with the major differences being the longer metal–histidine distances (His66, His68 and His243 in AtzA at 2.7–3.1 Å instead of 2.0–2.1 Å in TrzN; Fig. 9*b*), Fe²⁺ instead of Zn²⁺ as the metal, the additional His276 interaction with Fe²⁺ and the additional Fe–Asp327 interaction (at 2.4–2.5 Å). In the TrzN structure, the histidine equivalent to His276 (His274) makes a 2.7 Å hydrogen bond to the water coordinated to the Zn ion. Uniquely among the amidohydrolase family, the TrzN structure has a threonine residue at position 325, making it the first member of the ninth subgroup of metal-coordination sites of the amidohydrolase family (Seibert & Raushel, 2005; Jackson *et al.*, 2014; Khurana *et al.*, 2009).

Unlike the active site of AtzA, the TrzN active site has excellent complementarity for its substrate, which almost certainly accounts for the lower *K*_m of TrzN for atrazine (~20 μ*M*; Shapir *et al.*, 2002) when compared with that of AtzA for the same substrate (>159 μ*M*). The mechanism of TrzN differs from any proposed for AtzA, in which a zinc-activated water forms the tetrahedral intermediate and highly activated leaving groups (such as halides) depart readily. Leaving groups with higher p*K*_a values (such as amines) can be hydrolyzed by TrzN as they are protonated; Glu241 first donates a proton to the triazine ring, which then protonates the leaving group. Thus, TrzN is an efficient hydrolase with a broad substrate range, while AtzA is a less active enzyme that can perform either deamination or dehalogenation reactions depending on the identities of the amino-acid residues at positions 328 and 331. AtzA variants that are effectively intermediates between AtzA and TriA possess both of these activities at the same time, but with a lower specificity than either AtzA or TriA (Noor *et al.*, 2012).

4. Concluding remarks

For the first time since its discovery ~20 years ago, we have been able to determine the structure of AtzA, the 'archetypal' atrazine chlorohydrolase. AtzA appears to be structurally ill-adapted to perform its physiological function, with a metal-coordination geometry that leads to a low affinity for its metal cofactor and a substrate-binding pocket that appears to be unable to accept its physiological substrate without substantial reorganization. In addition, four of the active-site amino-acid residues appear to be held in strained conformations, suggesting that the active site is far from optimized, which is consistent with the relatively poor kinetic performance of

AtzA (k_{cat}/K_m of $\sim 1 \times 10^4 M^{-1} s^{-1}$ and a K_m in excess of the substrate's solubility limit of 153 μM ; Scott *et al.*, 2009; Noor *et al.*, 2012). AtzA is certainly less well adapted to its physiological function than the alternative atrazine chlorohydrolase TrzN (k_{cat}/K_m of $\sim 1 \times 10^5 M^{-1} s^{-1}$ and a K_m value of $\sim 20 \mu M$; Jackson *et al.*, 2014), which has good complementarity for its substrate and a catalytic mechanism that permits a broader substrate range than AtzA. Both *atzA* and *trzN* are prone to horizontal gene transfer, as they are located on transmissible plasmids (de Souza *et al.*, 1998; Sajjaphan *et al.*, 2004), and it will be interesting to see whether the catalytically superior TrzN begins to outcompete and replace AtzA in the environment.

Acknowledgements

We would like to thank Drs Tara Sutherland and Judith Scoble (CSIRO) for their constructive comments during the preparation of this manuscript. We thank the beamline scientists of the Australian Synchrotron and the CSIRO Collaborative Crystallization Centre (<http://www.csiro.au/C3>), Melbourne, Australia. We also thank OpenEye Scientific Software for a license for *Afit*, *QuacPac* and *Omega*.

References

Balotra, S., Newman, J., Cowieson, N., French, N. G., Campbell, P. M., Briggs, L. J., Warden, A. C., Easton, C. J., Peat, T. S. & Scott, C. (2015). *Appl. Environ. Microbiol.* **81**, 470–480.

Balotra, S., Newman, J., French, N. G., Briggs, L. J., Peat, T. S. & Scott, C. (2014). *Acta Cryst.* **F70**, 310–315.

Boundy-Mills, K. L., de Souza, M. L., Mandelbaum, R. T., Wackett, L. P. & Sadowsky, M. J. (1997). *Appl. Environ. Microbiol.* **63**, 916–923.

Cameron, S. M., Durchschein, K., Richman, J. E., Sadowsky, M. J. & Wackett, L. P. (2011). *ACS Catal.* **1**, 1075–1082.

Copley, S. D. (2000). *Trends Biochem. Sci.* **25**, 261–265.

Cowtan, K. (2006). *Acta Cryst.* **D62**, 1002–1011.

Davis, I. W., Raha, K., Head, M. S. & Baker, D. (2009). *Protein Sci.* **18**, 1998–2002.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.

Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.

Frisch, M. J. *et al.* (2009). *Gaussian 09*. Gaussian Inc., Wallingford, Connecticut, USA.

Fruchey, I., Shapir, N., Sadowsky, M. J. & Wackett, L. P. (2003). *Appl. Environ. Microbiol.* **69**, 3653–3657.

Hall, R. S., Fedorov, A. A., Marti-Arbona, R., Fedorov, E. V., Kolb, P., Sauder, J. M., Burley, S. K., Shoichet, B. K., Almo, S. C. & Raushel, F. M. (2010). *J. Am. Chem. Soc.* **132**, 1762–1763.

Ireton, G. C., McDermott, G., Black, M. E. & Stoddard, B. L. (2001). *J. Mol. Biol.* **315**, 687–697.

Jackson, C. J., Coppin, C. W., Carr, P. D., Aleksandrov, A., Wilding, M., Sugrue, E., Ubels, J., Paks, M., Newman, J., Peat, T. S., Russell, R. J., Field, M., Weik, M., Oakshott, J. G. & Scott, C. (2014). *Appl. Environ. Microbiol.* **80**, 4003–4011.

Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.

Khurana, J. L., Jackson, C. J., Scott, C., Pandey, G., Horne, I., Russell, R. J., Herlt, A., Easton, C. J. & Oakshott, J. G. (2009). *Biochem. J.* **418**, 431–441.

Kube, M. *et al.* (2013). *Nature Commun.* **4**, 2156–2159.

Lennox, E. S. (1955). *Virology*, **1**, 190–206.

Mahan, S. D., Ireton, G. C., Knoeber, C., Stoddard, B. L. & Black, M. E. (2004). *Protein Eng. Des. Sel.* **17**, 625–633.

Martinez, B., Tomkins, J., Wackett, L. P., Wing, R. & Sadowsky, M. J. (2001). *J. Bacteriol.* **183**, 5684–5697.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.

Mulbry, W. W., Zhu, H., Nour, S. M. & Topp, E. (2002). *FEMS Microbiol. Lett.* **206**, 75–79.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.

Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A., Holmes, I. H., Schindler, J. F., Unkefer, C. J. & Terwilliger, T. C. (1999). *Biochemistry*, **38**, 16105–16114.

Noor, S., Changey, F., Oakshott, J. G., Scott, C. & Martin-Laurent, F. (2014). *Biodegradation*, **25**, 21–30.

Noor, S., Taylor, M. C., Russell, R. J., Jermini, L. S., Jackson, C. J., Oakshott, J. G. & Scott, C. (2012). *PLoS One*, **7**, e39822.

Peat, T. S., Balotra, S., Wilding, M., French, N. G., Briggs, L. J., Panjikar, S., Cowieson, N., Newman, J. & Scott, C. (2013). *Mol. Microbiol.* **88**, 1149–1163.

Raillard, S., Kriebber, A., Chen, Y., Ness, J. E., Bermudez, E., Trinidad, R., Fullem, R., Davis, C., Welch, M., Seffernick, J., Wackett, L. P., Stemmer, W. P. C. & Minshull, J. (2001). *Chem. Biol.* **8**, 891–898.

Russell, R. J., Scott, C., Jackson, C. J., Pandey, R., Pandey, G., Taylor, M. C., Coppin, C. W., Liu, J. W. & Oakshott, J. G. (2011). *Evol. Appl.* **4**, 225–248.

Sadowsky, M. J., Tong, Z., de Souza, M. & Wackett, L. P. (1998). *J. Bacteriol.* **180**, 152–158.

Sajjaphan, K., Shapir, N., Wackett, L. P., Palmer, M., Blackmon, B., Tomkins, J. & Sadowsky, M. J. (2004). *Appl. Environ. Microbiol.* **70**, 4402–4407.

Scott, C., Jackson, C. J., Coppin, C. W., Mourant, R. G., Hilton, M. E., Sutherland, T. D., Russell, R. J. & Oakshott, J. G. (2009). *Appl. Environ. Microbiol.* **75**, 2184–2191.

Seffernick, J. L., Aleem, A., Osborne, J. P., Johnson, G., Sadowsky, M. J. & Wackett, L. P. (2007). *J. Bacteriol.* **189**, 6989–6997.

Seffernick, J. L., de Souza, M. L., Sadowsky, M. J. & Wackett, L. P. (2001). *J. Bacteriol.* **183**, 2405–2410.

Seffernick, J. L., Erickson, J. S., Cameron, S. M., Cho, S., Dodge, A. G., Richman, J. E., Sadowsky, M. J. & Wackett, L. P. (2012). *J. Bacteriol.* **194**, 4579–4588.

Seffernick, J. L., Reynolds, E., Fedorov, A. A., Fedorov, E., Almo, S. C., Sadowsky, M. J. & Wackett, L. P. (2010). *J. Biol. Chem.* **285**, 30606–30614.

Seibert, C. M. & Raushel, F. M. (2005). *Biochemistry*, **44**, 6383–6391.

Shapir, N., Osborne, J. P., Johnson, G., Sadowsky, M. J. & Wackett, L. P. (2002). *J. Bacteriol.* **184**, 5376–5384.

Shapir, N., Pedersen, C., Gil, O., Strong, L., Seffernick, J., Sadowsky, M. J. & Wackett, L. P. (2006). *J. Bacteriol.* **188**, 5859–5864.

Shapir, N., Rosendahl, C., Johnson, G., Andreina, M., Sadowsky, M. J. & Wackett, L. P. (2005). *Appl. Environ. Microbiol.* **71**, 2214–2220.

Shapir, N., Sadowsky, M. J. & Wackett, L. P. (2005). *J. Bacteriol.* **187**, 3731–3738.

Souza, M. L. de, Sadowsky, M. J. & Wackett, L. P. (1996). *J. Bacteriol.* **178**, 4894–4900.

Souza, M. L. de, Wackett, L. P. & Sadowsky, M. J. (1998). *Appl. Environ. Microbiol.* **64**, 2323–2326.

Udiković-Kolić, N., Scott, C. & Martin-Laurent, F. (2013). *Appl. Microbiol. Biotechnol.* **96**, 1175–1189.

Verschueren, K. H. G., Seljée, F., Rozeboom, H. J., Kalk, K. H. & Dijkstra, B. W. (1993). *Nature (London)*, **363**, 693–698.

Wackett, L. P. (2009). *Curr. Opin. Microbiol.* **12**, 244–251.



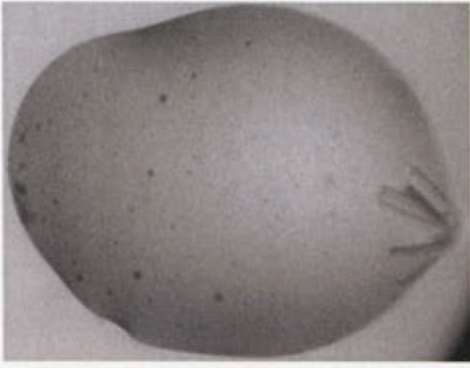
BIOLOGICAL
CRYSTALLOGRAPHY

Volume 71 (2015)

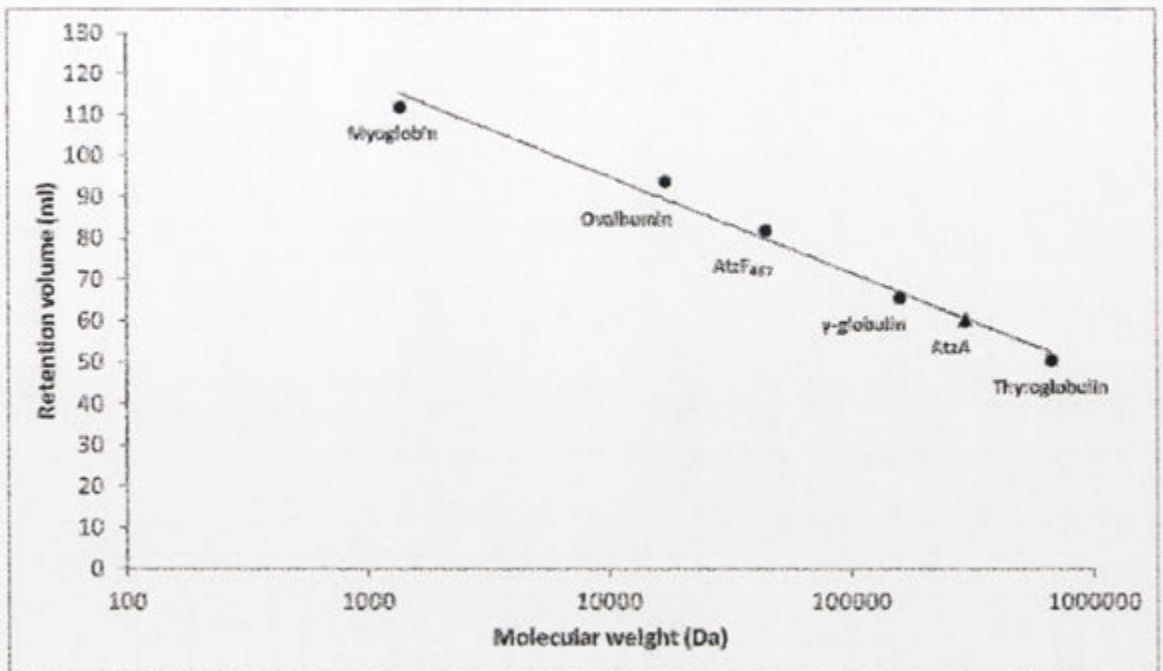
Supporting information for article:

The structure of the hexameric atrazine chlorohydrolase AtzA

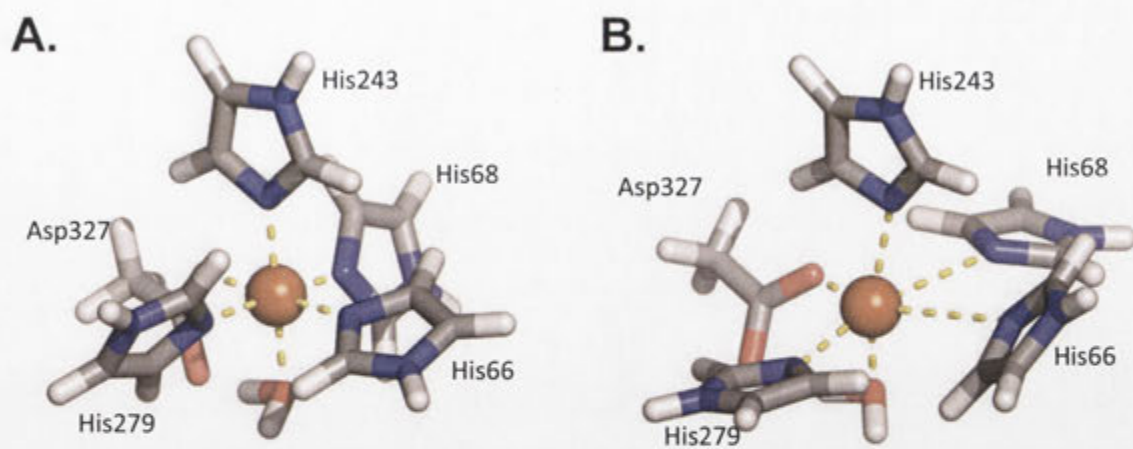
T. S. Peat, J. Newman, S. Balotra, D. Lucent, A. C. Warden and C. Scott



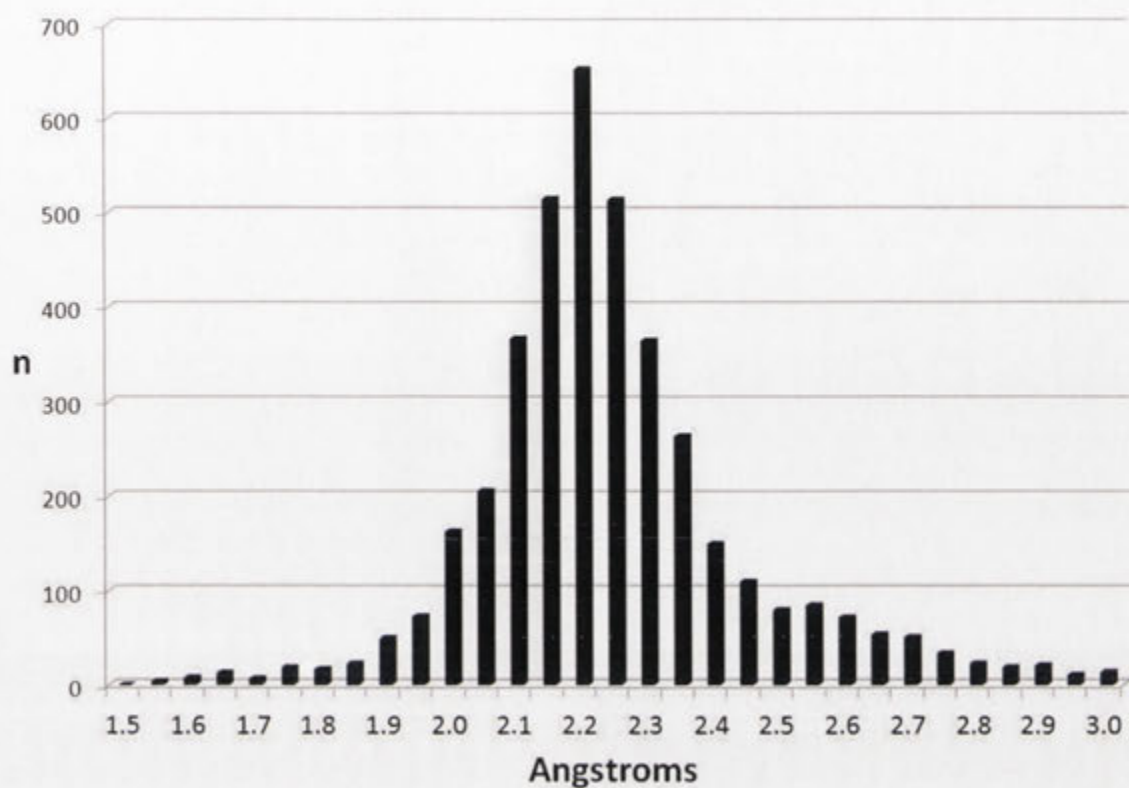
Supplementary Figure S1. AtzA Crystal.



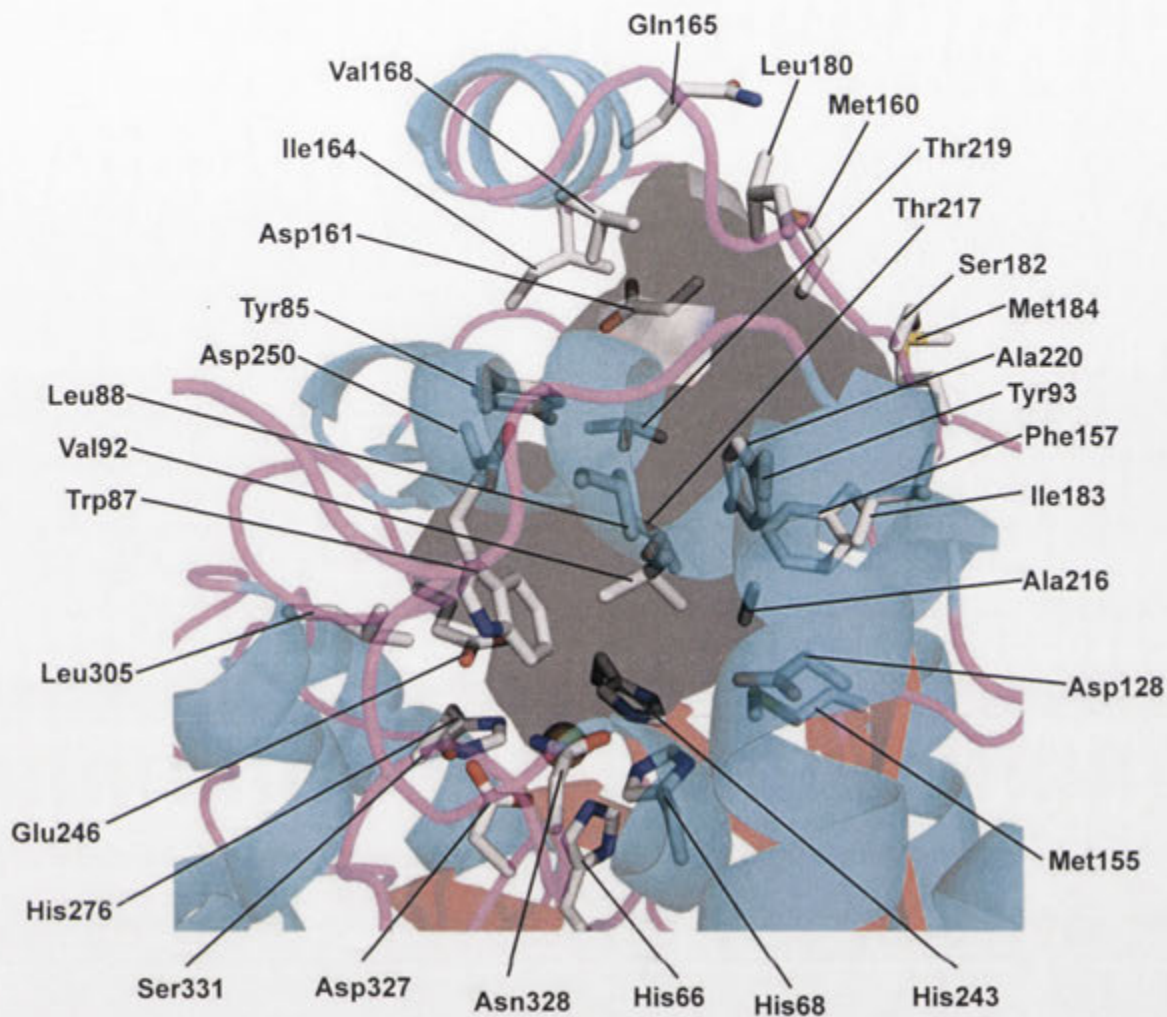
Supplementary Figure S2. Size exclusion chromatography (SEC) the AtzA hexamer.



Supplementary Figure S3. The coordination geometry of the Fe^{2+} in AtzA predicted by gas-phase electronic structure calculation (A) and from the crystal structure (B).



Supplementary Figure S4. Histogram showing the distribution of coordination bond lengths (< 3 Angstroms) between Fe²⁺ and the coordinating nitrogen in histidines in 517 structures in the Brookhaven Protein Databank.



Supplementary Figure S5. The surface of the active site and solvent access channel is shown in an AtzA monomer, which has been colored by secondary structure. The amino acids that form this surface (His66, His68, Gln71, Phe84, Tyr85, Trp87, Leu88, Phe89, Val92, Tyr93, Asp128, Met155, Phe157, Met160, Asp161, Ile164, Gln165, Val168, Leu180, Ser182, Ile183, Met184, Ala216, Thr217, Thr219, Ala220, His243, Glu246, Asp250, His276, Leu305, Asp327, Asn328 and Ser331) are shown as sticks, and are labelled. The active-site metal (Fe²⁺) is shown as an orange sphere.

Chapter 3. Purification and crystallographic studies with AtzB (hydroxyatrazine hydrolase).

Pages 50-60

Overview

This chapter describes the over-expression and purification of AtzB (hydroxyatrazine hydrolase). First isolated from *Pseudomonas* sp. strain ADP, this enzyme performs the second step in atrazine mineralization pathway, i.e., removal of alkylamine side chain. Unfortunately, efforts to crystalize AtzB have yet to succeed.

Contributions

Sahil Balotra: I performed the DNA cloning, enzyme expression and purification and enzyme characterization described in this chapter.

Drs Janet Newman and Tom Peat: Crystallographic trials

Lygie Esquirol: DSF experiment

Dr Colin Scott: Planning and supervision

3.1 Introduction

The environmental fate of herbicide atrazine has been extensively studied over the last six decades [1, 2]. In bacteria, the mineralisation of these compounds occurs via a six step process that is dependent upon six different hydrolases which are atrazine chlorohydrolase (AtzA, or TrzN, EC 3.8.1.8), hydroxyatrazine hydrolase (AtzB, EC 3.5.99.3), *N*-isopropylammelide isopropyl amidohydrolase (AtzC, EC 3.5.99.4), cyanuric acid hydrolase (AtzD, E.C. 3.5.2.15), biuret hydrolase (AtzE, EC 3.5.1.84), and allophanate hydrolase (AtzF, EC 3.5.1.54) [3-13]. The hydrolytic cascade is initiated by one of two enzymes, AtzA or TrzN [14, 15]. AtzA has very strict substrate specificity acts on 2-halo-1,3,5-triazine compounds, whereas TrzN can hydrolyse broader range of *s*-triazines such as ametryn, atrazine, and cyanoatrazine and ametryn sulfoxide, that carry different substituents at the 2 position (Fig. 1) [14, 15]. Although, a number of different *s*-triazine compounds can be transformed by AtzA and TrzN, they are all transformed to 2-hydroxy-1,3,5-triazine compounds.

AtzB, a metal-dependent amidohydrolase with 37% sequence similarity with AtzA, is the only enzyme identified to date that can transform 2-hydroxy-1,3,5-triazine compounds to 2,4-dihydroxy-1,3,5-triazines (Fig. 2) [3]. AtzB is homodimeric, with a mononuclear metal centre in each monomer as inferred from a 1:1 stoichiometry of enzyme to metal [3]. Both Fe^{2+} and Zn^{2+} were found in similar abundance in the purified enzyme; therefore, it may be that AtzB does not distinguish between the two cations and a mixed population of Zn^{2+} -AtzB and Fe^{2+} -AtzB is present in the cell (albeit, the relative affinities for these metals have not yet been determined) [3]. Surprisingly, AtzB can perform both dechlorination and deamination reactions and acts on fairly large number of substrates (Fig. 2) [3] but in the

absence of structural information the mechanistic basis for such a broad substrate specificity is not known.

Herein, we have purified this enzyme and have presented the attempts made so far to crystallise this enzyme to obtain its X-ray structure for greater insight into its enzyme-substrate interactions.

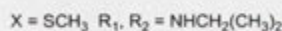
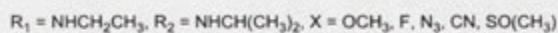
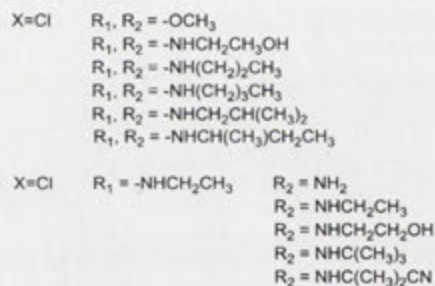
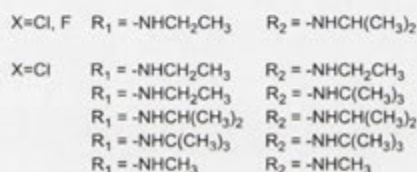


Figure 1. Substrate range of the two atrazine chlorohydrolases, AtzA and TrzN. Substituents at the 2-, 4-, and 6- positions (X, R₁ and R₂, respectively) in s-triazine substrates for AtzA and TrzN are shown. Figure adapted from Shapir *et al.*, 2005, and Seffernick *et al.*, 2000. [14, 15]

3.2 Materials and methods

3.2.1 Cloning

The *atzB* gene was obtained from GenScript (New Jersey, USA); the coding sequence used was identical to that described by Martinez *et al.* [7] (accession No. U66917) and was provided as an *NdeI/BamHI* insert in pUC57 (supplied by GenScript). The *NdeI/BamHI* restriction fragment containing the *atzB* gene was subcloned into the *NdeI* and *BamHI* sites of pETCC2 vector. The pETCC2 expression vector was received from Dr. Christopher Coppin (CSIRO Land and Water) as a gift and has been described elsewhere [6]. As a result of subcloning, a hexahis-tag along with thrombin cleavage site (MGSSHHHHHSSGLVPRGSH) is added to the N-terminus of the enzyme.

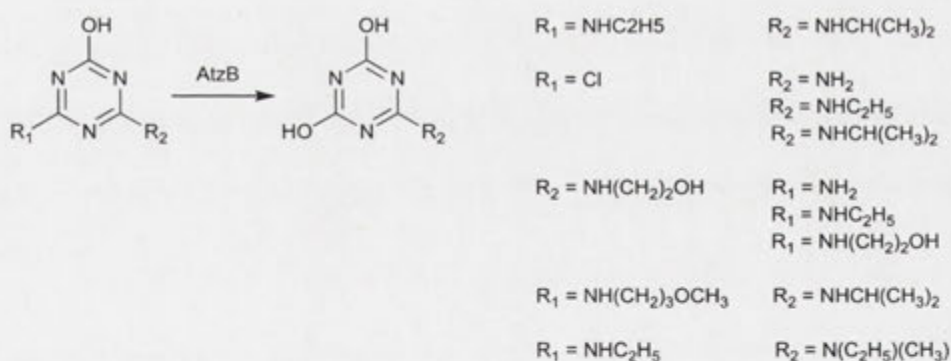


Figure 2. Substrate range of AtzB. Substituents at the 2-, 4-, and 6- positions (X, R_1 and R_2 , respectively) in s-triazine substrates for AtzB. Figure adapted from Seffernick *et al.*, 2007 [3]

3.2.2 Expression and purification of AtzB

The *atzB* gene, cloned into the *NdeI* and *BamHI* sites of the pETCC2 plasmid, was used to transform electrocompetent *Escherichia coli* BL21 (λ DE3) cells (Invitrogen). Transformants

were grown at 310 K on Luria–Bertani (LB) [16] agar plates supplemented with $100 \mu\text{g.mL}^{-1}$ ampicillin. Starter cultures were prepared by using single colony of transformed cells to inoculate 50 mL of LB supplemented with $100 \mu\text{g.mL}^{-1}$ ampicillin and shaken at 200 rev min^{-1} at 310 K overnight. The overnight starter cultures were diluted 1:20 with 950 mL LB medium and shaken at 310 K and 200 rev.min^{-1} until an OD_{600} of 0.6-0.8 was attained. Protein expression was induced by adding $200 \mu\text{M}$ isopropyl β -D-1-thiogalactopyranoside (IPTG) and the induced cultures were kept at 303.2 K overnight, whilst shaking at 200 rev.min^{-1} .

The cells were harvested by centrifugation at $4000 g$ for 10 min in an Avanti J-E centrifuge (Beckman Coulter, Indianapolis, USA). The collected cells were then resuspended in lysis buffer (50 mM HEPES pH 7.5, 100 mM NaCl) and lysed by passage through an Avestin C3 homogenizer three times at 124 MPa. The insoluble cellular debris was removed from the homogenised cells by centrifugation at $21,000 g$ using an Avanti J-E centrifuge.

AtzB was purified from the soluble cell-free extract in one step: His-tag affinity chromatography using an Ni–NTA Superflow cartridge (Qiagen, Maryland, USA) with a first two step gradients of 25 mM imidazole and 57.5 mM imidazole in 50 mM HEPES, 100 mM NaCl (pH 7.5) and finally the protein was eluted in 250 mM imidazole in 50 mM HEPES, 100 mM NaCl buffer (pH 7.5). The protein was concentrated in an Amicon Ultra-15 Centrifugal Filter Unit with an Ultracel-30 membrane (Millipore, Carrigtwohill, Ireland) to 0.37 mg.mL^{-1} and snap-frozen in liquid nitrogen in 100 mL aliquots. The final purity was estimated by Coomassie staining.

3.2.3 Stability analysis

Heat-denaturation curves of hexaHis-tagged AtzB were generated using differential scanning fluorimetry. The sample was tested in a suite of different buffers and pHs in triplicate [17]. The assay was performed in a CFX96 RT-PCR machine (Bio-Rad) with 19.6 μL of each screening condition, 300 nL protein at 6-7 $\text{mg}\cdot\text{mL}^{-1}$ and 300 nL of a 1:10 (aqueous) dilution of SYPRO Orange dye (Sigma).

3.2.4 Crystallisation

All crystallisation experiments were performed by Dr Janet Newman at the C3 facility in CSIRO Manufacturing Flagship, Melbourne, Australia. Screens were done in 96-well SD-2 plates (IDEX, USA) against a 50 μL reservoir. Either a Phoenix (Art Robbins Instruments, USA) or a Mosquito (TTP LabTech, UK) robot was used to place the crystallisation drops. The JCSG+, PACT_Com7, PS gradient, Morpheus and ProPlex screens [18] were tried during the crystallographic trials with *either in situ* thrombin treated AtzB or intact AtzB and the detailed information about these screens is available at <http://c6.csiro.au/>.

3.3 Results and discussion

3.3.1 Expression and purification of AtzB

The *atzB* gene was successfully expressed as a soluble protein in BL21 (λDE3) strain of *E. coli* and the expressed protein was purified by affinity chromatography using Ni-NTA column (Fig. 3). The single column purification resulted in the purification of the protein and the

purity of the protein as estimated by Coomassie stained SDS PAGE > 98% (Fig 3). The yield of AtzB was estimated to be 1.5-2.5 mg.L⁻¹.

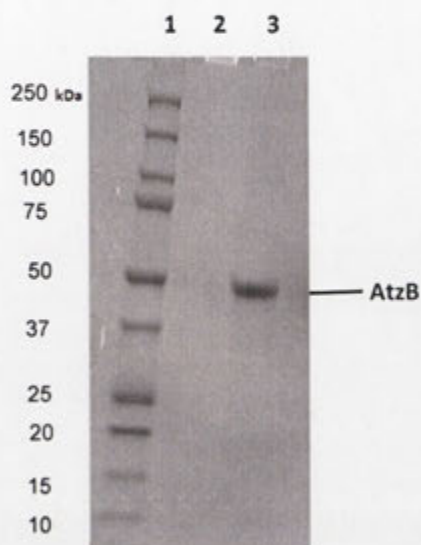


Figure 3. SDS-PAGE purified AtzB. Lane 1) molecular weight protein standards; lane 2) empty lane; lane 3) purified AtzB after single step purification by affinity chromatography.

3.3.2 Stability analysis

Thermal stability studies of AtzB showed that the protein is correctly folded as the transition from the native to the unfolded state follows a first-order phase transition with a T_m of ~62°C (Lygie Esquirol, personal communication).

3.3.3 Crystallographic trials

Unfortunately, initial attempts at crystallography with hexaHis-tagged AtzB proved unsuccessful. In part this was due to the purified protein precipitating upon concentration, which limited the concentration of hexaHis-tagged AtzB that could be tested under each crystallography condition to just 0.37 mg.ml⁻¹. The removal of his-tag *via* thrombin-

mediated hydrolysis resulted in a protein that could be concentrated to $> 2 \text{ mg.mL}^{-1}$. This concentrated AtzB will be used in further crystallography trials.

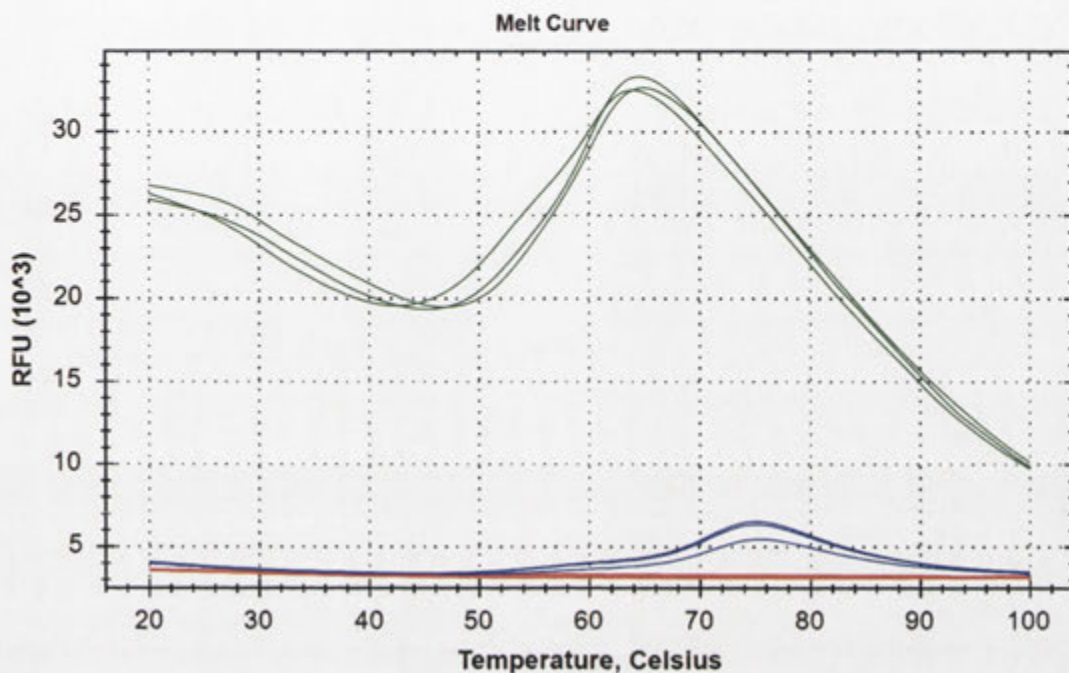


Figure 4. ThermoFluor assay of AtzB. A temperature melting curve of AtzB was performed in triplicate, using differential scanning fluorimetry with the dye SYPRO Orange. The green curves are the AtzB protein (in triplicate) in the buffer comprised of 50mM *N*-(2-Acetamido) iminodiacetic acid and 200mM NaCl at pH 6.5. The blue curves are a 0.1 mg.mL^{-1} lysozyme control and the red curves are dye only control. This experiment was conducted by Lygie Esquirol.

3.4 Conclusion and future work

The attempts to crystallize AtzB have failed so far. AtzB with intact his-tag precipitated upon concentrating it beyond 0.37 mg.mL^{-1} , which may be too low a concentration for crystal nucleation. As the removal of his-tag showed nearly five-fold increase in the concentration at which AtzB remained soluble, thrombin-treated AtzB may provide more suitable material for crystallization trials. Additional modifications to the crystallography

protocol, including co-crystallisation with substrates, inhibitors and protein-stabilising small molecules. Additionally, it may be possible to stabilise AtzB by mutagenesis, allowing structural determination of stable AtzB variants.

AtzB has been shown to be a mixed population of iron and zinc isoforms and it would be interesting to explore whether these isoforms are functionally identical, or contribute to the promiscuous dechlorinase activity observed in AtzB differentially. Experiments to establish the relative K_d of AtzB for a range of different metals, along with analysis of isoform catalytic function and physical characterisation (e.g. structural differences between isoforms, pH and temperature profiles, etc.) were planned. However, there was insufficient time to complete this work during the course of this PhD study.

References

1. Scott, C., Jackson, C.J., Coppin, C.W., Mourant, R.G., Hilton, M.E., Sutherland, T.D., Russell, R.J., and Oakeshott, J.G., Catalytic improvement and evolution of atrazine chlorohydrolase. *Applied and Environmental Microbiology*, 2009. **75**: p. 2184-2191.
2. Udikovic-Kolic, N., Scott, C., and Martin-Laurent, F., Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology*, 2012. **96**: p. 1175-1189.
3. Seffernick, J.L., Aleem, A., Osborne, J.P., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Hydroxyatrazine N-ethylaminohydrolase (AtzB): An amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *Journal of Bacteriology*, 2007. **189**: p. 6989-6997.
4. deSouza, M.L., Sadowsky, M.J., and Wackett, L.P., Atrazine chlorohydrolase from *Pseudomonas* sp strain ADP: Gene sequence, enzyme purification, and protein characterization. *Journal of Bacteriology*, 1996. **178**: p. 4894-4900.
5. Sadowsky, M.J., Tong, Z.K., de Souza, M., and Wackett, L.P., AtzC is a new member of the amidohydrolase protein superfamily and is homologous to other atrazine-metabolizing enzymes. *Journal of Bacteriology*, 1998. **180**: p. 152-158.
6. Peat, T.S., Balotra, S., Wilding, M., French, N.G., Briggs, L.J., Panjekar, S., Cowieson, N., Newman, J., and Scott, C., Cyanuric acid hydrolase: Evolutionary innovation by structural concatenation. *Molecular Microbiology*, 2013. **88**: p. 1149-1163.

7. Martinez, B., Tomkins, J., Wackett, L.P., Wing, R., and Sadowsky, M.J., Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. Strain ADP. *Journal of Bacteriology*, 2001. **183**: p. 5684-5697.
8. 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., X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. *Applied and Environmental Microbiology*, 2015. **81**: p. 470-480.
9. Topp, E., Mulbry, W.M., Zhu, H., Nour, S.M., and Cuppels, D., Characterization of *s*-triazine herbicide metabolism by a *Nocardioide*s sp isolated from agricultural soils. *Applied and Environmental Microbiology*, 2000. **66**: p. 3134-3141.
10. Balotra, S., Newman, J., French, N.G., Briggs, L.J., Peat, T.S., and Scott, C., Crystallization and preliminary x-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp strain ADP. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*, 2014. **70**: p. 310-315.
11. Wackett, L.P., Sadowsky, M.J., Martinez, B., and Shapir, N., Biodegradation of atrazine and related *s*-triazine compounds: From enzymes to field studies. *Applied Microbiology and Biotechnology*, 2002. **58**: p. 39-45.
12. Wang, Y., Li, X., Chen, X., and Chen, D., Directed evolution and characterization of atrazine chlorohydrolase variants with enhanced activity. *Biochemistry-Moscow*, 2013. **78**: p. 1104-1111.
13. Schomburg, I., Chang, A., and Schomburg, D., Brenda, enzyme data and metabolic information. *Nucleic Acids Research*, 2002. **30**: p. 47-49.
14. Shapir, N., Rosendahl, C., Johnson, G., Andreina, M., Sadowsky, M.J., and Wackett, L.P., Substrate specificity and colorimetric assay for recombinant TrzN derived from *Arthrobacter aurescens* tc1. *Applied and Environmental Microbiology*, 2005. **71**: p. 2214-2220.
15. Seffernick, J.L., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Substrate specificity of atrazine chlorohydrolase and atrazine-catabolizing bacteria. *Applied and Environmental Microbiology*, 2000. **66**: p. 4247-4252.
16. Jeffrey, H.M., *Experiments in molecular genetics*. 1972, New York: Cold Spring Harbor Laboratory Press.
17. Seabrook, S.A. and Newman, J., High-throughput thermal scanning for protein stability: Making a good technique more robust. *ACS Combinatorial Science*, 2013. **15**: p. 387-392.
18. Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J.L., Stuart, D.I., and Perrakis, A., Towards rationalization of crystallization screening for small- to medium-sized academic laboratories: The PACT/JCSG plus strategy. *Acta Crystallographica Section D-Biological Crystallography*, 2005. **61**: p. 1426-1431.

Chapter 4. X-ray structure and mutagenesis studies of the *N*-isopropylammelide isopropylaminohydrolase, AtzC

Pages 61-87

Overview

The *N*-isopropylammelide isopropylaminohydrolase (AtzC) isolated from *Pseudomonas* sp. strain ADP performs the third hydrolytic step in the atrazine mineralization pathway to remove the isopropylamine side chain from the *N*-isopropylammelide. Notwithstanding the relatively low resolution X-ray structure of AtzC that was available in the PDB prior to the commencement of these studies, there was no available structure/function analysis of this enzyme. The work herein describes the provision of an AtzC structure with bound inhibitor, which guided *in silico* substrate-docking experiments. In concert with mutagenesis studies, these experiments allowed us to propose a plausible reaction mechanism for AtzC. The reaction mechanism of AtzC appears to be similar to that of alpha carbonic anhydrase. However, the low sequence identity between alpha carbonic anhydrase and AtzC suggest that these enzymes have evolved this mechanism independently of each other.

Contributions

Sahil Balotra: DNA cloning, mutagenesis, enzyme expression and purification, experiment design and data analysis, assay development and enzyme characterization. I co-authored the paper, contributed to writing and figure preparation.

Drs Janet Newman and Tom Peat: Crystallography and structural determination

Lyndall J. Briggs: Gene cloning

Dr Andrew C. Warden: Substrate docking

Dr Colin Scott: Planning and supervision

The work reported herein is under review for publication in PLoS One.

X-ray structure and mutagenesis studies of the *N*-isopropylammelide isopropylaminohydrolase, AtzC

Sahil Balotra^{1,2*}, Andrew C. Warden^{1*}, Janet Newman³, Lyndall J. Briggs¹, Colin Scott¹, Thomas S. Peat³

¹ *CSIRO Land and Water Flagship, Black Mountain, Canberra, Australia*

² *Research School of Chemistry, Australian National University, Canberra, Australian Capital Territory, Australia*

³ *CSIRO Biomedical Manufacturing, Parkville, Australia*

* Authors contributed equally

Corresponding authors:

Colin.scott@csiro.au

Thomas.peat@csiro.au

ABSTRACT

The *N*-isopropylammelide isopropylaminohydrolase from *Pseudomonas* sp. strain ADP, AtzC, provides the third hydrolytic step in the mineralization of *s*-triazine herbicides, such as atrazine. We obtained the X-ray crystal structure of AtzC at 1.84 Å with an inhibitor bound in the active site and then used a combination of *in silico* docking and site-directed mutagenesis to understand the interactions between AtzC and its substrate, isopropylammelide. The substitution of an active site histidine residue (His249) for an alanine abolished the enzyme's catalytic activity. We propose a plausible catalytic mechanism, consistent with the biochemical and crystallographic data obtained that is similar to that found in carbonic anhydrase and other members of subtype III of the amidohydrolase family.

Introduction

Since its introduction in the 1950s, atrazine has been used extensively for weed control in corn, sorghum and sugarcane, which has resulted in the detection of atrazine in ground and surface water at concentrations of up to 4.6 μM . At such concentrations, atrazine is reportedly responsible for endocrine disruption in vertebrates, causing chemical feminization in frogs [1]. The availability of high concentrations of atrazine in the environment has promoted the evolution of new metabolic pathways in bacteria that allow the use of atrazine as a carbon and nitrogen source [2-5]. Amongst the atrazine mineralization pathways that have evolved, the one from *Pseudomonas* sp. strain ADP has been studied most extensively [6-8]. This pathway mineralizes atrazine *via* the step-wise removal of the side chains (chlorine, ethylamine and isopropylamine) by the sequential action of three hydrolases (AtzA, AtzB and AtzC, respectively) followed by the hydrolytic decomposition of the ring by a further three hydrolases (AtzD, AtzE and AtzF) [9,10]. Genetic and biochemical studies have been conducted with each of these enzymes [11-20], and more recently there has been considerable interest in obtaining their X-ray crystal structures, with the structures of AtzA [21], AtzD [9], AtzF [22] reported. Additionally the structure of a functional analogue of AtzA, termed TrzN [23,24], has also been reported.

A structure for AtzC has also been deposited in the PDB (2QT3)[25], albeit neither the structure itself nor an analysis of structure have been published. AtzC is an isopropylammelide isopropylaminohydrolase (E.C. 3.5.99.4), catalyzing the hydrolysis of *N*-isopropylammelide (IPA) to form isopropylamine and cyanuric acid in the third step of atrazine catabolism (Fig. 1) [15]. AtzC is a member of amidohydrolase protein superfamily and is the third enzyme of the atrazine mineralization pathway. It had been demonstrated by biochemical analysis that AtzC contains one Zn^{2+} ion per monomer, which is also seen in 2QT3 [15,25].

AtzC has 30-35% amino acid sequence identity with three enzymes of known function, ammelide aminohydrolase (TrzC, 3.5.3.-; 31% identity), bacterial cytosine deaminase (CodA, E.C. 3.5.4.1; 35% identity), and *N*-acyl-D-amino acid deacylase (D-AAase, 3.5.1.81; 35% identity). TrzC is involved in hydrolysis of the amine substituent of ammelide (2-amino-1,3,5-triazine-4,6-dione, Fig. 1) in the melamine (1,3,5-triazine-2,4,6-triamine) mineralization pathway [26]. CodA is involved in the pyrimidine salvage pathway and

deaminates cytosine to produce uracil and ammonia (Fig. 1) [27]. D-AAases are involved in hydrolysis of N-acyl-D-amino acids to their corresponding D-amino acids and acetate (Fig. 1) [28]. The structures and reaction mechanisms of CodA and D-AAase are well characterized [28,29]. The X-ray structure of AtzC deposited by Fedorov *et al.* [25] was obtained at moderate resolution (2.24 Å) without bound substrate or inhibitor. In this report, we present a 1.84 Å structure of AtzC with an inhibitor in the active site; this has guided substrate docking and mutagenesis experiments that have allowed us to propose a plausible reaction mechanism.

METHODS AND MATERIALS

DNA manipulation

The unmodified *atzC* gene from *Pseudomonas* sp. strain ADP (accession number: U66917.2; [10]) was obtained from GenScript (New Jersey, USA) as a pUC57 derivative. The *atzC* gene was then subcloned into pETCC2, an expression vector derived from pET14b that is described elsewhere [9] using *NdeI* and *BamHI* restriction sites that had been included in the synthetic gene, but which did not affect the coding sequence. Mutagenesis was carried out using the method of overlap PCR of Ho *et al.* [30]. The list of primers used in this study is given in Table 1. The reaction conditions for PCR were 1× Phusion HF Buffer, 2 picograms pETCC2-*atzC* DNA template, 0.5 μM primers, 200 μM dNTPs and a half unit of Phusion DNA polymerase in a total reaction volume of 25 μL. The cycle conditions for the PCR reaction were a 30 s denaturation step at 98°C followed by annealing at 53°C for 20 s and then extension for 120 s at 72°C for 30 cycles.

The size confirmation and separation of the resulting amplicon was done on 0.6% agarose gel by electrophoresis. The DNA band was excised from the gel and purified using a NucleoSpin Gel and a PCR Clean-up kit (Macherey-Nagel). The amplicons were digested with *BamHI* and *NdeI* restriction enzymes as per manufacturer's instructions and purified by gel extraction as described above. The digested amplicon was cloned into the *BamHI* and *NdeI* restriction sites of pETCC2 using T4 DNA ligase as per the manufacturer's instructions. The DNA polymerase, dNTPs, T4 ligase and restriction enzymes were purchased from New England Biolabs (Ipswich, USA).

Protein expression and purification

The *atzC* gene or its modified derivatives in pETCC2 were used to transform electrocompetent *Escherichia coli* BL21 (λ DE3) cells (Invitrogen). The cells were grown at 310 K on Luria–Bertani (LB) agar plates [31] supplemented with 100 $\mu\text{g.mL}^{-1}$ ampicillin. A single colony was picked from these plates and was used to inoculate a 50 mL starter culture using LB[31] as growth media supplemented with 100 $\mu\text{g.mL}^{-1}$ ampicillin. This starter culture was kept at 310 K overnight whilst shaking at 200 rev.min^{-1} and the next morning was subsequently diluted to 1:20 into LB, Miller medium [31] supplemented with 100 $\mu\text{g.mL}^{-1}$ ampicillin. The diluted culture was shaken at 310 K and 200 rev.min^{-1} and upon reaching an OD_{600} of 0.6–0.8, it was induced for protein expression by the adding isopropyl β -D-1-thiogalactopyranoside (IPTG) at 100 μM final concentration. The induced culture was kept at 310 K overnight whilst shaking at 200 rev.min^{-1} .

The cells were harvested from the overnight culture by centrifugation at 4000 g for 10 min in an Avanti J-E centrifuge (Beckman Coulter, Indianapolis, USA) and then resuspended in 15 mL lysis buffer (50 mM TRIS pH 7.5, 100 mM NaCl) per liter of harvested culture. The resulting cell suspension was lysed by passage through an Avestin C3 homogenizer three times at 124 MPa. After cell lysis, the insoluble cellular debris was removed by centrifugation at 21,000 g using an Avanti J-E centrifuge. Wild-type AtzC was purified from the soluble cell-free extract in two steps: His-tag affinity chromatography using an Ni–NTA Superflow cartridge (Qiagen, Maryland, USA) with a gradient of 0–500 mM imidazole in 50 mM TRIS (pH 7.5) and 100 mM NaCl, and size-exclusion chromatography using a 130 mL column packed with Superdex 200 prep-grade resin (GE Healthcare Life Sciences) with a buffer comprised of 50 mM TRIS (pH 7.5), 100 mM NaCl. Twenty column volumes of buffer were used to achieve the gradients used for the Ni–NTA Superflow cartridge. The protein was concentrated in an Amicon Ultra-15 Centrifugal Filter Unit with an Ultracel-30 membrane (Millipore, Carrigtwohill, Ireland) to 10 mg.mL^{-1} and snap-frozen in liquid nitrogen in 100 μL aliquots. The final purity was estimated to be 98% from a Coomassie-stained gel. The mutants were purified by His-tag affinity chromatography using a 1 mL Ni–NTA Superflow cartridge (Qiagen) with a five column wash by 25 mM imidazole in 50 mM HEPES (pH 7.5) followed by eight column volume of second wash with 50 mM

imidazole in the same buffer and finally eluted with six column volume of 250 mM imidazole in the same buffer.

X-ray crystallography and data collection

2QT3 was described as “complexed with zinc”, and crystallized out of 2-methyl-2,4-pentanediol at pH 5.5 at room temperature (Screen JCSG+ condition H5). Based on this, the purified AtzC protein was thawed, supplemented with 1 mM ZnCl₂, and screened for the suitable crystallization conditions using JCSG+ screen at two different temperatures. This gave some indications of crystallization, at 281 K but not at 293 K, and suggested that the protein was over-concentrated. Subsequent trials were set up with the JCSG+ screen diluted to half strength with water, as well as the PACT screen (also at half strength), with zinc-supplemented protein at 10 mg.mL⁻¹. Initial crystallization hits were all from trials containing polyethylene glycols. An optimization screen using the Optisalts additive screen (Qiagen) and a base condition of 0.1 M lithium sulfate, 0.05 M phosphate-citrate pH 4.2, 10% w/v PEG 1000 gave a crystal that resulted in a sub-3 Å diffraction data set.

Crystals of the tagged protein were generally poorly ordered, and subsequent trials focused on protein treated *in situ* with thrombin, to remove the N-terminal His-tag.

100 µL of tagged AtzC protein at 10 mg.mL⁻¹ in 50 mM HEPES pH 7.5, 100 mM NaCl was added to a PCR tube containing 10 µg of lyophilized thrombin, and subsequently set up (without further purification) in JCSG+ and PACT screens. Well diffracting crystals were obtained from condition F9 of JCSG+ (2.4 M sodium malonate pH 7; Suppl. Fig. 1). Malonate grown crystals of thrombin treated protein (using protein at either 5 or 10 mg.mL⁻¹) were used for soaking experiments, where atrazine, ammeline, barbaturic acid, atrazine 2-OH and atrazine desisopropyl 2-OH were dissolved in either water or DMSO and added to the formed crystals. In all cases, the subsequent X-ray diffraction analysis showed malonate in the active site.

Crystals of the mutants His219Ala, His249Ala, Asp188Ala, Asn304Asp and Trp309Phe were grown in the same way as the native protein (*in-situ* thrombin treatment of the protein; using sodium malonate as the crystallant).

Malonate grown wild-type and mutant crystals were cryo-cooled (without the addition of further cryo-protectant) by plunging into liquid nitrogen, taken to the Australian Synchrotron MX-2 beamline and placed in the N₂ stream at 100 K for data collection. 360 one degree images were taken from each crystal to obtain a complete data set. The reflections were indexed using XDS [32] and scaled using Aimless [33]. 2QT3 was used as the model for molecular replacement with the program Phaser [34]. The model was subsequently rebuilt manually using Coot [35] and refined using Refmac [36]. For the wild-type protein crystals the data were 99.3% complete to a resolution of 1.84 Å and the final model had a R_{work} of 19.1% with an R_{free} of 22.9% (see Table 1 for the crystallographic statistics). The mutant structures were obtained using the native structure as the starting model, but otherwise the data were treated the same as the native.

Differential scanning Fluorimetry (DSF)

Thermal melt analyses, using a bank of formulations that varied pH, buffer chemical and salt concentration ("Buffer screen 9") [37] were performed on the (tagged) wild-type, His219Ala and His249Ala proteins and analyzed using the program Meltdown. All three showed clean unfolding curves, and were quite stable (with T_m values of 333.75 K (+/- 0.08), 336.45 K (+/- 0.75) and 330.15 K (+/- 1.6) respectively) in the 50 mM HEPES pH 7.5, 100 mM NaCl standard formulation, but displayed quite different stability profiles across the different formulations tested (data not shown).

Enzymatic activity and pH dependence

AtzC activity was measured at 301 K using ammelide as a substrate. A 100 mM stock solution of ammelide (Sigma Aldrich, MO) was prepared in 100 mM sodium hydroxide and was subsequently diluted to the required concentrations using 200 mM sodium phosphate buffer at pH 7.5 supplemented with 10% v/v glycerol. The reaction was started by adding AtzC to a final concentration of 40.3 nM and the decrease in absorbance at 230 nm was measured using SpectraMax M2 spectrophotometer (Molecular Devices, California, USA). In experiments performed to determine the pH-dependence of AtzC, the ammelide

concentration was 400 μM and a glutamate dehydrogenase (GDH)-coupled assay [38] was used to measure the rate of ammonia release. In these experiments, a pH range from 6.5 to 8.5 was tested. The V_{max} and K_{M} values were obtained by linear regression in GraphPad version 6.01 for windows (GraphPad Software, La Jolla California USA, www.graphpad.com; Suppl. Fig 2).

For testing malonic acid as an inhibitor (Suppl. Fig. 3), a stock solution was prepared of malonic acid in 200 mM phosphate buffer after which the pH was adjusted to 7.5. Dilutions of this malonic acid solution were made in 200 mM phosphate buffer at pH 7.5 prior to testing. For the assay, enzyme was added to the malonate dilutions followed by the addition of the ammelide substrate.

Docking of substrate in active site

Isopropylammelide was manually positioned in the active site of AtzC, guided by the position of malonate found in the malonate-grown AtzC crystals (PDB 4CQB). Geometry optimization using density functional theory (DFT) was then performed on the active site to refine the positions of the substrate, residue side chains, the coordinated water and the Zn^{2+} center. Prior to optimization, His219 was epsilon protonated as it was within hydrogen bonding distance of the aromatic nitrogen atom on the substrate. His249 was delta protonated as the epsilon nitrogen was within hydrogen bonding distance of the metal-coordinated water. Positional constraints were placed upon backbone atoms of all residues and the side chains of Lys65, Asp127 and Asp303 were set to their charged states. The carbonyl groups of the backbone atoms were modelled as simple aldehydes. Calculations were performed using Dmol3 as implemented in Accelrys Materials Studio 7.0. The PBE functional [39] was employed in conjunction with the DND basis set (comparable to the 6-31G* basis set) and basis file 4.4. Defaults for a coarse-grained geometry optimization applied for all other parameters

RESULTS AND DISCUSSION

The structure of AtzC

AtzC is a tetramer, according to PISA [40], our SEC data (native molecular mass of ~210 kDa), and previous reports [16], with the asymmetric unit containing half of the tetrameric structure (Fig. 2). Each AtzC tetramer is a dimer of dimers (Fig. 2C). The interface between monomers that form dimers occludes about 1870 Å² of surface, whereas each of the interfaces between the two dimers of the tetramer are only 545 Å². Each protomer has two lobes (Fig. 3B), with the N- and C-termini coming together to form a small β-barrel (residues 2 to 55, then 358 to 403) with four β-strands each and a crossover strand (51-55) from the N-terminal side hydrogen bonding with the C-terminal sheet. The second and larger domain which contains the active site is a TIM-barrel fold (Fig. 3A).

The malonate molecule (an inhibitor of AtzC, Suppl. Fig. 3) adopts a planar conformation in the structure and packs against Trp309 (less than 3.5 Å away) as well as interacting with Lys65, Gln160, His219 and Asn304 (2.8-2.9 Å, ~3.2 Å, 2.6-2.7 Å and 2.5 to 3.0 Å, respectively; Fig. 2B). Residue Trp309 packs flat against the malonate found in the structure presented here, whereas Trp309 is found at a significantly different orientation when there is no inhibitor in the active site as seen in 2QT3, suggesting that Trp309 moves in the occupied active site to provide interactions that improve substrate binding. The active sites of the independently solved structures show similar architectures, each having three histidine residues coordinated to the zinc ion (His60, His62, His217; 2.1 to 2.3 Å) as well as Asp303 (about 2.4 Å) and a water (2.0 to 2.1 Å) bound to the zinc (Fig. 2B). The bound water is also within 2.7 Å of His249 and Asp303 as well as close to the bound malonate moiety. The 2QT3 structure has slightly longer bond lengths for all of the zinc interactions with the AtzC protein residues and water described above. The metal center is typical of a subtype III amidohydrolase (as defined by Seibert and Raushel [41,42]).

In the inhibitor:enzyme structure presented herein, additional electron density was found less than 3.5 Å from the malonate, above the plane and above the water molecule (about 4 Å) bound to the zinc ion. This density has been modelled as dimethylsulfoxide (DMSO), which was the solvent used for inhibitors and substrates which were used in both soaks and co-crystallization experiments. Oxygen atoms from both carboxylates of malonate are within hydrogen bonding distance of the water bound to zinc (2.4 to 2.8 Å) for both instances in the asymmetric unit.

The AtzC structure presented here is similar to that of the previously deposited AtzC structure (2QT3) with a 0.6 to 0.7 Å rmsd at the C α level, respectively (Fig. 3A and B). There is some divergence in the two structures between residues 37 to 43 and a more significant difference between the loop and helix seen at residues 79 to 97 (up to a 4.3 Å shift in C α atoms in this region). The closest homologs to AtzC in the PDB (beyond 2QT3) are cytosine deaminases from *E. coli* and *K. pneumoniae*. *E. coli* cytosine deaminase (PDB 3R0D) has the greatest structure homology with the present structure and shows about 1.5 Å rmsd over the C α backbone (17 gaps with 354 aligned residues using the SSM algorithm implemented in Coot [35]). *E. coli* cytosine deaminase (CodA) is a mononuclear Fe²⁺-dependent enzyme, albeit it has a high affinity for both zinc and iron [27,43,44]. The active sites of AtzC and CodA are similar with conservation of the three histidine and aspartate ligands, as well as several other residues (AtzC to 3R0D: His249 to His246, Gln160 to Gln156, Phe158 to Phe154, and Trp309 to Trp319, although this latter residue is in a different rotameric conformation in the two structures). However, there are several important residues that are different, including Asn304 (Asp314), His219 (Glu217), Lys65 (Thr66), Tyr82 (Leu81), Ser280 (Val278), Val310 (Tyr320) and there is a hole in the AtzC structure where Leu282 sits in the 3R0D structure, although this is partially occluded by Tyr82.

Exploring the substrate-binding pocket

Although an inhibitor bound AtzC structure was obtained, the structure of AtzC with its substrate or product was elusive, despite concerted efforts. Although malonate is a weak inhibitor of AtzC (Suppl. Fig. 3), the high malonate concentration (2.4 M) used in the crystallization trials likely prevented other small molecules from binding in the active site. Attempts to find different crystallants more amenable to soaking or co-crystallization efforts were unsuccessful. Substrate docking, guided by the position of the inhibitor in the active site, revealed that five hydrogen-bonding interactions are likely to hold the substrate in a position facilitating nucleophilic attack at the isopropylamine-bearing carbon by the metal-bound nucleophile (Fig. 4); Gln160 and Asn304 each donate and receive hydrogen bonds from IPA, Lys65 donates a hydrogen bond to the oxygen receiving a hydrogen bond from Asn304 and His219 donates a hydrogen bond to the aromatic

nitrogen atom in the heterocycle. There is a tryptophan (Trp309, shown in Fig. 2A) that sits directly above IPA in a π -stacking interaction that, during the geometry optimization, tilted approximately 1 Å out of its crystallographic position to accommodate the substrate. Trp309 has been shown to be motile in the comparison of the ligand-free AtzC structure and the inhibitor-bound structure, suggesting that this amino acid moves upon substrate binding. The N-H...O hydrogen bond donated by His219 is not perfectly linear and forces the hydrogen atom out of the plane of the ring of the side-chain, which would, in turn, exert an attractive force on the substrate in the direction of the nucleophile. The metal-bound water itself is donating hydrogen bonds to His249 and Asp303, the latter of which is also coordinated to the Zn^{2+} center. This orients the water with its lone pairs directed towards the substrate carbon bearing the isopropyl group.

The roles of substrate-binding pocket residues Lys65, Gln160, Asp188, His219, His249, Asp303, Asn304 and Trp309 were probed *via* mutagenesis of the *atzC* gene (Table 2). Crystal structures of WT, His219Ala, His249Ala and Asn304Asp have been deposited in the PDB (PDB: 4CQB, 4CQC, 4CQD and 5AKQ, respectively). Converting His249 to alanine resulted in a complete loss of activity, although the DSF results confirmed that the protein was well-folded. This suggests that His249 plays a pivotal role in the reaction mechanism of AtzC, perhaps in the generation of the metal-bound hydroxide as the nucleophile as an initial step. The pH optimum of AtzC is close to pH 7.25 (Suppl. Fig. 2), consistent with the suggested role of His249 in catalysis.

Substituting Lys65 to alanine gave a 30-fold increase in k_{cat} . The Lys65Ala variant has one fewer hydrogen bonding interactions with the substrate than the wild-type, and it is possible that this may increase k_{cat} by increasing the K_D of AtzC for the product, thereby facilitating more rapid product egress. Similarly for the Lys65Arg substitution (12-fold increase in k_{cat}), the product will be less-ideally bound due to the greater steric bulk of the arginine side chain, which would disrupt hydrogen bonding with the adjacent Asn304. Other mutations that reduced the number of hydrogen bonding interactions with the ring product are Asn304Asp (7-fold increase in k_{cat}) and Asn304Ala (slight increase in k_{cat}). Similarly, Trp309Ala opens up the area above the ring and increased the catalytic rate by almost a factor of two, which may further support the hypothesis that product departure is a rate limiting step. Trp309Phe gave a very slight reduction in catalytic rate.

Gln160 is involved in interactions with the substrate oxygen atom that, we propose, bears the negative charge during hydrolysis. When this residue was mutated to glutamate there was a reduction in k_{cat} of ~60 %, which is likely due to the negative charge on the carboxylate impeding the development of the negative charge on the oxygen atom of the substrate which would otherwise facilitate catalysis. When the same residue was mutated to alanine, there was almost no change in k_{cat} , which suggests that although it is positioned appropriately, the NH_2 of the Gln160 side chain is not crucial for stabilizing the developing negative charge on the intermediate. Asp188 (not shown in Fig. 4) lies within hydrogen bonding distance of His219 and its conversion to alanine increased the k_{cat} 4-fold, however conversion to asparagine reduced k_{cat} by nearly half. An asparagine in position 188 would most likely act as a hydrogen bond acceptor from both His219 and Arg84 (the latter not shown in Fig. 4) given the proximity of one of the oxygen atoms on Asp188 to both residues. The inability of the asparagine oxygen atom to donate a hydrogen bond to His219 necessitates the latter to be doubly protonated in order to donate a hydrogen bond to the substrate as described above and exert the downwards force assisting nucleophilic attack by the coordinated water. This would not be the case if residue 188 was either alanine or aspartate, and this is reflected in the higher k_{cat} values observed for those cases. Furthermore, conversion of His219 to alanine reduced the k_{cat} by nearly the same amount as Asp188Asn, corroborating the geometries provided by the DFT calculations and our interpretation regarding the role of His219.

Asn304Asp provided a significant increase in k_{cat} (7-fold) but a corresponding increase in K_{M} to 25 mM (WT = 3.1 mM). This suggests that while this mutation provides a weaker product-binding site, it also significantly increases the proportion of non-productive substrate binding events. This is most likely due to the new unfavorable electrostatic interaction with the substrate oxygen. A transient repositioning of the nearby Lys65 to form an alternative salt bridge to the one between it and Asp127, sterically hindering substrate binding, is also possible. However, it should be noted that while the carbon atoms of the side chain of Lys65 are in a slightly different conformation in the crystal structure of the Asn304Asp mutant, the NH_3^+ group is occupying the same position.

The AtzC reaction mechanism

The mechanism proposed herein (Fig. 5A), based on the X-ray crystal structure, docking, biochemical and mutagenesis studies, appears straightforward. Conjugation on one side of the ring facilitates a shift of electron density onto the substrate oxygen atom receiving a hydrogen bond from the Gln160 side chain allowing nucleophilic attack to occur. A level of ambiguity arises in the first step, where the deprotonation of the metal-bound water could either occur concurrently with nucleophilic attack on the substrate, or occur in a two-step mechanism analogous to that of carbonic anhydrase [45](Fig. 5B), whereby in the latter, a stable, metal-bound hydroxide is created by His249, which acts as a proton shuttle, removing the proton and generating the active configuration (as shown in Fig. 5A). Both mechanisms are plausible from the experimental and simulation data at hand, however the results of the mutagenesis favor the 'carbonic anhydrase' model. Notably this mechanism is fairly common (although far from universal) among the subtype III amidohydrolases [41], suggesting a predisposition for adopting the mechanism.

Author Contributions

Conceived and designed the experiments: TSP JN SB ACW CS. Performed the experiments: SB TSP JN ACW LJB. Analyzed the data: TSP JN SB ACW CS. Wrote the paper: TSP JN ACW SB CS.

References

1. Hayes TB, Collins A, Lee M, Mendoza M, Noriega N, et al. (2002) Hermaphroditic, demasculinized frogs after exposure to the herbicide atrazine at low ecologically relevant doses. *Proceedings of the National Academy of Sciences of the United States of America* 99: 5476-5480.
2. 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.
3. Mandelbaum RT, Wackett LP, Allan DL (1993) Mineralization of the s-triazine ring of atrazine by stable bacterial mixed cultures. *Applied and Environmental Microbiology* 59: 1695-1701.
4. Vargha M, Takats Z, Marialigeti K (2005) Degradation of atrazine in a laboratory scale model system with Danube river sediment. *Water Research* 39: 1560-1568.
5. Singh P, Suri CR, Cameotra SS (2004) Isolation of a member of *Acinetobacter* species involved in atrazine degradation. *Biochemical and Biophysical Research Communications* 317: 697-702.
6. Strong LC, Rosendahl C, Johnson G, Sadowsky MJ, Wackett LP (2002) *Arthrobacter aurescens* TC1 metabolizes diverse s-triazine ring compounds. *Applied and Environmental Microbiology* 68: 5973-5980.
7. Balotra S, Newman J, Cowieson NP, French NG, Campbell PM, et al. (2014) X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. *Applied and Environmental Microbiology*.
8. Udikovic-Kolic N, Scott C, Martin-Laurent F (2012) Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology* 96: 1175-1189.
9. Peat TS, Balotra S, Wilding M, French NG, Briggs LJ, et al. (2013) Cyanuric acid hydrolase: evolutionary innovation by structural concatenation. *Molecular Microbiology* 88: 1149-1163.
10. 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.
11. de Souza ML, Wackett LP, Sadowsky MJ (1998) The *atzABC* genes encoding atrazine catabolism are located on a self-transmissible plasmid in *Pseudomonas* sp. strain ADP. *Applied and Environmental Microbiology* 64: 2323-2326.
12. de Souza ML, Sadowsky MJ, Wackett LP (1996) Atrazine chlorohydrolase from *Pseudomonas* sp strain ADP: Gene sequence, enzyme purification, and protein characterization. *Journal of Bacteriology* 178: 4894-4900.
13. Seffernick JL, Aleem A, Osborne JP, Johnson G, Sadowsky MJ, et al. (2007) Hydroxyatrazine *N*-ethylaminohydrolase (AtzB): an amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *Journal of Bacteriology* 189: 6989-6997.
14. Boundy-Mills 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. *Applied and Environmental Microbiology* 63: 916-923.

15. 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. *Journal of Bacteriology* 184: 5376-5384.
16. 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. *Journal of Bacteriology* 180: 152-158.
17. Seffernick JL, Erickson JS, Cameron SM, Cho S, Dodge AG, et al. (2012) Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. *Journal of Bacteriology* 194: 4579-4588.
18. 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.
19. 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.
20. 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.
21. Peat TS, Newman J, Balotra S, Lucent D, Warden AC, et al. (2015) The structure of the hexameric atrazine chlorohydrolase, AtzA *Acta Crystallographica Section D-Biological Crystallography* 71: 710-720.
22. Balotra S, Newman J, Cowieson NP, French NG, Campbell PM, 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.
23. Jackson CJ, Coppin CW, Carr PD, Aleksandrov A, Wilding M, et al. (2014) 300-Fold increase in production of the Zn²⁺-dependent dechlorinase TrzN in soluble form via apoenzyme stabilization. *Applied and Environmental Microbiology* 80: 4003-4011.
24. Seffernick JL, Reynolds E, Fedorov AA, Fedorov E, Almo SC, et al. (2010) X-ray structure and mutational analysis of the atrazine chlorohydrolase TrzN. *Journal of Biological Chemistry* 285: 30606-30614.
25. Fedorov AA, Fedorov, E.V., Seffernick, J., Wackett, L.P., Burley, S.K., Almo, S.C. (2007) Crystal structure of *N*-Isopropylammelide isopropylaminohydrolase AtzC from *Pseudomonas* sp. strain ADP complexed with Zn.
26. Eaton RW, Karns JS (1991) Cloning and comparison of the DNA encoding ammelide aminohydrolase and cyanuric acid amidohydrolase from 3 *s*-triazine-degrading bacterial strains. *Journal of Bacteriology* 173: 1363-1366.
27. Ireton GC, Black ME, Stoddard BL (2003) The 1.14 angstrom crystal structure of yeast cytosine deaminase: Evolution of nucleotide salvage enzymes and implications for genetic chemotherapy. *Structure* 11: 961-972.
28. Liaw SH, Chen SJ, Ko TP, Hsu CS, Chen CJ, et al. (2003) Crystal structure of D-aminoacylase from *Alcaligenes faecalis* DA1 - A novel subset of amidohydrolases and insights into the enzyme mechanism. *Journal of Biological Chemistry* 278: 4957-4962.
29. Ko TP, Lin JJ, Hu CY, Hsu YH, Wang AHJ, et al. (2003) Crystal structure of yeast cytosine deaminase - Insights into enzyme mechanism and evolution. *Journal of Biological Chemistry* 278: 19111-19117.

30. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR (1989) Site-directed mutagenesis by overlap extension using the polymerase chain-reaction. *Gene* 77: 51-59.
31. Miller JH (1972) *Experiments in molecular genetics*. New York: Cold Spring Harbor Laboratory.
32. Kabsch W (2010) XDS. *Acta Crystallographica Section D-Biological Crystallography* 66: 125-132.
33. Evans PR (2011) An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallographica Section D-Biological Crystallography* 67: 282-292.
34. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. *Journal of Applied Crystallography* 40: 658-674.
35. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallographica Section D-Biological Crystallography* 66: 486-501.
36. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D-Biological Crystallography* 67: 355-367.
37. Seabrook SA, Newman J (2013) High-throughput thermal scanning for protein stability: making a good technique more robust. *ACS Combinatorial Science* 15: 387-392.
38. Balotra S, Newman J, French NG, Briggs LJ, Peat TS, et al. (2014) Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp. strain ADP. *Acta crystallographica Section F, Structural biology communications* 70: 310-315.
39. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. *Physical Review Letters* 77: 3865-3868.
40. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* 372: 774-797.
41. Seibert CM, Raushel FM (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* 44: 6383-6391.
42. Sugrue E, Fraser NJ, Hopkins DH, Carr PD, Khurana JL, et al. (2015) Evolutionary expansion of the amidohydrolase superfamily in bacteria in response to synthetic compounds: the molinate and diuron hydrolases. *Applied and Environmental Microbiology*.
43. Porter DJT, Austin EA (1993) Cytosine deaminase - the roles of divalent metal ions in catalysis. *Journal of Biological Chemistry* 268: 24005-24011.
44. Hall RS, Fedorov AA, Xu CF, Fedorov EV, Almo SC, et al. (2011) Three-Dimensional Structure and Catalytic Mechanism of Cytosine Deaminase. *Biochemistry* 50: 5077-5085.
45. Mikulski RL, Silverman DN (2010) Proton transfer in catalysis and the role of proton shuttles in carbonic anhydrase. *Biochimica et Biophysica Acta-Proteins and Proteomics* 1804: 422-426.
46. Gilmour KM (2010) Perspectives on carbonic anhydrase. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology* 157: 193-197.
47. Hewett-Emmett D, Tashian RE (1996) Functional diversity, conservation, and convergence in the evolution of the alpha-, beta-, and gamma-carbonic anhydrase gene families. *Molecular Phylogenetics and Evolution* 5: 50-77.

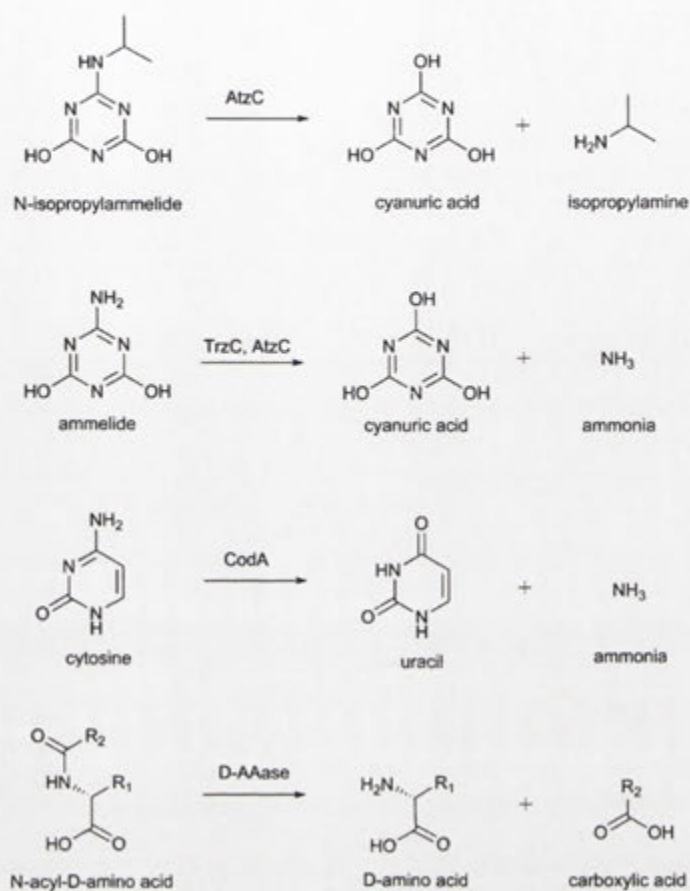


Figure 1. Reactions catalyzed by AtzC and its closest relatives, TrzC, CodA and D-AAase.

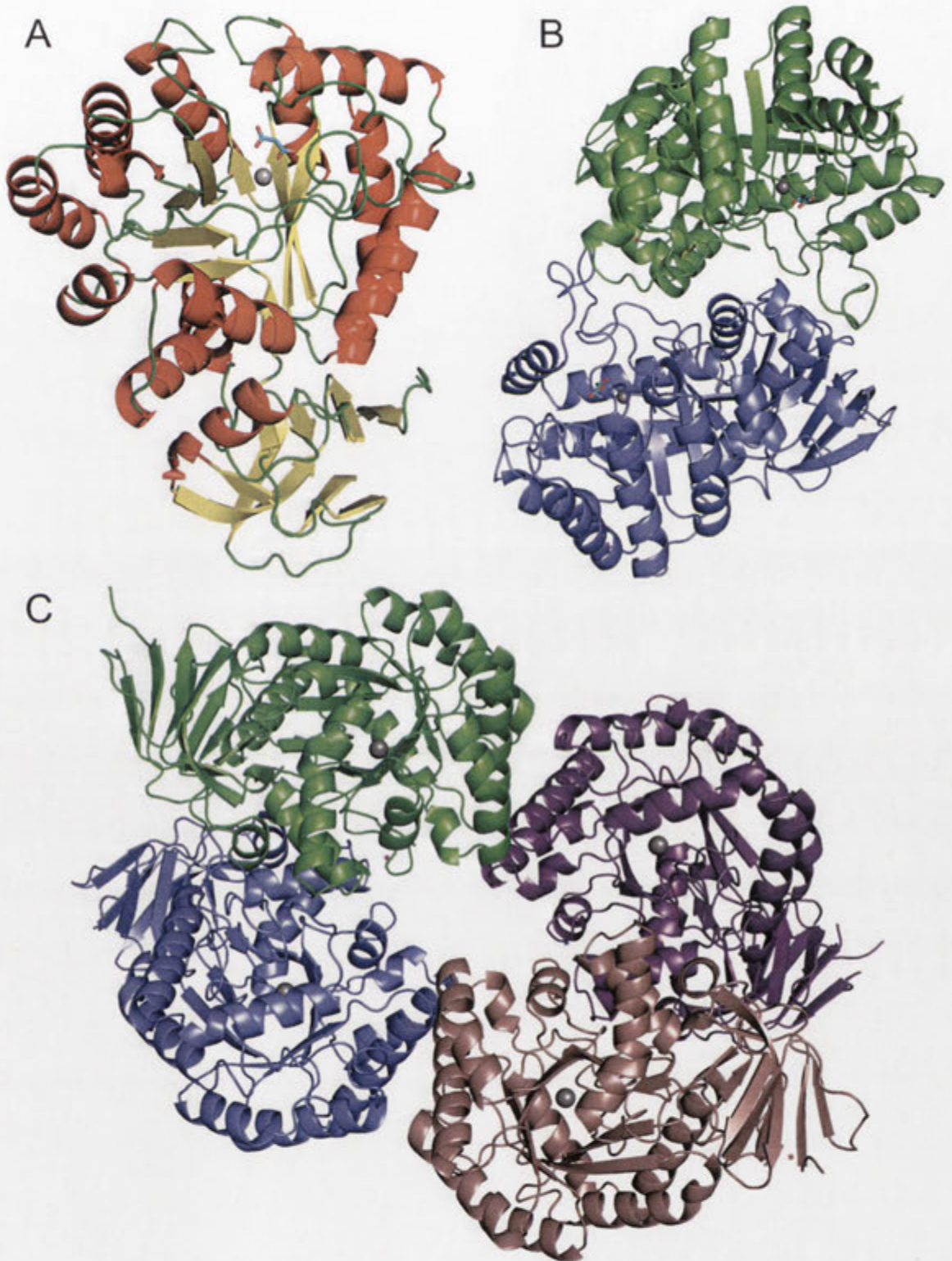


Figure 2. Structure of AtzC. A: A monomer of AtzC with the secondary structure highlighted by color: helices in red, strands in yellow and loops in green, with the Zn ion shown as a grey sphere. B: A dimer of AtzC and C: a tetramer of AtzC showing these interfaces with each monomer in a different color.

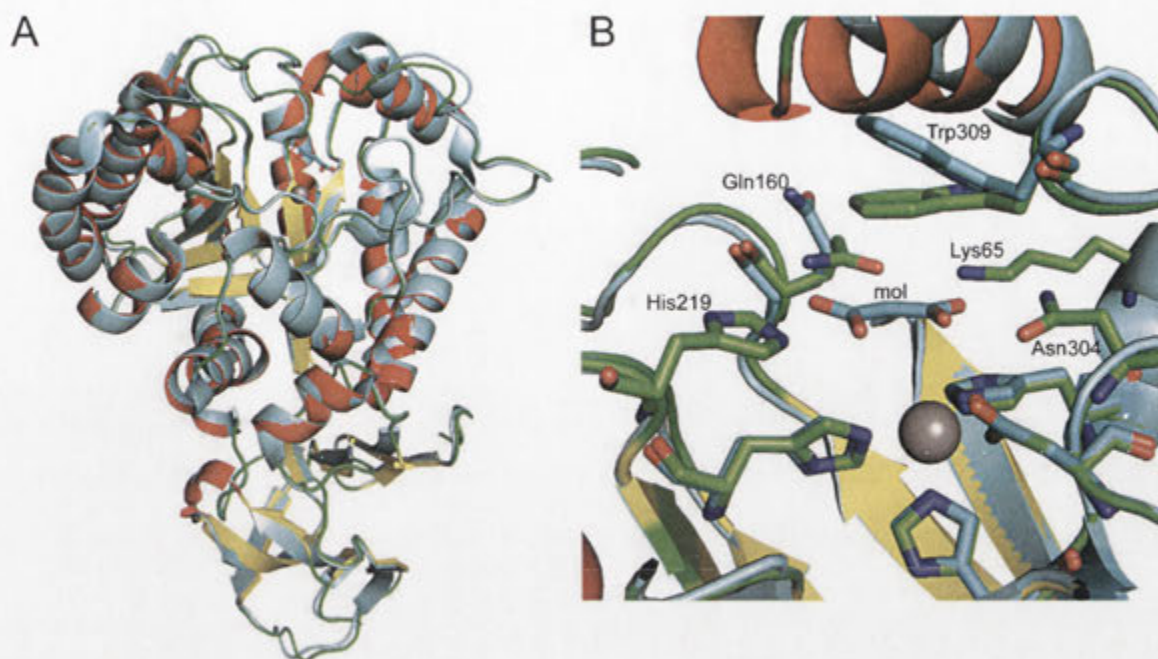


Figure 3. A comparison between the previously reported AtzC structure and the high resolution AtzC structure with bound inhibitor. A: The AtzC is shown as multicolored: helices in red, strands in yellow and loops in green whereas the 2QT3 structure is shown in cyan. B: Residues in the active site of AtzC are shown in stick representation and named along with the malonate (mal) inhibitor. 2QT3 carbons are shown in cyan whereas the structure reported here have green carbons. The Zn ion is shown as a grey sphere coordinated to His60, His62, His217 and Asp303. The Trp309 and Gln160 residues are seen to be in a different orientation between the ligand bound and apo structures.

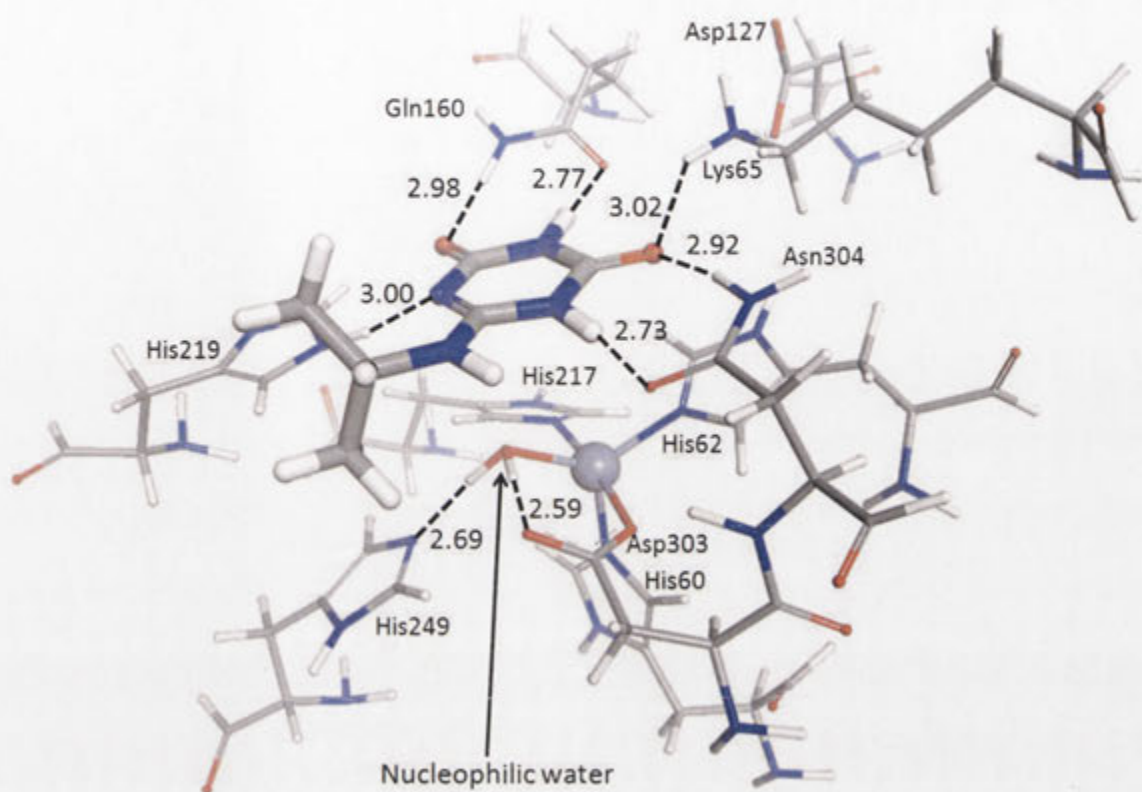


Figure 4. Docking the physiological substrate into the AtzC active site. Isopropylammelide (IPA, drawn with bonds of double thickness) docked into the active site of AtzC. The Zn²⁺ metal center is drawn as a grey sphere. Hydrogen bonding distances (between heteroatoms) are shown in Angstroms with the bonds themselves depicted as dashed lines.

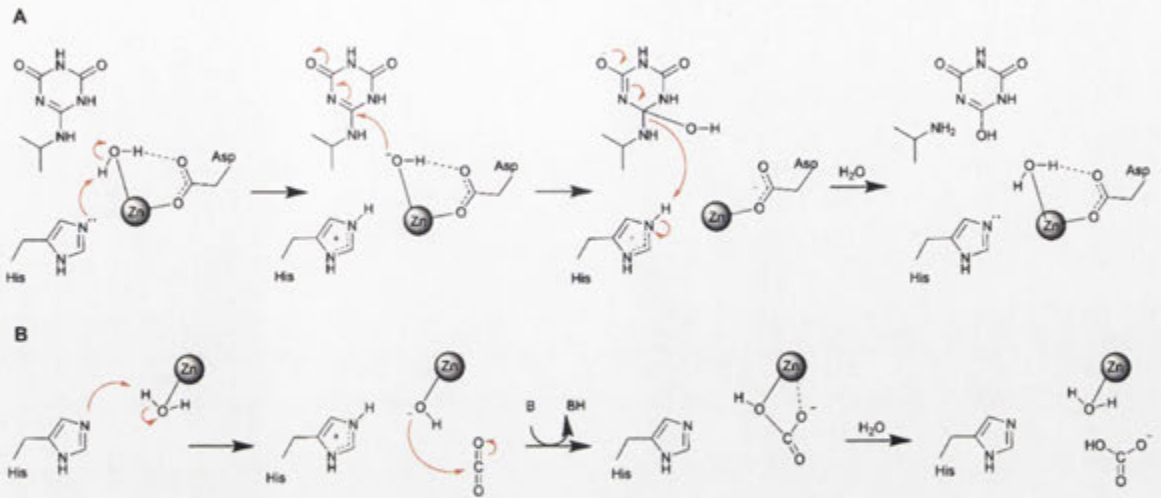


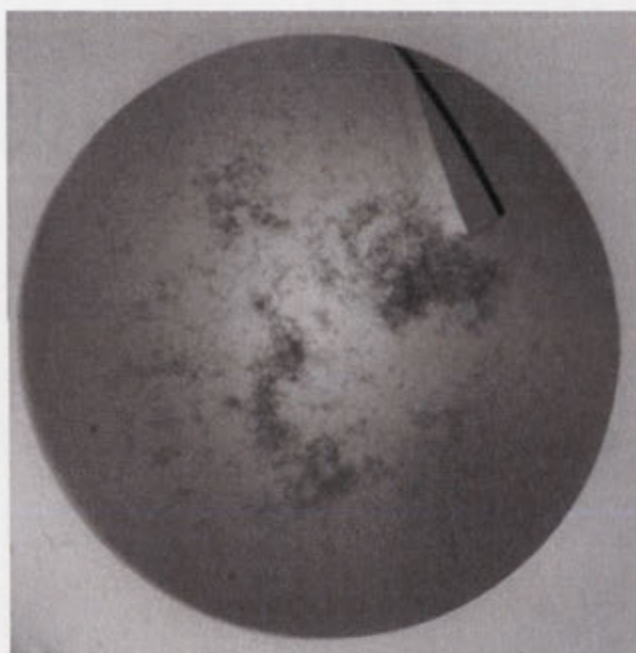
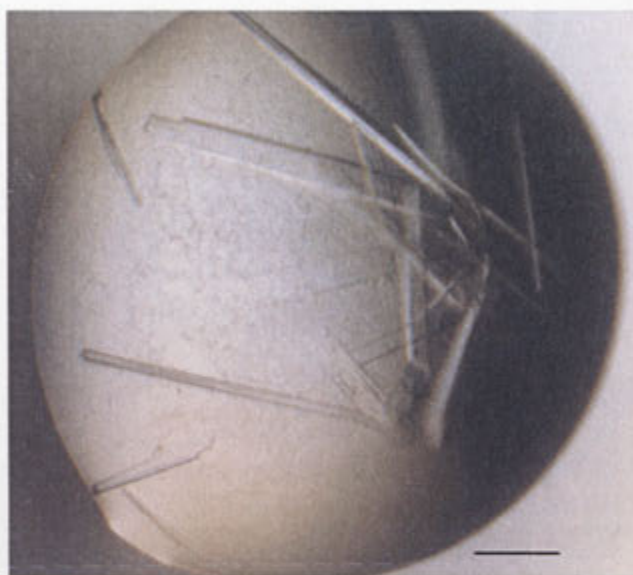
Figure 5. Reaction mechanism of AtzC and comparison with Carbonic anhydrase. The reaction mechanisms for: A) AtzC (proposed here) and B) carbonic anhydrase [45-47] are shown.

Table 1. Crystallographic Data.

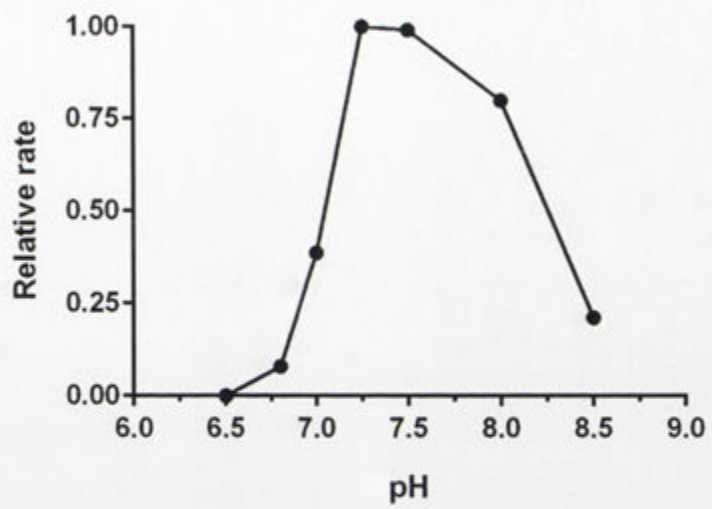
PDB	4CQB (WT)	4CQC (H219A)	4CQD (H249A)	5AKQ (N304D)
Space group	C2	C2	C2	C2
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	106.5, 86.7, 114.2	106.2, 87.3, 114.4	106.3, 87.1, 113.9	106.2, 86.6, 114.1
α , β , γ (°)	90.0, 104.7, 90.0	90.0, 103.9, 90.0	90.0, 104.6, 90.0	90.0, 104.4, 90.0
Resolution (Å)	46.3-1.84 (1.94-1.84)	46.3-2.20 (2.32-2.20)	46.3-2.25 (2.37-2.25)	46.2-2.60 (2.72-2.60)
<i>R</i> _{merge}	0.136 (0.703)	0.136 (0.799)	0.189 (0.954)	0.256 (1.044)
<i>R</i> _{pim}	0.054 (0.285)	0.052 (0.306)	0.078 (0.391)	0.155 (0.632)
<i>CC1/2</i>	0.996 (0.820)	0.997 (0.852)	0.991 (0.699)	0.986 (0.663)
<i>I</i> / σ <i>I</i>	10.7 (2.7)	12.5 (2.8)	9.0 (3.0)	6.9 (2.1)
Completeness (%)	99.3 (95.1)	99.4 (98.9)	99.3 (98.6)	100 (100)
Redundancy	7.4 (6.8)	7.7 (7.7)	6.7 (6.8)	7.3 (7.4)
Refinement				
Resolution (Å)	40.4 - 1.84	46.3 - 2.20	46.3 - 2.25	46.2 - 2.60
Unique reflections	81,509	48,598	45,047	29,424
<i>R</i> _{work} / <i>R</i> _{free} (%)	19.1 / 22.9	19.6 / 22.3	16.6 / 19.4	20.5 / 24.1
No. atoms	6,923	6,544	6,717	6,314
Protein	6,392	6,362	6,448	6,276
Zn ions / other	2 and 39	2 and 32	2 and 40	2 and 2
Water	490	186	227	34
<i>B</i> -factors (Å ²)	19.4	32.2	27.2	36.0
Protein	19.0	32.1	27.1	36.2
Zn ions / other	16.8 / 24.8	46.9 / 49.2	26.1 / 42.7	33.1 / 40.0
Water	24.3	28.6	25.8	21.0
R.m.s. deviations				
Bond lengths (Å)	0.019	0.010	0.008	0.019
Bond angles (°)	1.777	1.269	1.180	1.772

Table 2. AtzC variant characterization. NDA = No detectable activity.

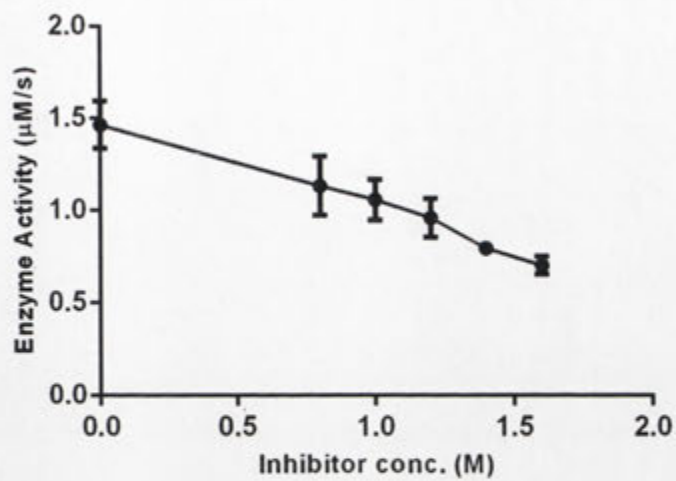
Variant	k_{cat} (s^{-1})	K_{M} (mM)	$K_{\text{cat}}/K_{\text{M}}$ ($\text{s}^{-1} \cdot \text{M}^{-1}$)
Wild-type	26.2 ± 1.8	3.1 ± 0.2	8.5 × 10 ³
Lys65Ala	778.5 ± 79.7	6.1 ± 0.6	1.3 × 10 ⁵
Lys65Arg	332.3 ± 34.3	2.4 ± 0.3	1.4 × 10 ⁵
Gln160Ala	23.0 ± 3.5	1.8 ± 0.3	1.3 × 10 ⁴
Gln160Glu	10.4 ± 0.5	1.3 ± 0.1	8.1 × 10 ³
Asp188Ala	106.5 ± 11.4	2.2 ± 0.2	4.9 × 10 ⁴
Asp188Aln	14.1 ± 3.1	1.9 ± 0.5	7.4 × 10 ³
His219Ala	16.8 ± 4	1.6 ± 0.4	1.1 × 10 ⁴
His249Ala	NDA	NDA	NDA
Asp303Ala	35.8 ± 3.5	1.7 ± 0.2	2.1 × 10 ⁴
Asp303Aln	90.4 ± 8.9	6.1 ± 0.6	1.5 × 10 ⁴
Aln304Ala	29.3 ± 1.7	4.2 ± 0.2	7.0 × 10 ³
Aln304Asp	189.0 ± 34	25.0 ± 4.5	7.6 × 10 ³
Trp309Ala	42.4 ± 4	5.6 ± 0.7	7.6 × 10 ³
Trp309Phe	21.3 ± 2	3.0 ± 0.3	7.1 × 10 ³



Supplementary Figure 1. Crystals of AtzC grown in a PEG-based buffer (top) and a malonate- based buffer (bottom). The bar is approximately 200 μM in length.



Supplementary Figure 2. pH dependency of AtzC-mediated ammelide deamination.



Supplementary Figure 3. Inhibition of AtzC by malonic acid.

Supplementary Table 1. Oligonucleotides used in this study.

Oligonucleotide name	Oligonucleotide sequence (5'-3')
AtzC D303E fwd	CTTGGCTGTGCTTCGGAAAATATCAGAGATTTTTGG
AtzC D303E rev	CAAAAATCTCTGATATTTCCGAAGCACAGCCAAGA
AtzC D303A fwd	ATCTTGGCTGTGCTTCGGCGAATATCAGAGATTTTTG
AtzC D303A rev	CCAAAATCTCTGATATTCGCCGAAGCACAGCCAAGA
AtzC Q160N fwd	GTCGTAGCCTTTGCAAATAGTGGATTTTTCGTTG
AtzC Q160N rev	AACGAAAAATCCACTATTTGCAAAGGCTACGACT
AtzC Q160A fwd	AGTCGTAGCCTTTGCAGCGAGTGGATTTTTCGTTG
AtzC Q160A rev	AACGAAAAATCCACTCGCTGCAAAGGCTACGACTT
AtzC K65R fwd	GCACATACCCACATGGATCGTTCATTTACGAGCACAGG
AtzC K65R rev	GTGTATGGGTGTACCTAGCAAGTAAATGCTCGTGTCCA
AtzC K65A fwd	ACATACCCACATGGATGCGTCATTTACGAGCACAG
AtzC K65A rev	TGTGCTCGTAAATGACGCATCCATGTGGGTATGTG
AtzC N304Q fwd	TTGGCTGTGCTTCGGACCAGATCAGAGATTTTTGGG
AtzC N304Q rev	ACCCAAAATCTCTGATCTGGTCCGAAGCACAGCCA
AtzC N304A fwd	GGCTGTGCTTCGGACGCGATCAGAGATTTTTGGG
AtzC N304A rev	CCAAAATCTCTGATCGCGTCCGAAGCACAGCCA
AtzC D188E fwd	GTTGGGGGAGTTGAACCTGCTACGCGGG
AtzC D188E rev	TCCCGCTAGCAGGTTCAACTCCCCCAA
AtzC D188A fwd	TTGGGGGAGTTGCACCTGCTACGCG
AtzC D188A rev	CCCGCTAGCAGGTGCAACTCCCCCA
AtzC W309F fwd	GGACAATATCAGAGATTTTTTTGTTCCCTTTGGCAACGGTG
AtzC W309F rev	TCACCGTTGCCAAAGGGAACAAAAAATCTCTGATATTGTC
AtzC W309A fwd	CAATATCAGAGATTTTGC GGTTCCCTTTGGCAACG
AtzC W309A rev	GTTGCCAAAGGGAACCGCAAAATCTCTGATATTGT
AtzC H219A fwd	TGATATCGACTATCACATAGCTGATATTGGAAGTGTGG
AtzC H219A rev	TCCAACAGTTCCAATATCAGCTATGTGATAGTCGATATC
AtzC H249A fwd	GGTAGAGTAACTACGAGTGCTGCCTGGTGTGTTTGCAG
AtzC H249A rev	TCTGCAAAACACCAGGCAGCACTCGTAGTTACTCTAC

Chapter 5. Cyanuric acid hydrolase: evolutionary innovation by structural concatenation.

Pages 88-115

Overview

Cyanuric acid hydrolase, AtzD, is a member of a newly identified family of ring-opening amidases. Herein, we report the first X-ray structure for any member of this family. AtzD possesses a novel protein fold (termed the 'Toblerone' fold), which is proposed to have evolved by the concatenation of monomers of the trimeric YjgF superfamily and the acquisition of a metal binding site. Threefold rotational symmetry can be observed in the AtzD monomer, its active site and its substrate, to the extent that the active site possesses three potential Ser-Lys catalytic dyads. Based on biochemical evidence and crystallographic data, out of three dyads, a single dyad (Ser85-Lys42) has been hypothesized to be catalytic. A plausible catalytic mechanism based on these observations is also presented. Active-site residues conferring substrate specificity were explored by making a comparison with a homology model of the related barbiturase, Bar, and the phylogeny of the 68 AtzD-like enzymes in the database were analysed in light of this structure-function relationship.

Contributions

Sahil Balotra: DNA cloning, mutagenesis, assay development, enzyme expression and purification and enzyme characterization. I co-authored the paper, contributing to writing and figure preparation.

Drs Janet Newman, Tom Peat and Santosh Panjekar: Crystallography and structural determination

Dr Matthew Wilding: Phylogenetic analysis

Nigel G. French and Lyndall J. Briggs: Gene cloning and enzyme expression optimisation

Dr Nathan Cowieson: SAXS analysis

Dr Colin Scott: Planning and supervision

This paper has been published: Peat *et al.* (2013) *Molecular Microbiology* **88**: 1149-1163.

Cyanuric acid hydrolase: evolutionary innovation by structural concatenation

Thomas S. Peat,¹ Sahil Balotra,² Matthew Wilding,²
Nigel G. French,² Lyndall J. Briggs,²
Santosh Panjikar,^{3,4} Nathan Cowieson,³
Janet Newman¹ and Colin Scott^{2*}

¹CSIRO Materials, Science and Engineering, Parkville,
Vic. 3801, Australia.

²CSIRO Ecosystem Sciences, Black Mountain,
Canberra, ACT 2601, Australia.

³Australian Synchrotron, Clayton, Vic. 3168, Australia.

⁴Department of Biochemistry and Molecular Biology,
Monash University, Vic. 3800, Australia

Summary

The cyanuric acid hydrolase, AtzD, is the founding member of a newly identified family of ring-opening amidases. We report the first X-ray structure for this family, which is a novel fold (termed the 'Toblerone' fold) that likely evolved via the concatenation of monomers of the trimeric YjgF superfamily and the acquisition of a metal binding site. Structures of AtzD with bound substrate (cyanuric acid) and inhibitors (phosphate, barbituric acid and melamine), along with mutagenesis studies, allowed the identification of the active site. The AtzD monomer, active site and substrate all possess threefold rotational symmetry, to the extent that the active site possesses three potential Ser–Lys catalytic dyads. A single catalytic dyad (Ser85–Lys42) is hypothesized, based on biochemical evidence and crystallographic data. A plausible catalytic mechanism based on these observations is also presented. A comparison with a homology model of the related barbiturase, Bar, was used to infer the active-site residues responsible for substrate specificity, and the phylogeny of the 68 AtzD-like enzymes in the database were analysed in light of this structure–function relationship.

Introduction

The deployment of the *s*-triazines, such as atrazine (1-chloro-3-ethylamino-5-isopropylamino-2,4,6-triazine), into the environment has resulted in the evolution of new

atrazine catabolic pathways in bacteria (Wackett, 2009). These pathways, and the enzymes that comprise them, have garnered considerable attention in light of their potential application in the bioremediation of atrazine (Udikovic-Kolic *et al.*, 2012), which is a potential human carcinogen, endocrine disrupter and teratogen (Wiegand *et al.*, 2001; MacLennan *et al.*, 2002; Hayes *et al.*, 2003). *s*-Triazine degrading systems are also model systems for the study of evolving enzymes and pathways (Wackett, 2009; Noor *et al.*, 2012; Udikovic-Kolic *et al.*, 2012).

Atrazine is mineralized in bacteria by a series of hydrolases, with the first three hydrolytic steps resulting in the dechlorination and deamination of atrazine to produce cyanuric acid (CA; 2,4,6-trihydroxy-1,3,5 triazine; Seffernick *et al.*, 2012), which is then hydrolysed by the ring-opening amide hydrolase, cyanuric acid hydrolase (AtzD; E.C. 3.5.2.15). AtzD catalyses the hydrolytic ring opening of CA to yield 1-carboxybiuret which then spontaneously decomposes to biuret and carbon dioxide (Fig. 1A; Seffernick *et al.*, 2012). Although CA is a naturally occurring compound, occasionally formed during oxidative damage of DNA (Wackett, 2009), its environmental abundance has increased markedly since the introduction of synthetic *s*-triazine compounds.

AtzD, is the archetype for a family of ring-opening amidases (Seffernick *et al.*, 2012) that includes a related CA hydrolase (TrzD; 56% identical to AtzD; Karns, 1999) and barbiturase (Bar; E.C. 3.5.2.1; Soong *et al.*, 2002; Seffernick *et al.*, 2012). Barbituric acid (2,4,6-trihydroxy-1,3-pyrimidine; BA), is an intermediate in the oxidative degradation of pyrimidines (Soong *et al.*, 2002). Pyrimidines are degraded reductively in eukaryotes and most prokaryotes (Soong *et al.*, 2001); however, an oxidative pathway was described from a small number of eubacteria in the 1950s in which pyrimidines are first oxidized to BA, which is then hydrolysed by Bar and ureidomalonase to yield malonic acid and urea (Soong *et al.*, 2001). This metabolic pathway is rare, as AtzD/Bar homologues were described in only 3% of 6423 surveyed genomes in a recent study (Seffernick *et al.*, 2012).

Bar catalyses an equivalent ring-opening reaction to that of AtzD, using BA as substrate and yielding the stable product 3-oxo-3-ureidopropanoate (Soong *et al.*, 2002; Fig. 1B). BA differs from CA at just a single position in the ring (C5 in BA and N5 in CA; Fig. 1A and B). It is therefore

Accepted 28 April, 2013. *For correspondence. E-mail colin.scott@csiro.au; Tel. (+44) 2 6246 4090; Fax (+44) 2 6246 4173.

© 2013 The Authors. Molecular Microbiology published by John Wiley & Sons Ltd.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

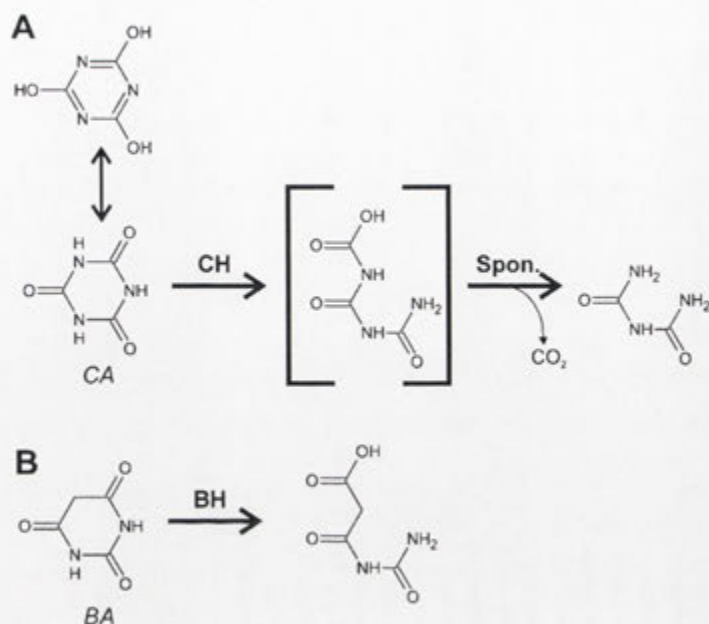


Fig. 1. Enzymatic hydrolysis of cyanuric acid and barbituric acid.

A. Cyanuric acid (CA) tautomerizes between its trihydro- and triketo-forms, the latter acting as the substrate for cyanuric acid hydrolase (CH). The product of hydrolysis is 1-carboxybiuret, which spontaneously decomposes to biuret (liberating CO₂). **B.** Barbituric acid (BA) is hydrolysed by barbiturate hydrolase (BH) to form 3-oxo-3-uridopropionate.

perhaps surprising that BA is an inhibitor, rather than a substrate, for AtzD (Seffernick *et al.*, 2012) and that CA is an inhibitor of Bar (Soong *et al.*, 2002). With no available structure for AtzD or Bar, the determinants of specificity were not identified.

Early studies of this family suggested that these enzymes were members of the metal-dependent amidohydrolase family; Soong and co-workers were able to inactivate Bar by treatment with a chelator and the activity was restored upon incubation with Zn²⁺ (Soong *et al.*, 2002). However, more recent work from Seffernick *et al.* (2012) demonstrated that this family is unrelated to amidohydrolases, forming a discrete evolutionary class of their own. Moreover, it was suggested that the newly defined family was metal-independent.

Herein, we describe the first X-ray structure for this family, an unusual concatenated knot fold (which we term the 'Toblerone' fold), that possesses a threefold rotational symmetry that extends to the active-site architecture. Structures of AtzD with the substrate and inhibitors, and kinetic and mutagenesis analysis of these with wild-type and mutant proteins are presented. We identify Lys42 and Ser85 as a plausible catalytic dyad and hypothesize a catalytic mechanism consistent with that assignment.

Results

Structure of AtzD

The AtzD monomer is a unique arrangement of a core of 12 β -strands with six helices on the outside (Fig. 2A). This structure is formed from three 'repeating units' (RUs) that

share the same structure as one another; they have been designated RU A (residues 3–104), B (residues 113–250) and C (residues 256–364). Each RU is comprised of a four strand antiparallel β -sheet and two helices. This arrangement of RUs confers threefold rotational symmetry in the AtzD monomer reminiscent of a 'Toblerone' Bar (viewed end on; Fig. 2A), and it is proposed that this new fold be termed the Toblerone fold. The three RUs of the monomer align with the following root mean square deviation (rmsd) over the C α atoms and sequence identities: B to A = 3.0 Å/9.3%, C to A = 2.3 Å/16.2%, C to B = 2.2 Å/17.8% (Fig. 2B).

There were no homologues to the full-length AtzD in the PDB (using SSM). However, a single RU has homology with proteins of the YgjF superfamily: perchloric acid-soluble protein (PSP) from *Pseudomonas syringae* (PDB: 3K0T), YGJF from *Streptococcus pyogenes* (PDB: 3EWC) and RutC from *Escherichia coli* (PDB: 3V4D). PSP overlays an AtzD RU with a rmsd of ~1.3 Å over 109 residues (sequence identity ~16.5%; Fig. 2C), notwithstanding that the YgjF proteins have a six strand antiparallel β -sheet (rather than four). YgjF proteins form homotrimers that align poorly with the structure of an AtzD monomer (Fig. 2D), due to the two additional β -strands per monomer/RU.

AtzD possesses a single metal binding site per monomer, in RU C, for which there is no equivalent in the YgjF family of proteins (Fig. 3A–C; Zhang *et al.*, 2010; Knapik *et al.*, 2012), but which is completely conserved in 90% of the AtzD homologues currently available (data not shown). The metal binding site consists of backbone carbonyls from residues Ala347, Gln350, Pro352, Gly355,

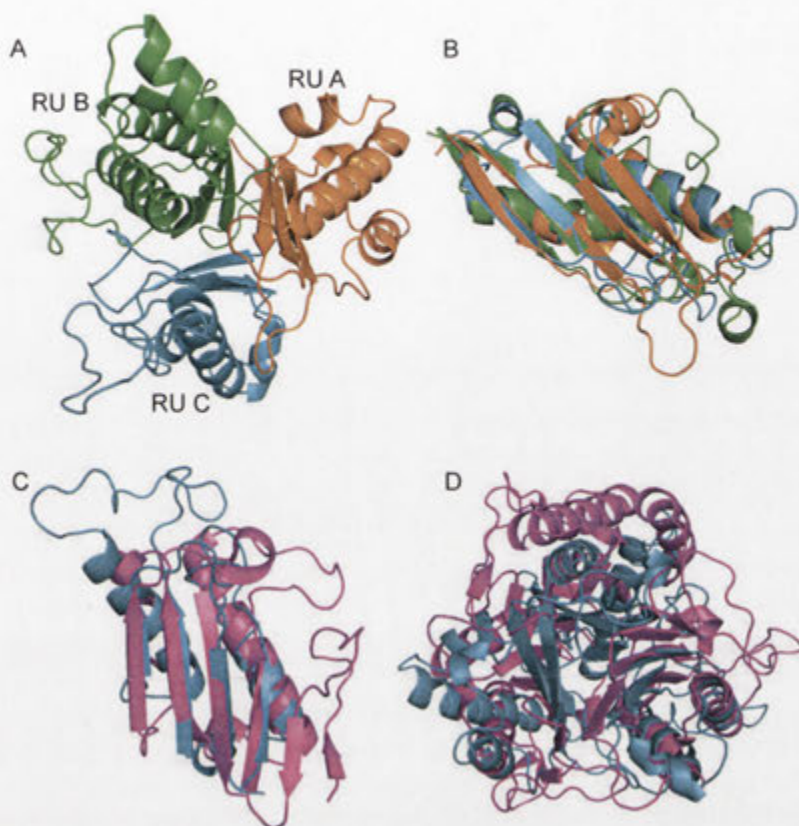


Fig. 2. Structure of AtzD monomer and comparison to the perchloric acid soluble protein (PSP).
 A. Structure and substructure of an AtzD monomer. The RUs are coloured as follows: RU A (residues 2–104) is orange, RU B (residues 113–250) is green and RU C (residues 256–364) is cyan.
 B. Overlay between RUs A, B and C (coloured as Fig. 2A).
 C. PSP (PDB: 3K0T) monomer (magenta) superposed with RU C of AtzD (cyan).
 D. A trimer of PSP (magenta) superposed with a monomer of AtzD (cyan).

the δ -carboxylate of Glu298 and a water molecule. The octahedral co-ordination, electron density and B factor associated with the cation suggest that the metal bound by AtzD is likely to be magnesium (Dudev *et al.*, 1999; Fig. 3D). Anomalous scattering data show that the native metal ion is either magnesium or sodium (Supplemental Fig. S1). Unfortunately, it is not possible to distinguish between Na^+ and Mg^{2+} by anomalous scattering as both cations have an identical number of electrons and the theoretical anomalous difference at accessible X-ray wavelengths is consequently near identical. However, it is certain from these data that the density found is not Zn^{2+} , nor is it any other transition metal. The metal appears to stabilize an extended loop that contains residues that make polar contact with amino acids in the adjacent RU (Fig. 3B and C), which may stabilize the β -strand that holds Lys296 in the active site and contribute to the stability of the enzyme more generally. In addition, this highly ordered structure orients Glu348 in an interaction with Arg195, which sits in the active site (Fig. 3C).

Gel filtration during purification suggested that the native enzyme was a dimer (data not shown); however, Small Angle X-ray Scattering (SAXS; Supplemental Fig. S2) shows that AtzD forms a compact tetramer with D2 symmetry in solution (Fig. 4A and B). The diameter of

the compact tetramer is near identical to that of the dimer, explaining the underestimate in native molecular weight using a lower resolution method (gel filtration) during purification. The core of the D2 tetramer is highly ordered, as indicated by the low B-factors found in the crystal structure (Fig. 4C and D). The surface of the tetrameric protein assembly has two cavities per monomer that could provide access to each active site for substrate ingress/product egress (Fig. 4E). One is formed by residues Tyr188, Met191, Thr321 and Gly345. The second is formed by residues Thr58, Phe82, Glu237 and Val242, and between the two cavities sits Lys162 (hydrogen bonded to the Met84 carbonyl oxygen).

Active site of AtzD

A phosphate ion was found in a polar, positively charged cavity in the AtzD monomer, making several hydrogen bonds with the protein. This cavity contains three lysine/serine dyads (Lys42, Ser85, Lys162, Ser233, Lys296 and Ser344) any, or all, of which could plausibly form the catalytic site (Botos and Wlodawer, 2007; Ekici *et al.*, 2008).

Structures of AtzD were then obtained bound to CA or to melamine (1,3,5-triazine-2,4,6-triamine; Table 1,

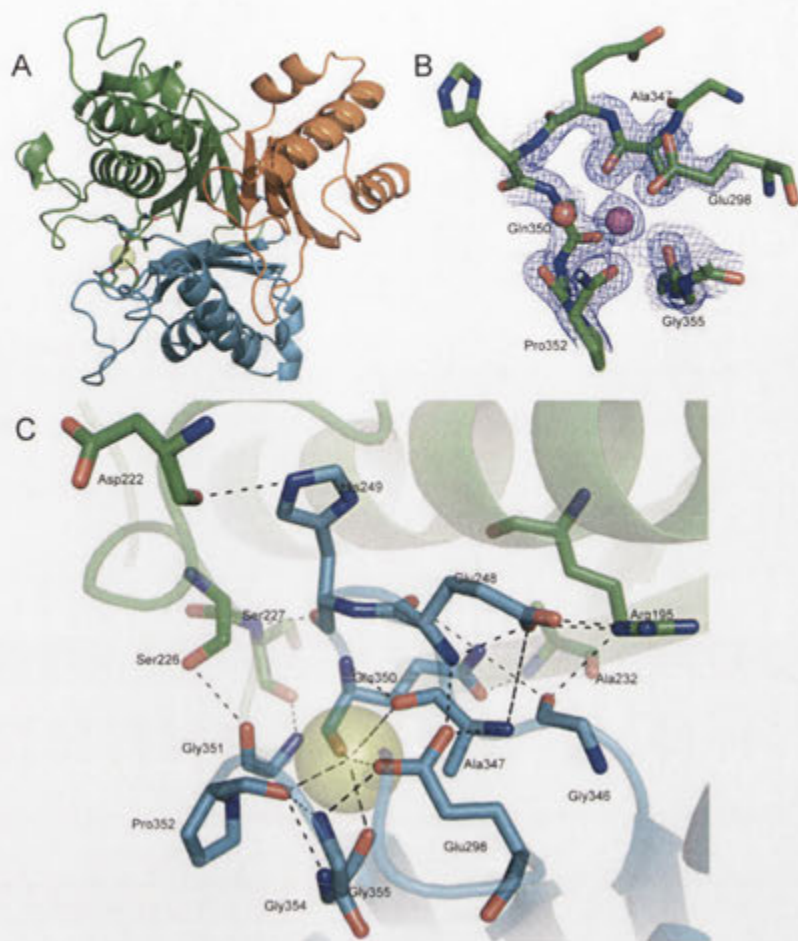


Fig. 3. The metal binding site of AtzD. A. The position of the bound metal (green sphere) in the AtzD monomer (coloured as per Fig. 2A). B. The bound metal (magenta sphere), the metal-binding residues, the water ligand (red sphere) and their electron densities are shown. C. The interactions between the metal, metal-stabilized loop and the surrounding protein.

Fig. 5A–C). There was clear density for melamine and CA, displacing two well-ordered water molecules and the phosphate found in the 'native' structure (Fig. 5D). Melamine and CA bound in an identical fashion (0.47 Å rmsd) via an extensive hydrogen bonding network between CA/melamine and Arg54, Arg195, Arg325, Gly86, Ala234, Gly345, Ser85, Ser233 and Ser344 (Fig. 5B). BA-containing crystals were also obtained, but at lower resolution. The position of BA superposed well with the position of the melamine and CA.

The threefold symmetry of the Toblerone fold is also present in the active site (Fig. 5B and C), with each RU contributing an arginine, a serine, a lysine and a main-chain carbonyl (glycine/alanine) to the active site (Fig. 5B and C). Although all three serine/lysine pairs are found in the vicinity of the substrate, only the side-chain of Ser85 is within strong hydrogen bonding distance of the molecules (i.e. < 3 Å). In addition, we were able to detect significant polarization of the substrate's electron density that suggested that a covalent bond could form between the substrate and Ser85 (Fig. 6).

Identification of catalytic amino acids

Consistent with earlier observations, the optimal pH for the reaction was 8.5–9.0 (Supplemental Fig. S3; Fruchey *et al.*, 2003) and similar to other Lys–Ser containing hydrolases, such as the leader peptidase of *E. coli*, which uses a buried lysine with an apparent pK_a of 8.7 as a general base to activate a nucleophilic serine (Paetzel *et al.*, 1998). Inhibitor studies were also consistent with a Lys–Ser-dependent hydrolytic mechanism, as AtzD was inhibited by phenylmethanesulphonyl fluoride (PMSF) and lysine methylation (Table 2).

Tryptic fingerprinting of AtzD that had been treated with PMSF was used in an attempt to determine which of the active-site serine residues participates in catalysis. In the untreated control, the tryptic fragments of AtzD containing all three active-site serine residues were observed (Supplemental Table S1). When the enzyme was treated with PMSF, the abundance of the Ser85-containing fragment was far lower than in the untreated control, while the other tryptic fragments (including those containing Ser233 and

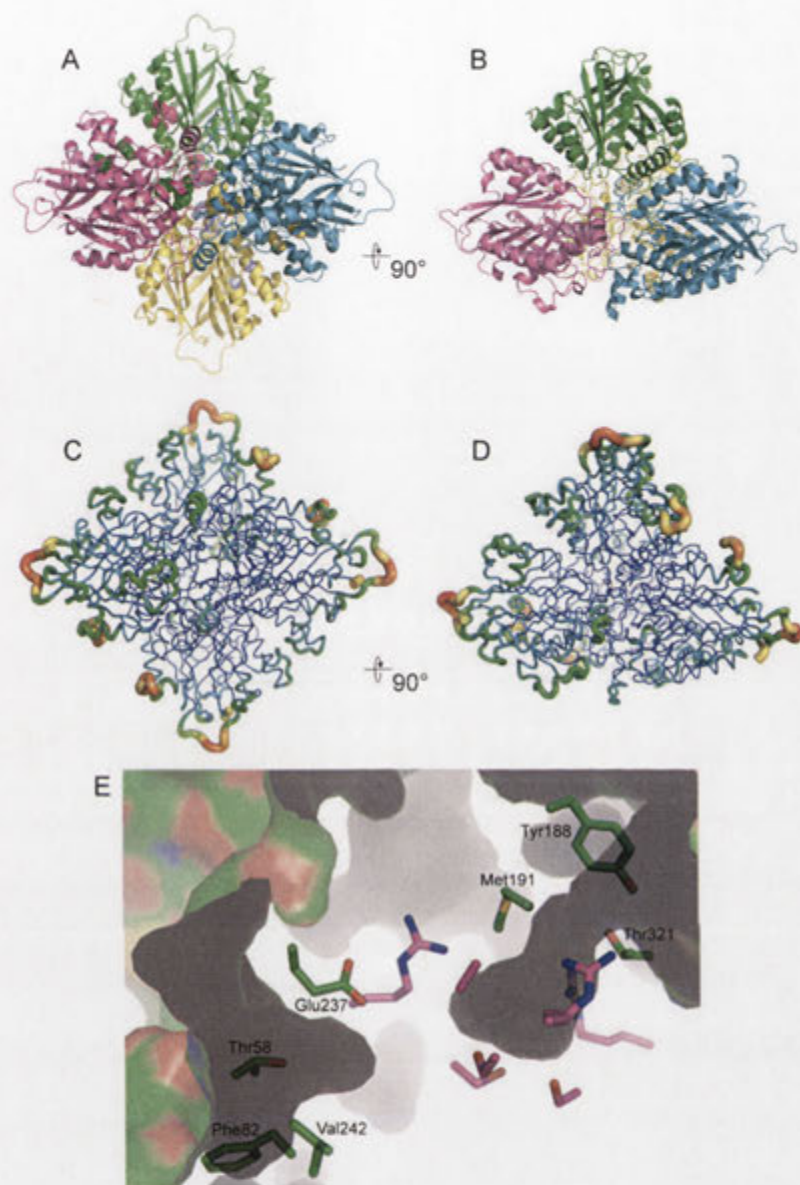


Fig. 4. The AtzD tetramer.

A and B. Ribbon depictions of the tetrameric structure of AtzD and are approximate 90 degree rotations of each other. In (A) residues 58 and 81 (towards the periphery of the tetramer), and 188 and 191 (towards the centre of the tetramer) are shown as balls in colours that contrast with the main chain; these residues contribute to the two potential substrate channels.

C and D. The same rotations, but in this case the chains are coloured by B factors.

E. Two channels potentially provide access to the active site of each AtzD monomer. Active-site amino acid residues are shown in pink, but are otherwise unlabelled. The amino acid side-chains that form the channels are shown in green and their identities given.

Ser344) do not vary substantively from the control (Supplemental Table S1). Although it was not possible to detect masses equivalent to those expected from the covalently modified Ser85-containing fragments, these observations imply that Ser85 may be the active-site serine, consistent with the crystallographic evidence observed (Fig. 6).

Unexpectedly, an unreported substrate inhibition was observed when substrate was supplied in molar excess (Supplemental Fig. S3; Fruchey *et al.*, 2003; Seffernick *et al.*, 2012). As it was not possible to delineate the effects of inhibition from the overall rate of the reaction, it was not possible to derive accurate steady-state estimates of k_{cat} and K_M . Reactions conducted in the presence of

biuret were not inhibited (M. Wilding, unpubl. obs.), suggesting that the inhibition is mediated by either the substrate or 1-carboxybiuret.

AtzD activity was significantly reduced by incubation with *o*-phenanthroline (Table 2), suggesting that the bound metal is required for activity, albeit it appears to be too distant from the substrate-binding pocket to be involved directly in catalysis. Removal of the metal also resulted in a destabilization of the protein (> 80% precipitated at room temperature), which hampered our efforts to obtain a crystal of metal-free AtzD or accurately determine the K_d value for the bound metal. The activity of apoAtzD could be recovered by providing magnesium ($MgCl_2$), but not Zn^{2+} as reported by Soong *et al.* for Bar (Soong *et al.*,

Table 1. Crystallographic statistics.

Data collection						
PDB code	3ZGR	3ZGT	3ZGU	3ZGS		
Ligand	Native	Melamine	Barbituric acid	Cyanuric acid	Xe-derivative	
Spacegroup	R32:H	R32:H	R32:H	R32:H	Spacegroup	R32:H
Cell (a = b × c)	130.7 × 236.9	128.4 × 228.4	129.2 × 233.8	129.1 × 229.7	Cell (a = b × c)	130.7 × 233.1
Resolution (Å)	1.90	2.60	3.09	2.58	Resolution (Å)	2.55
Completeness (%)	99.8 (99.1)	99.9 (99.9)	99.5 (96.5)	100 (100)	Completeness (%)	99.8 (100)
Rmerge %	0.056 (0.573)	0.137 (0.784)	0.145 (0.621)	0.141 (0.799)	Rmerge %	0.101 (0.739)
Mean I/sigI	27.9 (4.4)	12.8 (3.3)	12.9 (3.6)	12.5 (3.1)	Mean I/sigI	21.9 (3.7)
# unique reflections	61 745	22 546	13 929	23 460	# unique reflections	25 199
Multiplicity	11.1	11.1	9.7	11.1	Multiplicity	9.8
					Anomalous completeness	99.8 (100)
					Anomalous multiplicity	5.0
					# Xe	4
					Wavelength (Å)	1.378
Refinement						
Resolution (Å)	102.1–1.90	99.9–2.60	100.9–3.10	100.5–2.58		
No. Reflections	58 623	21 401	13 210	22 272		
Rwork %	17.5 (23.1)	17.9 (25.7)	15.5 (23.6)	18.0 (24.8)		
Rfree %	19.9 (25.1)	21.9 (29.3)	22.3 (26.3)	23.0 (29.0)		
# atoms (total)	5861	5376	5361	5440		
# waters	264	20	5	37		
# metal ions	2	2	2	2		
Mean B value overall (Å ²)	33.0	53.8	60.0	47.8		
Mean B value inhibitor (Å ²)	NA	28.2	40.4	37.9		
r.m.s.d. bond lengths (Å)	0.006	0.006	0.006	0.007		
r.m.s.d. bond angles (°)	1.10	1.12	1.10	1.09		
Ramachandran analysis (%)	96.9/2.5/0.6	95.4/3.3/1.3	95.3/3.6/1.1	95.1/3.9/1.0		
	preferred/allowed/outliers					

Values in parentheses are for the high-resolution bin.

2002) or other metal ions (Na⁺, K⁺, Mn²⁺, Co²⁺, Cu²⁺, Ca²⁺ and Fe²⁺). Along with the crystallographic and anomalous scattering data, these observations suggest that the native metal for AtzD is magnesium.

The three potential catalytic dyads were investigated by site-directed mutagenesis. Amino acid substitutions at Lys42 (Ala), Ser85 (Ala), Lys162 (Ala, Arg), Ser233 (Ala), Lys296 (Ala, Arg) and Ser344 (Ala) yielded soluble, but inactive, protein. The unfolding profiles for the variants, as judged by differential scanning fluorimetry (data not shown), were consistent with correctly folded protein. While this provides no direct evidence for the identities of the catalytic residues of AtzD, it does suggest that these residues are critical for catalysis.

Determinants of substrate specificity

AtzD is inhibited by BA and Bar is inhibited by CA, despite the near identical structures of the compounds (Fig. 1A and B). To understand the differences in substrate specificity between AtzD and Bar, the structure of Bar was modelled using the AtzD structure as a template (~44% identical in sequence; Supplemental Fig. S4). The model shows a similar active site with changes in active-site residues being (AtzD to Bar): Arg195/Asn194, Gly345/Val348, Arg325/Lys328, Thr321/His324 and Gly346/Ser349 (Fig. 7). The Gly345Val change is of particular

interest as it impinges sterically on the active-site space and introduces a hydrophobic residue in a predominantly hydrophilic pocket.

The major effect of these differences is that the Bar active site does not possess the almost perfect rotational symmetry that AtzD has in the active site, instead Bar has a single axis of bilateral symmetry (Fig. 7). The differences in the active site mirror those of the substrates, in that CA possesses rotational symmetry, while BA does not due to the atomic composition at position 5 in the ring (Fig. 1A and B). The introduction of the hydrophobic surface formed by Val348 in Bar and the loss of positive charge associated with the Arg195-Asn194 substitution in Bar when compared with AtzD also dramatically changes the electronic structure of the active site. It is speculated that these steric and electronic differences underpin the substrate specificity for AtzD and Bar, albeit further data are required to support this hypothesis.

As residues 195, 321, 325, 345 and 346 in AtzD appear to play a role in determining substrate specificity, the identities of the amino acids at equivalent positions in all 68 known AtzD homologues were analysed and compared with a phylogenetic analysis of these sequences (Fig. 8). Six major phylogenetic groups were discovered. Groups I, II and III possess AtzD-like signatures at positions equivalent to 195, 321, 325, 345 and 346 in AtzD (i.e. Arg195, Thr321, Arg325, Gly345 and Gly346). Indeed, seven

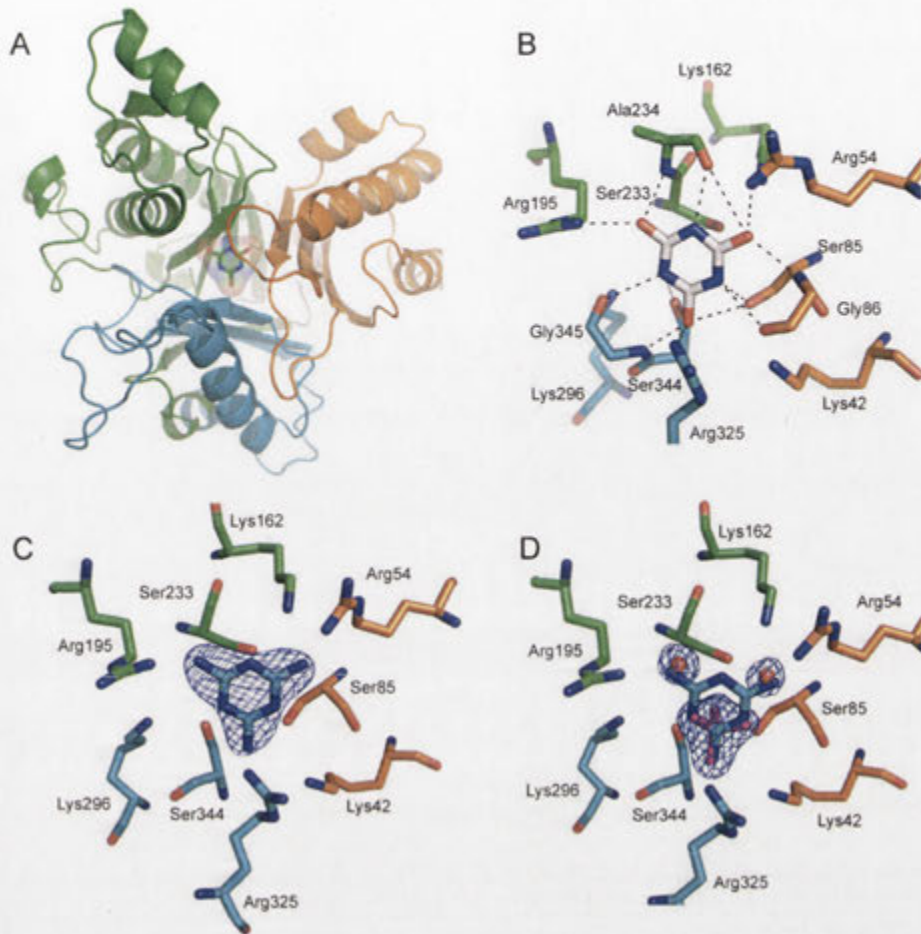


Fig. 5. The active site of AtzD.

A. The cyanuric acid-bound monomer showing the Toblerone fold (cartoon with RUs coloured as in Fig. 2A) with the substrate bound in the active site (stick and surface).

B. Active site of AtzD with cyanuric acid bound. Hydrogen bonds are shown (dashed lines).

C. The AtzD active site with bound melamine.

D. Electron density for the inhibitor is shown.

Density for the phosphate (magenta) and two water molecules (red balls) in the 'native' data set with the melamine bound AtzD structure superposed.

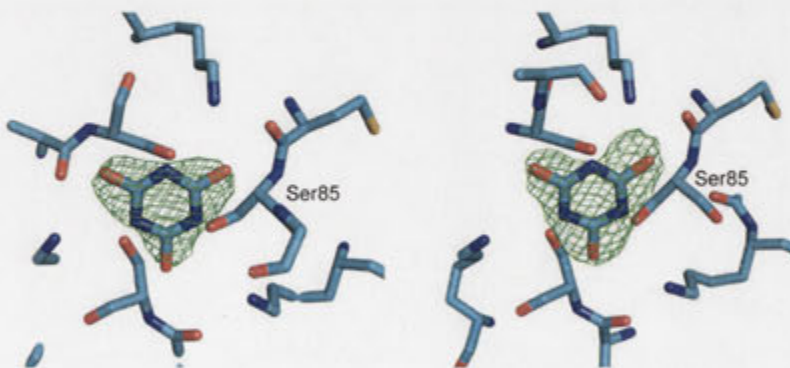


Fig. 6. Electron density of cyanuric acid bound in the AtzD active site. Left structure obtained at pH 4.5. Right structure obtained at pH 6.6. In each case Ser85 has been labelled and the electron density obtained for the substrate is shown. In the crystal obtained at pH 4.5 (too low for hydrolysis) the density has a near symmetrical distribution. In the crystal obtained at pH 6.5 (i.e. with hydrolytically active enzyme) the density has become polarized in the vicinity of Ser85, suggesting that Ser85 is the catalytic serine.

Table 2. Biochemical analysis of AtzD.

Inhibitor/treatment	Residual activity (%)
PMSF	8.5
<i>o</i> -Phenanthroline	1.2
EDTA	100
ApoAtzD plus Mg ²⁺ (50 μM) ^a	60
Lysine alkylation	7.5

a. ApoAtzD treated by the addition of 50 μM Zn²⁺, Na⁺, K⁺, Mn²⁺, Co²⁺, Cu²⁺, Ca²⁺ or Fe²⁺ failed to recover any catalytic activity.

characterized CA hydrolases (AtzD and homologues from *Pseudomonas* sp. ADP, *Moorella thermoacetica*, *Pseudomonas* sp. NRRLB-12227, *Bradyrhizobium japonicum* USDA 110, *Rhizobium leguminosarum* bv. *viciae* 3841, *Methylobacterium* sp. 4-46 and locus AZC_3892 from *Azorhizobium caulinodans* ORS 571; Karns, 1999; Seffernick *et al.*, 2009; 2012) are located within Groups II and Group III (Fig. 8).

Group VI contains the characterized Bar enzyme and there is strong conservation of a Bar-like signature within this group (i.e. Asn195, His321, Lys325, Val345 and Ser/Ala346). Groups IV and V possess signatures unlike either AtzD or Bar; Group IV has an Arg195, Met/Thr321, Arg325, Thr345 and Arg346 motif, while Group V has low conservation at positions 321 (a Gly, Arg or Trp), 325 (Lys/Arg) and 345 (Gly/Ala) and a conserved Gly at position 346. The equivalent of position 195 is absent in Group V. A single enzyme in Group V has been examined for CA and BA hydrolase activity (locus AZC_3203 from *A. caulinodans* ORS 571; Seffernick *et al.*, 2012) and no such activity could be detected. The remaining AtzD active-site residues (Lys42, Lys162, Lys296, Arg54, Arg325, Gly86, Ala234,

Ser85, Ser233 and Ser344) were conserved throughout the Toblerone fold proteins (not shown), suggesting that they are not involved in substrate specificity.

Discussion

Reaction mechanism

Structural, mutagenesis and biochemical studies suggest that AtzD, and the other members of the Toblerone fold family (TrzD and Bar), are lysine-serine hydrolases. The Michaelis complex is formed when CA binds to AtzD via an extensive hydrogen bonding network (Fig. 5B). Binding to this highly positively charged pocket promotes the formation of the triketide, eliminating the aromaticity of the substrate, with the hydrogen bonding interactions provided by Arg54, Arg195 and Arg325 possibly polarizing the substrate, making the carbonyl carbons more electrophilic and the lactones more scissile. Hydrolysis and formation of the enzyme:acyl intermediate then occurs via nucleophilic attack by a lysine-activated serine (evidenced by inhibitor and pH studies).

A number of potential catalytic mechanisms involving serine and lysine residues are possible, from a simple Ser-Lys dyad to more complex catalytic mechanisms involving multiple serine and lysine residues, with such configurations as Ser-*cis*Ser-Lys, Ser-Xxx-Xxx-Lys and Lys-Ser-Ser-Lys (McKinney and Cravatt, 2003; Ekici *et al.*, 2008; Pratt and McLeish, 2010). Although there are three pseudo-equivalent Lys-Ser dyads, it is presumed that a single serine acts as the nucleophile during a single catalytic event, as the product (carboxybiuret; Fig. 1A) is produced via hydrolysis of a single amide bond. It is not possible to definitively attribute the nucleophilic activity to any one of the three active-site serines, and it is possible

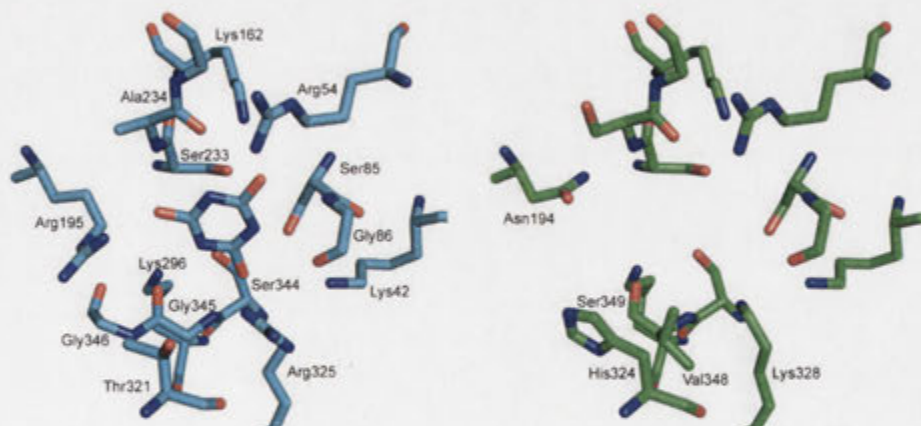


Fig. 7. Comparison of the active sites of AtzD and a homology model of Bar. The AtzD active site is shown (left, cyan) with cyanuric acid bound. The modelled Bar active site (right, green) is shown without cyanuric acid docked, only the amino acid residues that differ between AtzD and Bar are indicated on the Bar structure.

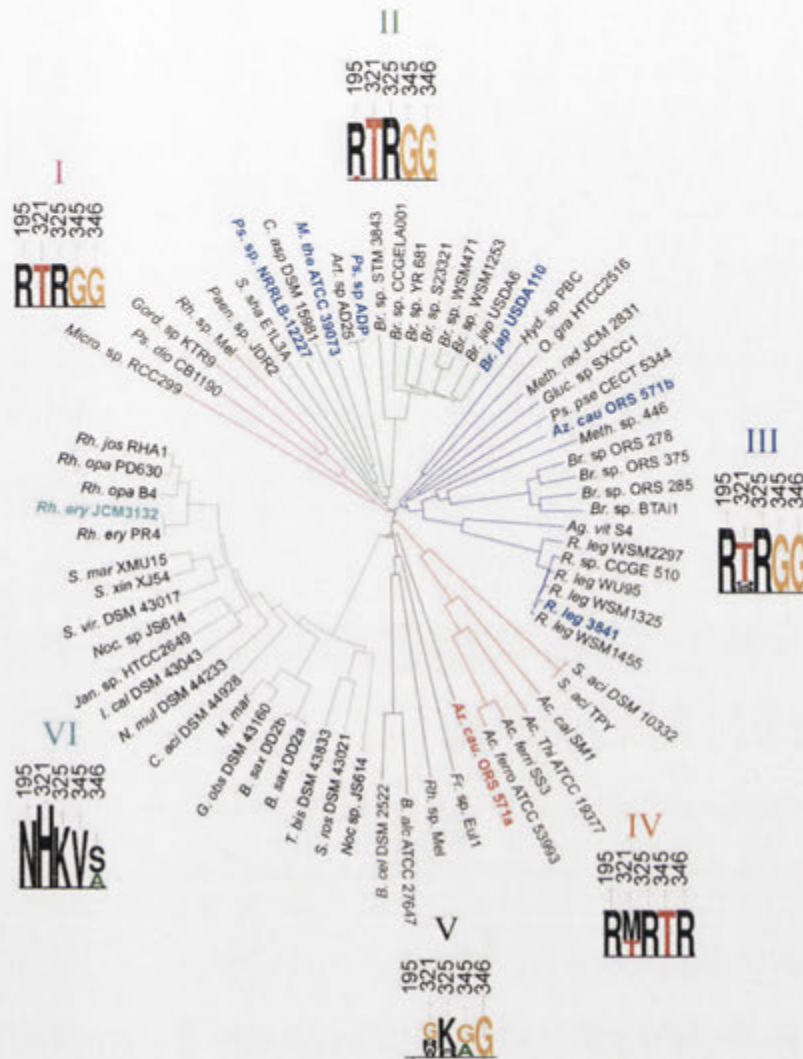


Fig. 8. Phylogeny of the AtzD/Bar family. The sequences for the AtzD/Bar homologues in the six groups were sourced from the following organisms (listed in order, clockwise): Group I (purple); *Micromonas* sp. RC299, *Pseudomonas dioxanivorans* CB1190, *Gordonia* sp. KTR8, *Rhodococcus* sp. Mel. Group II (green); *Paenibacillus* sp. JDR2, *Salinisphaera shabanensis* E1L3A, *Pseudomonas* sp. NRRLB-12227, *Clostridium asparagiforme* DSM 15981, *Moorella thermoacetica* ATCC 39073, *Arthrobacter* sp. AD25, *Bradyrhizobium* sp. STM3843, *Bradyrhizobium* sp. CCGELA001, *Bradyrhizobium* sp. YR681, *Bradyrhizobium* sp. S23321, *Bradyrhizobium* sp. WSM471, *Bradyrhizobium* sp. WSM1253 *Bradyrhizobium japonicum* USDA6, *Bradyrhizobium japonicum* USDA110. Group III (blue); *Hydrogenophaga* sp. PBC, *Oceanicola granulosus* HTCC2516, *Methylobacterium radiotolerans* JCM 2831, *Pseudomonas pseudoalcaligenes* CECT 5344, *Azorhizobium caulinodans* ORS 571, *Methylobacterium* sp. 446, *Bradyrhizobium* sp. ORS 278, *Bradyrhizobium* sp. ORS 375, *Bradyrhizobium* sp. ORS 285, *Bradyrhizobium* sp. BTA11, *Agrobacterium vitis* S4, *Rhizobium leguminosarum* bv. *trifolii* WSM2297, *Rhizobium* sp. CCGE 510, *Rhizobium leguminosarum* bv. *viciae* WU95, *Rhizobium leguminosarum* bv. *viciae* WSM1325, *Rhizobium leguminosarum* bv. *viciae* WSM1455. Group IV (red); *Sulfobacillus acidophilus* DSM 10332, *Sulfobacillus acidophilus* TPY, *Acidithiobacillus caldus* SM1, *Acidithiobacillus thiooxidans* ATCC 39073, *Acidithiobacillus ferrooxidans* ATCC 27647, *Acidithiobacillus caldus* SM1, *Acidithiobacillus thiooxidans* ATCC 39073, *Acidithiobacillus ferrooxidans* ATCC 27647, *Acidithiobacillus caldus* SM1, *Acidithiobacillus thiooxidans* ATCC 39073, *Acidithiobacillus ferrooxidans* ATCC 27647. Group V (black); *Azorhizobium caulinodans* ORS571, *Frankia* sp. Eul1c, *Rhodococcus* sp. Mel, *Bacillus alcalophilus* ATC 27647, *Bacillus cellulolyticus* DSM 2552. Group VI (cyan); *Nocardioides* sp. JS614, *Streptosporangium roseum* DSM 43021, *Thermobispora bisporea* DSM 43833, *Blastococcus saxobidens* DD2, *Blastococcus saxobidens* DD2, *Geodermatophilus obscures* DSM 44928, *Modestobacter marinus*, *Catenulispora acidiphia* DSM 44928, *Nakamurella multipartita* DSM 432021, *Intrasporangium calvum* DSM 43043, *Janibacter* sp. HTCC2649, *Nocardioides* sp. JS614, *Saccharomonospora viridis* DSM 43017, *Saccharomonospora xinjiangensis* XJ54, *Saccharomonospora marina* XMU15, *Rhodococcus erythropolis* PR4, *Rhodococcus erythropolis* JCM3132, *Rhodococcus opacus* B4, *Rhodococcus opacus* PD630, *Rhodococcus jostii* RHA1. The identities and conservation of amino acids found in each at positions equivalent to AtzD 195, 321, 325, 345 and 346 are shown as logos. The names of bacterial species from which Tolberone fold enzymes have had their catalytic activities characterized are emboldened and coloured to indicate their substrate specificity: blue for cyanuric acid hydrolase, cyan for barbituric acid hydrolase and red for the absence of hydrolytic activity with either cyanuric acid or barbituric acid.

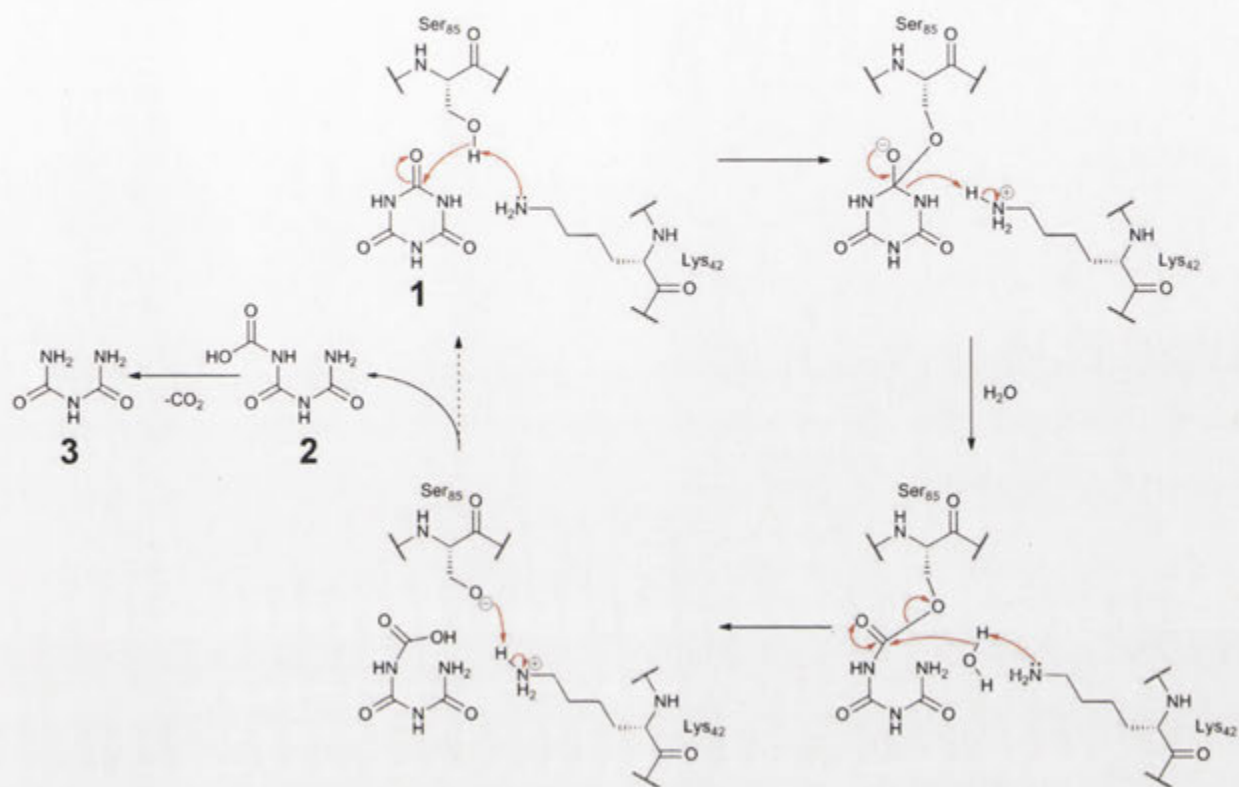


Fig. 9. Plausible catalytic mechanism for cyanuric acid hydrolysis by AtzD. Ser85 is activated by Lys42 and acts as nucleophile with bound CA (1), allowing nucleophilic attack at one of the three ketone positions of CA. The tetrahedral intermediate formed undergoes ring opening to form an acyl intermediate. The acylated serine is subsequently hydrolysed by a solvent water molecule, again activated by Lys42, via a tetrahedral intermediate to form carboxybiuret (2) which spontaneously decarboxylates to form biuret (3) and the active site is regenerated by intramolecular proton transfer.

that any or all three may act as nucleophile (albeit only one can do so *per* catalytic cycle).

The position of the substrate in the active site (Fig. 5A and B) places Ser85 such that it is closer to a scissile bond than are the other serine residues. Additionally, pH-dependent polarization of the substrate's electron density and a reduction in abundance of the Ser85 containing polypeptide in mass spectra of PMSF-treated tryptic digests of AtzD are suggestive that Ser85 is the dominant (if not sole) nucleophile in the hydrolytic mechanism. A hypothetical reaction mechanism consistent with a Ser85 nucleophile is therefore proposed: Lys42 is a general base, activating Ser85 and promoting formation of a tetrahedral intermediate between Ser85 and the closest substrate carbonyl carbon, this then resolves into the acyl:enzyme intermediate following ring opening (Fig. 9). Thereafter, a solvent water molecule is required to hydrolyse the acyl intermediate and regenerate the serine, liberating carboxybiuret (which then spontaneously decarboxylates to form biuret; Seffernick *et al.*, 2012).

Evolution of the Toblerone fold

To the authors' knowledge AtzD possesses a unique fold, related to, but distinct from, the YgjF-family of trimeric proteins. At some point in their evolutionary history, the subunits of the trimeric ancestor of AtzD and Bar concatenated, possibly following successive gene duplications. On the other hand, the RUs may have been 'recruited' from other members of the YgjF protein family. An alternative hypothesis is that an ancestral Toblerone fold protein was truncated to one-third its original length (i.e. a single RU) that was capable of forming stable trimers. However, YgjF-family members are found in all domains of life, catalysing a wide range of reactions and participating in diverse physiological processes (Kim *et al.*, 2001; Schaap *et al.*, 2002; Burman *et al.*, 2007). Until the introduction of *s*-triazine herbicides, the AtzD/Bar enzymes were only known to participate in a single physiological process (BA hydrolysis) in a limited number of species (Seffernick *et al.*, 2012). The concatenation scenario is therefore the more likely of the two possible evolutionary histories.

The AtzD active site is located in the central 'barrel' of the protein (Fig. 5). In contrast, the three binding pockets of each PSP/RutC trimer are formed by the interactions between adjoining monomers and are located on the outside of the protein (Zhang *et al.*, 2010; Knapik *et al.*, 2012). In both PSP and RutC, the central barrel provides the core interactions that maintain the trimer (Zhang *et al.*, 2010; Knapik *et al.*, 2012). In the AtzD family, the strength of these core interactions may have been reduced in an evolutionary trade-off with the optimization of the catalytic function, which could have resulted in a destabilization of the protein. It is plausible that concatenation of the three RUs of AtzD compensates for the loss of the stabilizing core interactions, although it is unclear if concatenation was a prerequisite or a consequence of this trade off.

Removal of the metal ion from AtzD greatly reduced its stability, resulting in protein precipitation at room temperature, suggesting that formation of the metal binding site is also an evolutionary innovation that provides enhanced stability. It is notable that none of the other cations tested appear to substitute for Mg²⁺. Bar has been reported to be Zn²⁺-dependent (Soong *et al.*, 2002), despite the fact that the metal binding sites of AtzD and Bar are completely conserved (as it is in 90% of the known AtzD/Bar homologues). Moreover, the metal binding site does not appear well adapted for binding Zn²⁺, albeit the structure of apoAtzD was not determined due to its inherent instability, and it is conceivable that the structure of the metal-binding pockets in holoAtzD/Bar are, in part, determined by the identity of the bound cation. It is also possible that there is a second, unidentified, Zn²⁺ binding site in Bar. However, there are no canonical Zn²⁺-binding motifs in the primary sequence of Bar (Seffernick *et al.*, 2012), and so such a site would necessarily be another novel metal binding site.

Distribution of CA and BA hydrolase activities

It appears that there are six phylogenetically distinct groups within the AtzD/Bar family. The six groups are also separated by the identities of the amino acids that are likely to distinguish substrate specificities of AtzD and Bar. Phylogenetic Group II contains AtzD and TrzD and the predicted determinants of substrate specificity for CA are highly conserved throughout this group (Fig. 8), as they are in Groups I and III. This may suggest that Groups I–III contain CA hydrolases. Group VI contains Bar and the Bar-like specificity signature is highly conserved within this group, suggesting that Group VI contains BA hydrolases. The high level of conservation for the 'specificity signature' residues within these groups is consistent with their inferred roles in determining substrate specificity. Moreover, a number of enzymes in Groups II, III and VI

have been characterized elsewhere and found to possess activities consistent with the proposed CA or BA hydrolases (Fig. 8; Seffernick *et al.*, 2012).

Groups IV and V are unlike the CA hydrolases, BA hydrolases or each other, and may fulfil different physiological roles. Indeed, a single Group V enzyme from *A. caulinodans* ORS571 has been characterized (Seffernick *et al.*, 2012), and found to possess neither AtzD-like nor Bar-like substrate specificities. Group V is composed of proteins from five bacterial species (from the genera *Rhizobium*, *Azorhizobium*, *Frankia* and *Bacillus*). There is far lower conservation in the amino acids identified in this study as specificity determinants, suggesting that these enzymes may not share substrate specificities.

It is interesting to note that the Group IV enzymes belong to the genera *Acidithiobacillus* (Gram-negative, γ -proteobacter) and *Sulfobacillus* (Gram-positive, Clostridiales). Although phylogenetically distinct, *Acidithiobacillus* and *Sulfobacillus* are physiologically similar (Waksman and Joffe, 1922; Temple and Colmer, 1951; Golovacheva and Karavaiko, 1978; Kovalenko and Malakhova, 1983; Karavajko *et al.*, 1990; Narayan and Sahana, 2009); both are acidophiles that oxidize iron and elemental sulphur, indeed both are exploited in biomining/bioleaching applications (Narayan and Sahana, 2009; Tang *et al.*, 2009). It may be that the Toblerone fold enzymes of these two genera fulfil a physiological role that is unique to their unusual physiologies.

Conclusion

AtzD is the archetype for the Toblerone fold, which has likely evolved by the concatenation of three genes encoding proteins of the YgjF superfamily. The concatenation, along with the evolution of a stabilizing metal binding site (that may also help organize the active site), has allowed the repurposing of the stabilizing central barrel of the monomer for catalysis, which has resulted in a unique family of ring-opening hydrolases. Active-site residues have been identified that differ between AtzD and Bar, and are therefore implicated in determining substrate specificity. The inferred 'specificity signature' varies between the six phylogenetic groups of the Toblerone fold family, and is well conserved within Groups I–IV and VI, suggesting that these residues may indeed confer substrate specificities to this family of enzymes.

Experimental procedures

DNA manipulation

The *atzD* gene (Accession No. U66917) was provided by GenScript, as an NdeI/BamHI insert in pUC57. The *atzD* gene was subcloned into the NdeI and BamHI sites of

pETCC2. The pETCC2 expression vector was a gift from Dr Christopher Coppin (CSIRO Ecosystem Sciences) and is derived from pET14b (Novagen). An in-frame N-terminal hexahis-tag along with thrombin cleavage site (MGSSHHH-HHHSSGLVPRGSH) was added to the encoded enzyme as a result of the subcloning.

Mutagenesis was by the over-lapping PCR method of Ho *et al.* (1989). The oligonucleotide primers used are detailed in Supplemental Table S2. Amplicons were resolved on a 0.8% agarose gel and purified using NucleoSpin® Gel kits (Machery Nagel). Purified BamHI and NdeI digested amplicons were cloned into the pETCC2 vector using T4 DNA ligase (NEB). Sequence verification was done by Macrogen (Korea).

Protein expression and purification

Electrocompetent *E. coli* BL21 λ (DE3) (Invitrogen) were transformed with appropriate plasmids (Suppl Table 2), then grown in Lennox medium (LB; Lennox, 1955) or on LB agar (15 mg ml⁻¹ agar, Merck) supplemented with 100 μ g ml⁻¹ ampicillin (Sigma Aldrich). Starter cultures (50 ml) were grown from single colonies at 37°C for 18 h while being shaken at 200 r.p.m., used to inoculate 950 ml of LB and incubated at 37°C to an OD₆₀₀ of 0.6–0.8. Cultures were induced by addition of 100 μ M isopropyl- β -D-thiogalactopyranoside (IPTG; Astral) and incubated at 37°C overnight while shaking at 200 r.p.m. Cultures were then harvested by centrifugation (4000 *g*, 10 min) in an Avanti J-E centrifuge (Beckman Coulter) and resuspended in lysis buffer (50 mM sodium phosphate pH 7.5, 50 mM NaCl) and lysed by passing through Avestin C3 homogenizer three times at 18 000 p.s.i. Insoluble material was removed by centrifugation (18 000 *g*) using an Avanti J-E centrifuge.

Metal ion affinity chromatography (Ni-NTA Superflow Cartridge; Qiagen) was used as per the manufacturer's instructions. AtzD eluted at 160 mM imidazole. AtzD was further purified by size exclusion chromatography using a 130 ml column packed with Superdex 200 prep grade resin (GE Healthcare Life Sciences) using lysis buffer. AtzD was estimated to be >98% pure by Coomassie (Sigma Aldrich) stained precast SDS-PAGE gel (4–20% Tris-HEPES-SDS, Thermo Scientific), and typical yields were 7.5–8.5 mg from 1 l of culture.

Crystallization

Crystals of native AtzD were prepared as follows: concentrated protein (10.4 mg ml⁻¹ in lysis buffer) was set up in crystallization droplets consisting of 200 nl protein solution, 250 nl crystallization cocktail and 50 nl additive solution [0.1 M sodium HEPES pH 6.9, 0.2 M NH₄Cl, 20% (w/v) polyethylene glycol (PEG) 6000]. The crystallization cocktail contained a low-molecular-weight PEG (either PEG 300, PEG monomethyl ether 350 or PEG 400) between 35% and 50% (v/v) in sodium phosphate (50–100 mM; pH 6.5–8) buffer, with or without NaCl (100–200 mM). All chemicals were obtained from Sigma Chemicals (St. Louis, USA). Droplets were set up with a Phoenix nanodispensing robot (ARI, Sunnyvale, CA) into SD-2 plates (Molecular Dimensions, UK) and

were stored at 20°C. Triangular prism-shaped crystals appeared within 24 h, and grew large enough to obtain X-ray data within a week (Supplemental Fig. S5).

Substrate/inhibitor bound AtzD crystals were prepared as follows: AtzD protein in 50 mM sodium phosphate pH 7.5, 50 mM NaCl was dialysed against HEPES buffer (50 mM, pH 7.5) containing NaCl (50 mM), DTT (5 mM) and melamine (10 mM), barbituric acid (10 mM) or cyanuric acid (0.1 mM) and then set up in crystallization experiments as described above, against crystallization conditions without phosphate.

Crystal data collection

Crystals were taken to the MX-2 beamline of the Australian Synchrotron and cryo-cooled to 100 K directly in the nitrogen stream. A xenon derivative was prepared by exposing a crystal to xenon gas at 20 p.s.i. for 2 min, and flash-cooled in liquid nitrogen. A total of 720 frames of 0.5° oscillation were used to collect 360 degrees of data for the Xe derivative data set. All further data sets consisted of 360 frames of 0.5° oscillation for a total of 180 degrees of data. Each was processed as described below. Melamine, barbituric acid and cyanuric acid were all placed in clear density using Afitt (OpenEye Scientific Software, USA) for those data sets collected after co-crystallization with these compounds.

Anomalous data were collected at 1.7149 Å (7230 eV) on 17 isomorphous AtzD native crystals and the best data were merged to give a data set with 140-fold multiplicity which gave exquisite anomalous difference maps, showing every sulphur atom, the phosphate ions as well as the two metal binding sites (Supplemental Fig. S1).

Structural determination

Data sets were indexed with XDS (Kabsch, 2010) and processed with SCALA (Evans, 2006), and entered into the SIRAS protocol in Auto-Rickshaw (Panjikar *et al.*, 2005) to locate the four xenon atoms. Phases from this derivative were sufficient to allow Buccaneer (Cowtan, 2006) to auto-build a partial model, which was used in Phaser (McCoy *et al.*, 2007) as a molecular replacement model (crystallographic statistics are found in Table 1). These data allowed the complete tracing of the chain from residue 2 to residue 364, using the program COOT (Emsley *et al.*, 2010). The model was refined using Refmac (Murshudov *et al.*, 1997) to a final resolution of 1.90 Å, and showed clear density for the entire chain excepting weak density for the residues of the loop from residues 105 to 109. This model was subsequently used for molecular replacement using Phaser (McCoy *et al.*, 2007) for subsequent data sets with inhibitors bound.

Small-angle X-ray scattering (SAXS)

AtzD was dialysed overnight into a 50 mM phosphate buffer, 50 mM NaCl, 1 mM DTT (pH 7.5). The same buffer was used as the buffer standard during data collection. A dilution series of AtzD (from 0.19 to 3 mg ml⁻¹) was prepared, and scattering data were collected for 1 s using a Pilatus 1 M photon counting detector (Dektris) with a sample to detector distance of 1.6 m. Ten replicate images were collected for each

sample and averaged, with outlier rejection, to control for radiation damage. Data were measured in a Q range from 0.01 to 0.5 Å⁻¹ and at the highest protein concentration the scattering remained above the noise threshold to the edge of the detector.

AtzD (75 µl, 6 mg ml⁻¹) was injected on to a size exclusion column (Wyatt Silica Bead column, 300 Å pore size) that had been pre-equilibrated with the PO₄/NaCl buffer. The column was developed at 0.2 ml min⁻¹, and a single peak eluted. The SAXS scattering showed no change over the peak (data not shown). A CRYSOLO (Svergun *et al.*, 1995) fit of various quaternary structures for the protein to the scattering data was performed (Supplemental Fig. S2).

Kinetic analysis

Hydrolysis of CA and BA was assayed by UV-visible spectroscopy as described elsewhere (Seffernick *et al.*, 2012), using a SpectraMax M2 spectrophotometer (Molecular Devices). Reactions [200 µl in 1 mM TAPS (N-Tris(hydroxymethyl)methyl-3-aminopropanesulphonic acid; pH 8.5); 1 µM AtzD] with 10–1000 µM substrate were assayed in UV-transparent 96-well microtitre plates (Greiner Bio-one). Absorbances were measured at 15 s intervals for 10 min. The pH-dependence of AtzD was determined in MOPS [3-(N-morpholino)propanesulphonic acid; 1 mM, pH 6.5, 6.9 and 7.5], TAPS (1 mM pH 7.5, 7.9 and 8.5) and CHES (N-Cyclohexyl-2-aminoethanesulphonic acid; 1 mM pH 9).

Enzyme inhibition

Alkylation of lysines in AtzD was carried out using JBS Methylation Kit (Jena Bioscience) following manufacturer's guidelines. A protein concentration of 6 mg ml⁻¹ in sodium phosphate (50 mM, pH 7.5)/NaCl buffer (50 mM) was used at the time of treatment. Phenylmethylsulphonyl fluoride (PMSF) was prepared at 100 mM in isopropanol. Enzyme solutions were prepared at 4 µM concentrations in MOPS buffer (1 mM, pH 6.5) and treated twice with PMSF (1 mM final concentration) for 30 min. Trypsin was used to as a positive control using N_ε-Benzoyl-L-arginine ethyl ester hydrochloride (BAEE) as a control substrate. Trypsin activity was monitored by UV spectrophotometrically at 253 nm.

Ethylenediaminetetraacetic acid (EDTA; 20 mM) and o-phenanthroline (75 mM) in 25 mM MOPS buffer (pH 8.5) were used to dialyse AtzD (117 µM) overnight at 4°C. Chelators were removed by centrifugation in a 30 kDa spin column (Amicon Ultra, Millipore). The zinc amidohydrolase TrzN (E.C. 3.8.1.8) was used as a positive control. Metals (Mg²⁺, Na⁺, K⁺, Mn²⁺, Co²⁺, Cu²⁺, Ca²⁺, Zn²⁺ and Fe²⁺; 50 µM) were added to o-phenanthroline treated ApoAtzD and incubated at 4°C for 16 h.

Mass spectrometry

Peptides generated from tryptic digestions of AtzD were acidified with 0.5% formic acid, filtered [Millex-LG, low protein binding hydrophilic LCR (PTFE), 0.2 µm], and analysed by nanoflow, reversed phase, liquid chromatography-tandem mass spectrometry (MS) using an Agilent Chip Cube system

coupled to an Agilent XCD ion trap mass spectrometer. MS/MS spectra were assigned to sequences by several rounds searching and 'autovalidation' using SpectrumMill software (Agilent Rev A.03.03.084 SR4). First a small database of common contaminant sequences such as trypsin and keratins was searched assuming tryptic cleavages. Remaining unassigned spectra were used to search just the AtzD sequence allowing trypsin- and chymotrypsin-specific cleavages. The trypsin (Sigma-Aldrich) had not been treated to prevent chymotrypsin-specific cleavages and this generated some mixed tryptic/chymotryptic cleavages that usefully extended the sequence coverage. Finally, remaining unassigned spectra were used to search the AtzD sequence in 'homology' mode allowing single unassigned mass gaps in sequences that were otherwise consistent with MS/MS spectra. Spectra assigned in this last mode were inspected individually to verify putative modifications of amino acid residues.

Modelling of Bar

Bar was modelled using the Swiss-Model software as found on the web server (Arnold *et al.*, 2006); the AtzD structure was used as the starting point (44% sequence identity over 364 residues). There are some minor differences between the structures in various loops: residues 74–78 (AtzD) has an extra residue in the AtzD loop; residues 106–111 (AtzD) this loop of weak density in the AtzD structure is built slightly differently in the Bar structure; residues 210–215 (AtzD) has two extra residues in the AtzD loop; residues 280–283 (AtzD) has six extra residues in the Bar modelled loop; the AtzD structure has a single amino acid residue extension at the C-terminus of the protein as well.

In silico sequence analysis

AtzD homologues were identified via a BlastP search of non-redundant databases. The following sequences were returned with *E*-values of less than 10⁻³⁰ and identity scores of greater than 40%, there were no other hits with *E*-values of less than 10¹: NP862537 (*Pseudomonas* sp. ADP), ABK41866 (*Arthrobacter* sp. AD25), YP430955 (*M. thermoacetica* ATCC 39073), ZP10581004 (*Bradyrhizobium* sp. YR681), ZP09650932 (*Bradyrhizobium* sp. WSM471), P0A3V4 (*Pseudomonas* sp. NRRRL-12227), ZP10083023 (*Bradyrhizobium* sp. WSM1253), YP005453208 (*Bradyrhizobium* sp. S23321), YP005606973 (*B. japonicum* USDA 6), ZP09433530 (*Bradyrhizobium* sp. STM 3843), EJZ29306 (*Bradyrhizobium* sp. CCGE-LA001), NP773921 (*B. japonicum* USDA 110), ZP03758143 (*Clostridium asparagiforme* DSM 15981), YP003013624 (*Paenibacillus* sp. JDR-2), YP001757420 (*Methylobacterium radiotolerans* JCM 2831), ZP10760424 (*Pseudomonas pseudoalcaligenes* CECT 5344), ZP08551100 (*Salinisphaera shabanensis* E1L3A), ZP08318418 (*Gluconacetobacter* sp. SXCC-1), YP004783181 (*Acidithiobacillus ferrivorans* SS3), YP002219377 (*Acidithiobacillus ferrooxidans* ATCC 53993), YP001526119 (*A. cauliodans* ORS 571), YP001208170 (*Bradyrhizobium* sp. ORS 278), YP002547456 (*Agrobacterium vitis* S4), YP004750255 (*Acidithiobacillus caldus* SM-1), EIW44662 (*R. leguminosarum* bv. *trifolii* WU95),

YP770629 (*R. leguminosarum* bv. *viciae* 3841), EJC71551 (*R. leguminosarum* bv. *viciae* WSM1455), YP002979447 (*R. leguminosarum* bv. *trifolii* WSM1325), ZP09420969 (*Bradyrhizobium* sp. ORS 375), YP001526808 (*A. caulinodans* ORS 571), ZP10837886 (*Rhizobium* sp. CCGE 510), EJC83804 (*R. leguminosarum* bv. *trifolii* WSM2297), ZP09473782 (*Bradyrhizobium* sp. ORS 285), YP001237458 (*Bradyrhizobium* sp. BTAi1), ZP01155857 (*Oceanicola granulosus* HTCC2516), ZP09997547 (*Acidithiobacillus thiooxidans* ATCC 19377), ZP10152173 (*Hydrogenophaga* sp. PBC), YP006671162 (*Gordonia* sp. KTR9), YP005257446 (*Sulfobacillus acidophilus* DSM 10332), YP004719285 (*S. acidophilus* TPY), YP001770627 (*Methylobacterium* sp. 4-46), AEX65082 (*Rhodococcus* sp. Mel), YP00433323 (*Pseudonocardia dioxanivorans* CB1190), XP002503480 (*Micromonas* sp. RCC299), CAC86669 (*Rhodococcus erythropolis*), YP004017027 (*Frankia* sp. Eul1c), YP004094229 (*Bacillus cellulosilyticus* DSM 2522), YP003112640 (*Catenulisporea acidiphila* DSM 44928), YP002769329 (*R. erythropolis* PR4), YP922706 (*Nocardioidea* sp. JS614), YP002779991 (*Rhodococcus opacus* B4), YP005331388 (*Blastococcus saxosidens* DD2), ZP00996765 (*Janibacter* sp. HTCC2649), YP003133937 (*Saccharomonospora viridis* DSM 43017), YP005327200 (*B. saxosidens* DD2), YP006364153 (*Modestobacter marinus*), ZP09986349 (*Saccharomonospora xinjiangensis* XJ-54), YP003204583 (*Nakamurella multipartita* DSM 44233), ZP09740740 (*Saccharomonospora marina* XMU15), EHI43897 (*R. opacus* PD630), YP004098318 (*Intrasporangium calvum* DSM 43043), YP003407358 (*Geodermatophilus obscurus* DSM 43160), YP003341997 (*Streptosporangium roseum* DSM 43021), YP705235 (*Rhodococcus jostii* RHA1), YP925454 (*Nocardioidea* sp. JS614), YP003652816 (*Thermobispora bispora* DSM 43833), ZP10819722 (*Bacillus alcalophilus* ATCC 27647), AEX65050 (*Rhodococcus* sp. Mel).

Alignments and bootstrapping (1000×) of the AtzD superfamily sequences were carried out using CLUSTALW (<http://www.bioinformatics.nl/tools/clustalw.html>). A phylogenetic tree was constructed from this alignment using the Interactive Tree of Life (iTol; Letunic and Bork, 2007; 2011) and consensus sequences were created using Weblogo sequence generator (Crooks et al., 2004).

Accession codes

The molecular models for AtzD and AtzD bound to CA, melamine and BA were deposited in the Protein Data Bank, with PDB codes 3ZGR, 3ZGS, 3ZGT and 3ZGT.

Acknowledgements

We thank Dr Peter Campbell (CSIRO Ecosystem Sciences) for his assistance with mass-spectrometry, Drs Carol Hartley and Andrew Warden (CSIRO Ecosystem Sciences) for their constructive comments, and the beamline scientists at the Australian Synchrotron for their help in data collection. We also thank James Holton and Christine Gee for help with anomalous data collection. Crystals were grown in the CSIRO Collaborative Crystallisation Centre (<http://www.csiro.au/C3>) at CSIRO Materials Science and Engineering. This work was supported in part by an Office of the Chief Executive Julius Career Development Award (C.S.).

This work was supported in part by an Office of the Chief Executive Julius Career Development Award (C.S.).

References

- Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**: 195–201.
- Botos, I., and Wlodawer, A. (2007) The expanding diversity of serine hydrolases. *Curr Opin Struct Biol* **17**: 683–690.
- Burman, J.D., Stevenson, C.E.M., Sawers, R.G., and Lawson, D.M. (2007) The crystal structure of *Escherichia coli* TdcF, a member of the highly conserved YjgF/YER057c/UK114 family. *BMC Struct Biol* **7**: 30.
- Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr* **62**: 1002–1011.
- Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**: 1188–1190.
- Dudev, T., Cowan, J.A., and Lim, C. (1999) Competitive binding in magnesium coordination chemistry: water versus ligands of biological interest. *J Am Chem Soc* **121**: 7665–7673.
- Ekici, O.D., Paetzel, M., and Dalbey, R.E. (2008) Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci* **17**: 2023–2037.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010) Features and development of COOT. *Acta Crystallogr D Biol Crystallogr* **66**: 486–501.
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**: 72–82.
- Fruchey, I., Shapir, N., Sadowsky, M.J., and Wackett, L. (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.
- Golovacheva, R.S., and Karavaiko, G.I. (1978) New genus of thermophilic spore-forming bacteria, *Sulfobacillus*. *Microbiology* **47**: 658–665.
- Hayes, T., Haston, K., Tsui, M., Hoang, A., Haeffele, C., and Vonk, A. (2003) Atrazine-induced hermaphroditism at 0.1 ppb in American leopard frogs (*Rana pipiens*): laboratory and field evidence. *Environ Health Perspect* **111**: 568–575.
- Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989) Site-directed mutagenesis by overlap extension using the polymerase chain-reaction. *Gene* **77**: 51–59.
- Kabsch, W. (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**: 125–132.
- Karavajko, G.I., Bulygina, E.S., Tsaplina, I.A., Bogdanova, T.I., and Chumakov, K.M. (1990) *Sulfobacillus thermonisulfooxidans* – a new lineage of bacterial evolution. *FEBS Lett* **261**: 8–10.
- Karns, J.S. (1999) Gene sequence and properties of an s-triazine ring-cleavage enzyme from *Pseudomonas* sp. strain NRRLB-12227. *Appl Environ Microbiol* **65**: 3512–3517.

- Kim, J.M., Yoshikawa, H., and Shirahige, K. (2001) A member of the YER057c/yjgF/Uk114 family links isoleucine biosynthesis and intact mitochondria maintenance in *Saccharomyces cerevisiae*. *Genes Cells* **6**: 507–517.
- Knapik, A.A., Petkowski, J.J., Otwinowski, Z., Cymborowski, M.T., Cooper, D.R., Chruszcz, M., *et al.* (2012) Structure of *Escherichia coli* RutC, a member of the YjgF family and putative aminoacrylate peracid reductase of the rut operon. *Acta Crystallograph Sect F Struct Biol Cryst Commun* **68**: 1294–1299.
- Kovalenko, E.V., and Malakhova, P.T. (1983) The spore-forming iron-oxidizing bacterium *Sulfobacillus thermosulfidoxidans*. *Microbiology* **52**: 760–763.
- Lennox, E.S. (1955) Transduction of linked genetic characters of the host by bacteriophage P1. *Virology* **1**: 190–206.
- Letunic, I., and Bork, P. (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* **23**: 127–128.
- Letunic, I., and Bork, P. (2011) Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res* **39**: W475–W478.
- 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.
- McKinney, M.K., and Cravatt, B.F. (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.
- MacLennan, P.A., Delzell, E., Sathiakumar, N., Myers, S.L., Cheng, H., Grizzle, W., *et al.* (2002) Cancer incidence among triazine herbicide manufacturing workers. *J Occup Environ Med* **44**: 1048–1058.
- 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.
- Narayan, S.J., and Sahana, S. (2009) Bioleaching: a review. *Res J Biotechnol* **4**: 72–75.
- Noor, S., Taylor, M.C., Russell, R.J., Jermini, L.S., Jackson, C.J., Oakeshott, J.G., and Scott, C. (2012) Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS ONE* **7**: e39822. doi:10.1371/journal.pone.0039822
- Paetzl, M., Dalbey, R.E., and Strynadka, N.C.J. (1998) Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature* **396**: 186–190.
- Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S., and Tucker, P.A. (2005) Auto-Rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr* **61**: 449–457.
- Pratt, R.F., and McLeish, M.J. (2010) Structural relationship between the active sites of beta-lactam-recognizing and amidase signature enzymes: convergent evolution? *Biochemistry* **49**: 9688–9697.
- Schaap, F.G., van der Vusse, G.J., and Glatz, J.F.C. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates. *Mol Cell Biochem* **239**: 69–77.
- Seffernick, L.J., Sadowsky, M.J., and Wackett, L.P. (2009) Thermostable cyanuric acid hydrolase from *Moorella thermoacetica* ATCC 39073. *Appl Environ Microbiol* **75**: 6986–6991.
- Seffernick, L.J., Erickson, J.S., Cameron, S.M., Cho, S., Dodge, A.G., Richman, J.E., *et al.* (2012) Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. *J Bacteriol* **194**: 4579–4588.
- Soong, C.L., Ogawa, J., and Shimizu, S. (2001) Novel amidohydrolytic reactions in oxidative pyrimidine metabolism: analysis of the barbiturase reaction and discovery of a novel enzyme, ureidomalonnase. *Biochem Biophys Res Commun* **286**: 222–226.
- Soong, C.L., Ogawa, J., Sakuradani, E., and Shimizu, S. (2002) Barbiturase, a novel zinc-containing amidohydrolytic enzyme involved in oxidative pyrimidine metabolism. *J Biol Chem* **277**: 7051–7058.
- Svergun, D., Barberato, C., and Koch, M.H.J. (1995) CRYSOLE – a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**: 768–773.
- Tang, K., Baskaran, V., and Nemati, M. (2009) Bacteria of the sulphur cycle: an overview of microbiology, biokinetics and their role in petroleum and mining industries. *Biochem Eng J* **44**: 73–94.
- Temple, K.L., and Colmer, A.R. (1951) The autotrophic oxidation of iron by a new bacterium *Thiobacillus ferrooxidans*. *J Bacteriol* **62**: 605–611.
- Udikovic-Kolic, N., Scott, C., and Martin-Laurent, F. (2012) Evolution of atrazine-degrading capabilities in the environment. *Appl Microbiol Biotechnol* **96**: 1175–1189.
- Wackett, L.P. (2009) Questioning our perceptions about evolution of biodegradative enzymes. *Curr Opin Microbiol* **12**: 244–251.
- Waksman, S.A., and Joffe, J.S. (1922) Microorganisms concerned in the oxidation of sulfur in the soil II. *Thiobacillus thiooxidans*, a new sulfur-oxidizing organism isolated from the soil. *J Bacteriol* **7**: 239–256.
- Wiegand, C., Krause, E., Steinberg, C., and Pflugmacher, S. (2001) Toxicokinetics of atrazine in embryos of the zebrafish (*Danio rerio*). *Ecotoxicol Environ Saf* **49**: 199–205.
- Zhang, H.M., Gao, Y., Li, M., and Chang, W.R. (2010) Crystal structure of the PSPTO-PSP protein from *Pseudomonas syringae* pv. *tomato* str. DC3000 in complex with D-glucose. *Biochem Biophys Res Commun* **397**: 82–86.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Differential Scanning Fluorimetry of AtzD and its active site variants

The data from DSF experiments are provided here for the reviewers. In each experiment the variants were analysed with 8-fold replication, with the wild-type as control in each case.

The first derivatives are shown, and there is a clear difference in apparent T_m between the wild-type and variants. Unexpectedly, in some of the variants the change in T_m indicates that the amino acid substitution has caused a significant gain in stability. We do not believe that these unusual observations make a material difference to work described in the submitted manuscript: i.e. the proteins appear to be folded. However, we would like to better understand the cause of the unexpected increases in T_m before releasing these data.

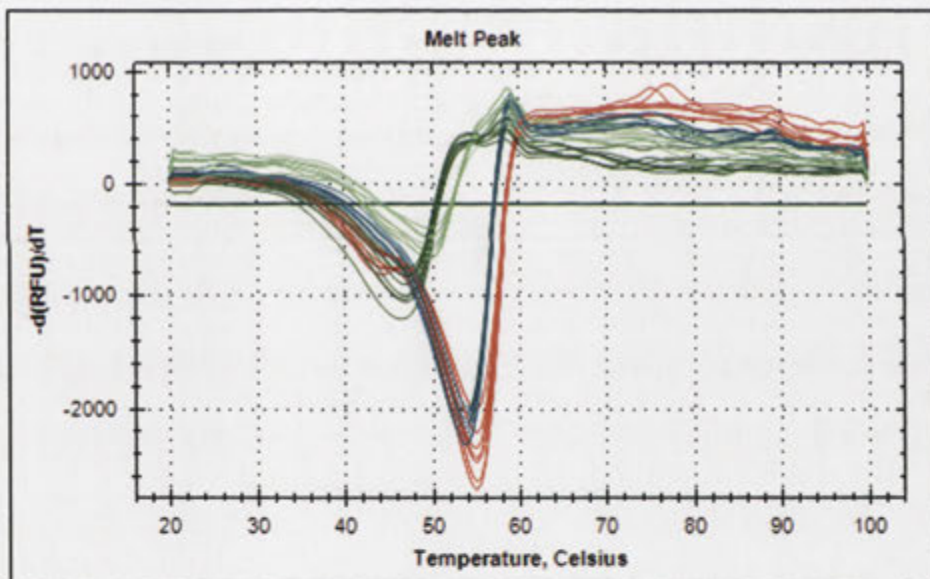
Serine to alanine substitutions:

Light green – S85A (49.5-50C)

Teal – S233A (53.5C)

Green - S344 A (46.5-47C)

Red – WT (55.0C)



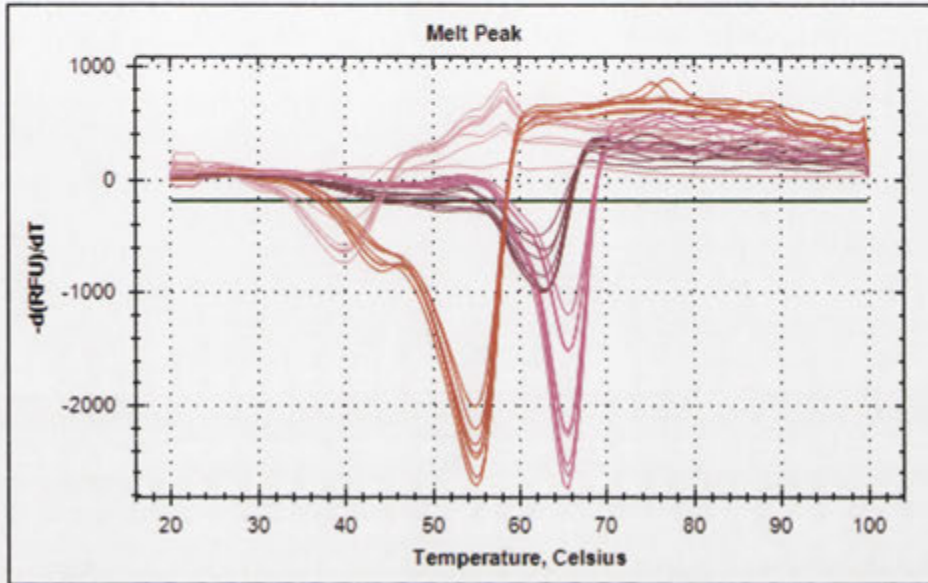
Lysine to alanine substitutions:

K42A – light pink (40C)

K163A – magenta (65.5C)

K296A – oxblood (62-63C)

WT – red (55C)

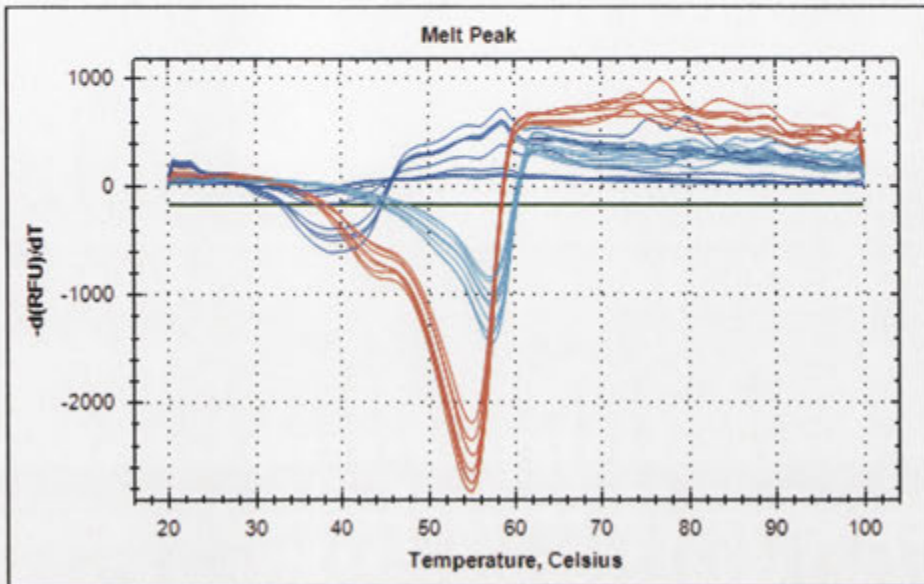


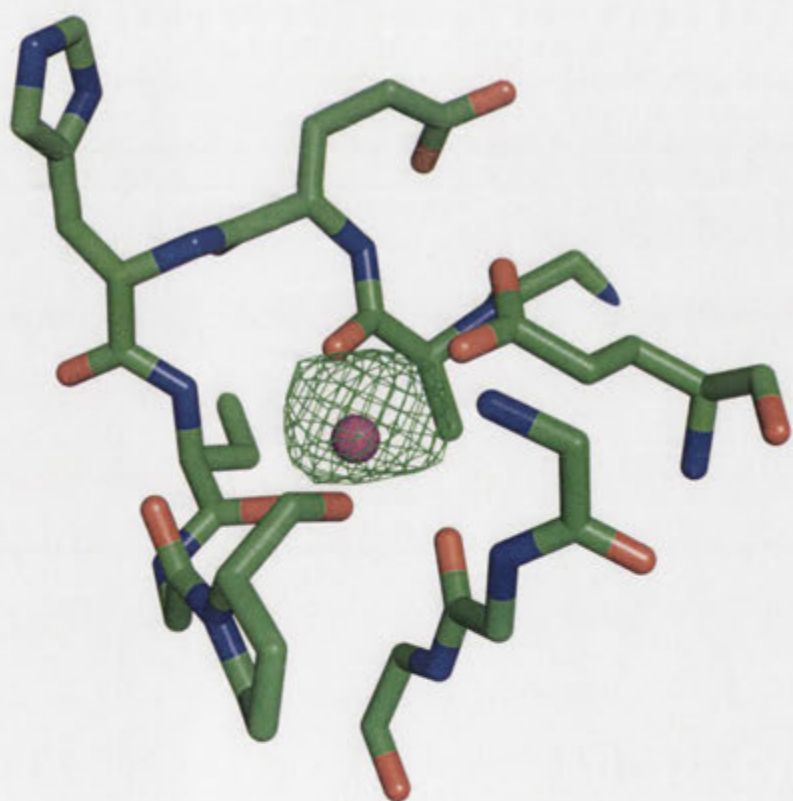
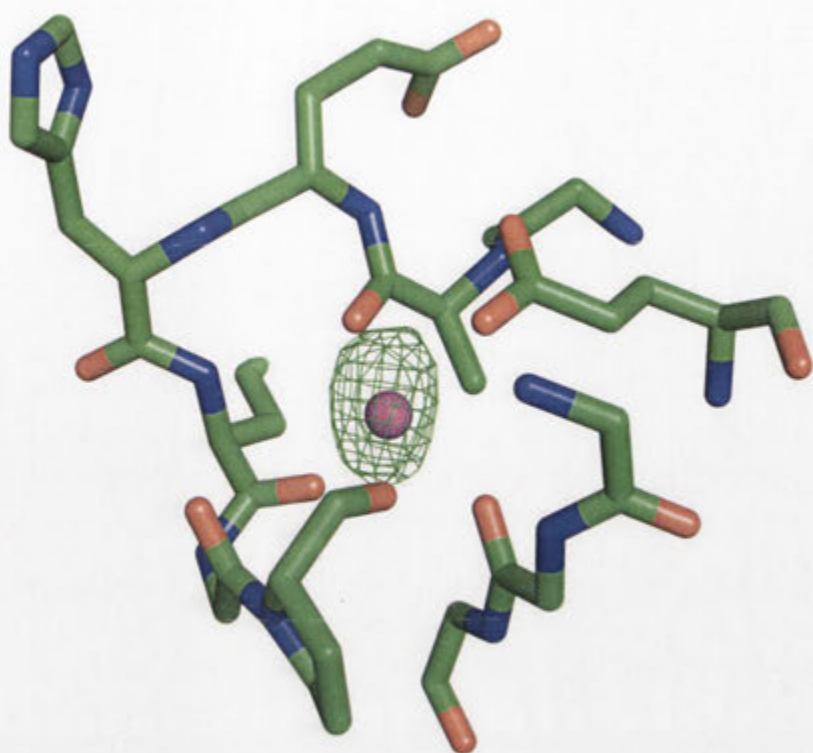
Lysine to arginine substitutions:

K163R – cyan (57C)

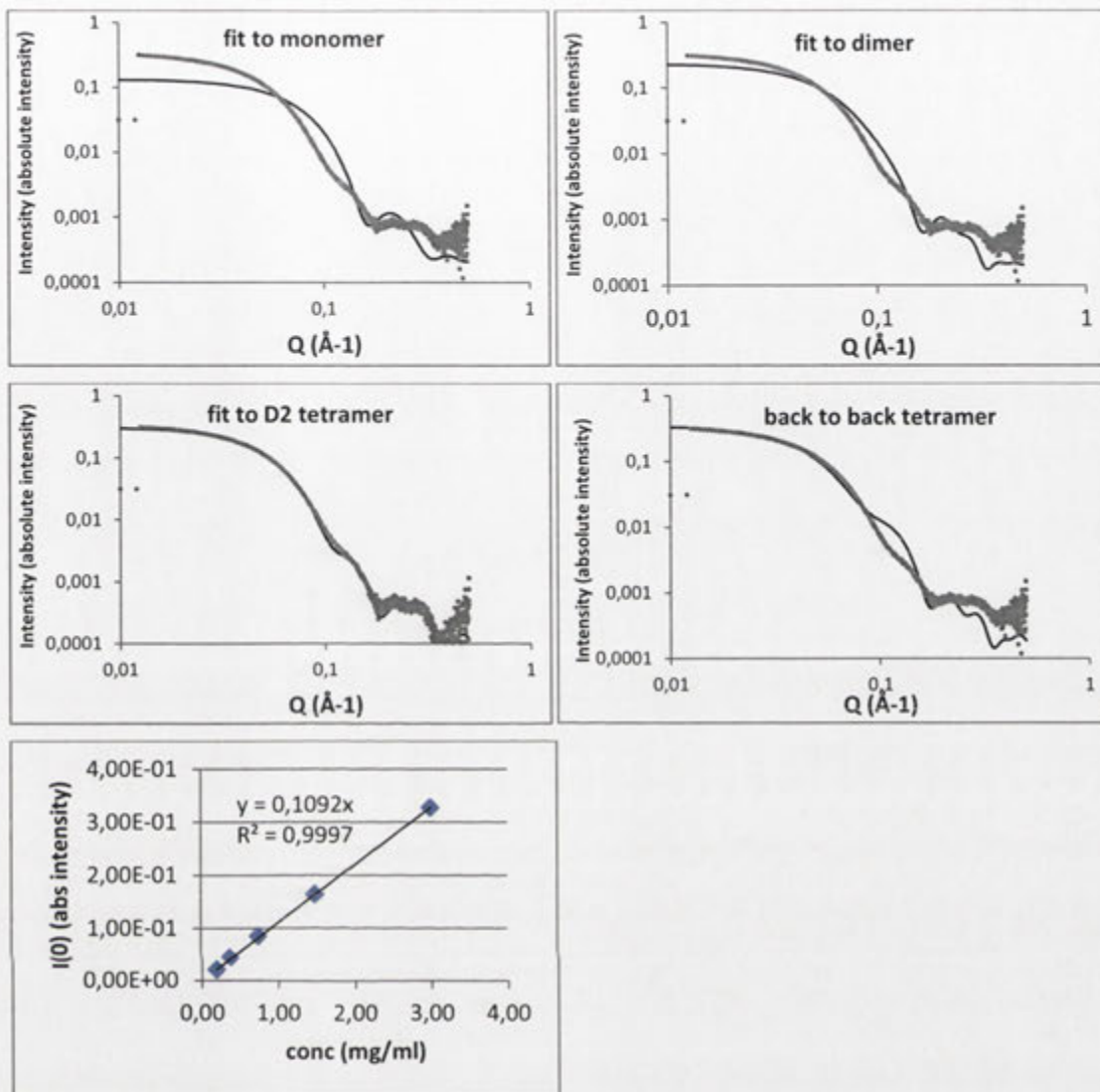
K296R – blue (38.5-39C)

WT – red (55C)

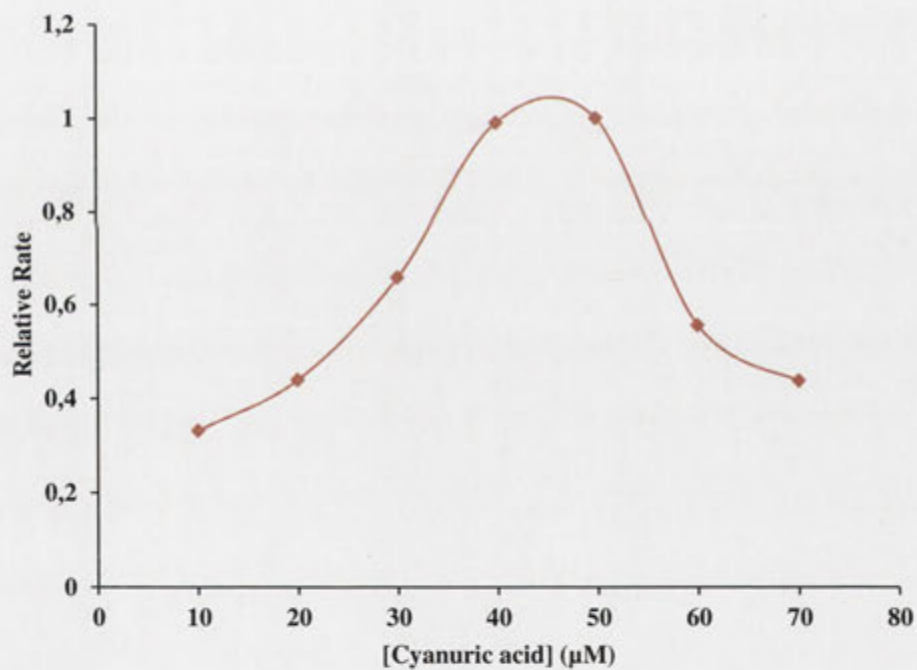
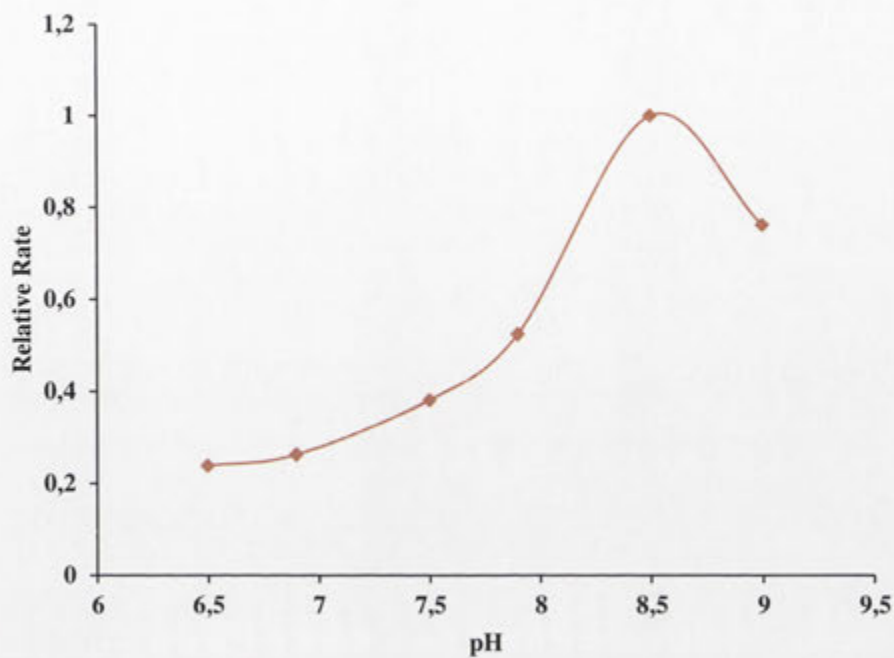




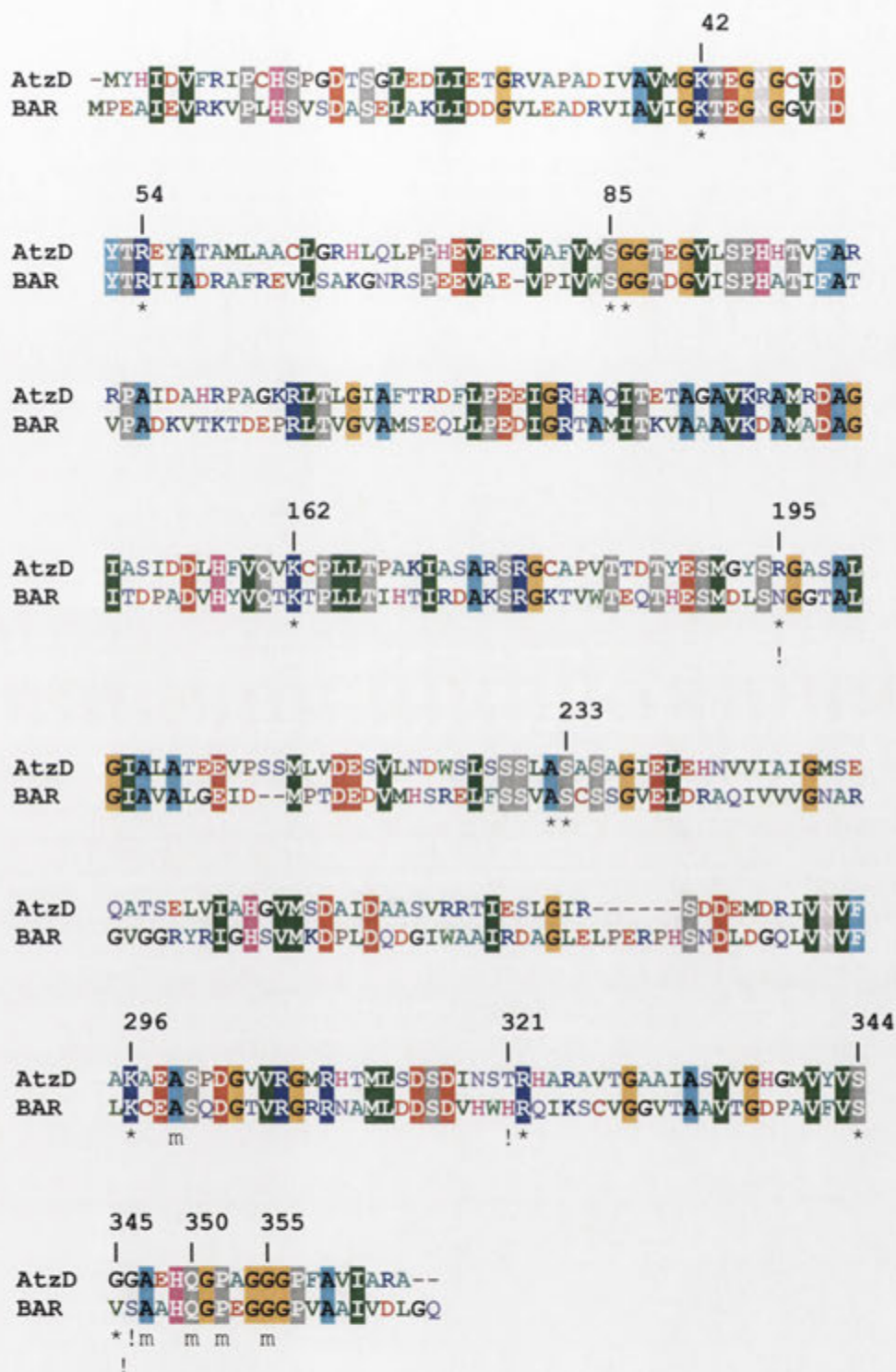
Supplemental Figure 1. Anomalous scattering for the two bound metals in the asymmetric unit. Top is the anomalous difference map (green mesh) set at 4.0 sigma at the metal site in protomer B of the asymmetric unit. Bottom is the anomalous difference map at 4.0 sigma at the metal site of protomer A. These sites were initially modelled as magnesium ions, but could also be sodium ions as sodium and magnesium are indistinguishable by this method. The anomalous difference maps were generated from data collected at 1.7149 Å (7230 eV) with 140 fold average redundancy. The orientation of both is approximately the same as that shown in Figure 3B.



Supplemental Figure 2. Panels 1-4 show CRYSOLE fits of different oligomeric states of the AtzD monomer, plotted against the measured SAXS data. Both the D2 and C4 (back to back) tetramers are found in the crystal, although only a dimer is found in the asymmetric unit. Panel 5 shows a plot of the total forward scatter vs. the concentration for the SAXS dilution series, showing an extremely good fit to the D2 tetramer with no variation across the concentration range.



Supplemental Figure 3. (Top) pH dependence of AtzD-dependent hydrolysis of cyanuric acid. (Bottom) Substrate inhibition of AtzD-dependent hydrolysis of cyanuric acid at pH 9.0. A relative rate of 1 is equivalent to the rate of AtzD (1µM) at pH8.5 at 50 µM cyanuric acid (30 µM.sec⁻¹).



Supplemental Figure 4. Alignment of AtzD and Bar. Where possible residues referred to in the results and discussion are numbered (above). Symbols below the sequence indicate the positions of active site residues (*), residues that appear to determine AtzD and Bar substrate specificities (!) and residues that form the metal binding site (m).



Supplemental Figure 5. Crystallization of native AtzD. Crystal of native AtzD in PEG 400. The crystal in the centre of the image is 0.1 mm x 0.1 mm

Treatment	No. spectra	No. Distinct peptides	% amino acid coverage	Total Protein Spectral Intensity	Intensity of peptide (R77)VAF...TVF(A)
Control	55	16	44	7.9×10^9	8 x spectra totaling 5.7×10^8
PMSF	14	10	34	1.7×10^9	1 spectrum at 5.0×10^6

Supplemental Table 1. Tryptic finger printing of AtzD covalently modified with PMSF. Peptide V78AF...TVF87 includes the putative active site serine residue (Ser85).

Supplemental Table 2. Plasmids and Primers. Site-directed mutations shown in red.

Plasmid/primer	Detail/sequence (5'-3')
pUC57: <i>atzD</i>	pUC57 containing the <i>atzD</i> gene based on Genbank Acc. No. U66917 (GenScript USA Inc.)
pETcc2	Expression vector based on pET14b (see methods for details)
pETCC2 Rev	TCATCGTCATCCTCGGCACCGTCA
K42A Fwd	ATCGTCGCGGTAATGGGCGCTACCGAGGGCAATGGCTG
K42A Rev	GCAGCCATTGCCCTCGGTAGCGCCCATTACCGCGACGA
K42R Fwd	CGCGGTAATGGGCAGAACCGAGGGCAATG
K42R Rev	CCATTGCCCTCGGTTCTGCCATTACCGC
S85A Fwd	TCGCGTTTGTGATGGCAGGTGGGACGGAAG
S85A Rev	TTCCGTCCCACCTGCCATCACAAACGCGAC
K162R Fwd	CATTTTGTGCAGGTGCGATGTCCGCTGCTGACAC
K162R Rev	TGTCAGCAGCGGACATCGCACCTGCACAAAATGC
K162A Fwd	TGCATTTTGTGCAGGTGGCATGTCCGCTGCTGACAC
K162A Rev	GGTGTGTCAGCAGCGGACATGCCACCTGCACAAAATGC
S233A Fwd	ATCGTCACTGGCGTCGGCGCAGCAGGCATCGAACTGGAG
S233A Rev	GCTCCAGTTCGATGCCTGCTGCCGCCGACGCCAGTGACGA
S344A Fwd	GGCATGGTGTATGTGGCAGGTGGCGCCGAGCATC
S344A Rev	ATGCTCGGCGCCACCTGCCACATACACCATGCCA
K296R Fwd	AACGTATTCGCCAGAGCGGAGGCGAG
K296R Rev	GCTCGCCTCCGCTCTGGCGAATACGT
K296A Fwd	TCAACGTATTCGCCGACGCGGAGGCGAGC
K296A Rev	GGCTCGCCTCCGCTGCCGGCGAATACGTTG

Chapter 6. X-ray structure of the amidase domain of AtzF, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex

Pages 116-139

Overview

The allophanate hydrolase, AtzF, from *Pseudomonas* sp. strain ADP carries the final step in s-triazine mineralization. The structure of amidase domain of AtzF was solved and compared with the recently solved structures of allophanate hydrolases from other organisms. Contrary to previous reports the C-terminus of AtzF does not appear to have a physiologically-relevant catalytic function (as reported for the allophanate hydrolase of *Kluyveromyces lactis*). However, one role of C-terminal domain was identified: coordinating the quaternary structure of AtzF. AtzF forms a large, *ca.* 660 kDa, multi-enzyme complex with other two enzymes of cyanuric acid mineralization pathway (AtzD and AtzE). The function of this complex may be to channel the unstable metabolites between active sites.

Contributions

Sahil Balotra: DNA cloning, mutagenesis, enzyme expression and purification, designed and performed experiments, data analysis, assay development and enzyme characterization. I co-authored the paper, contributed to writing and figure preparation.

Drs Janet Newman and Tom Peat: Crystallography and structural determination

Nigel G. French and Lyndall J. Briggs: Cloning and enzyme expression optimisation

Dr Nathan P. Cowieson: SAXS analysis

Dr Peter Campbell: Helped in peptide mass finger printing experiment

Dr Andrew Warden: Protein modelling

Dr Christopher J. Easton: Suggestions for substrate synthesis

Dr Colin Scott: Planning and supervision

This paper has been published in series of two manuscripts: Balotra *et al* (2014) *Acta Crystallographica F* **70**: 310-315; Balotra *et al* (2015) *Applied and Environmental Microbiology* **81**: 470-480

Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp. strain ADP

Sahil Balotra,^{a†} Janet Newman,^{b†} Nigel G. French,^a Lyndall J. Briggs,^a Thomas S. Peat^b and Colin Scott^{a*}

^aEcosystem Sciences, CSIRO, GPO Box 1700, Canberra, ACT 2601, Australia, and ^bMaterials, Science and Engineering, CSIRO, 343 Royal Parade, Parkville, VIC 3052, Australia

† These authors contributed to this work equally.

Correspondence e-mail: colin.scott@csiro.au

Received 12 November 2013
Accepted 27 December 2013

The allophanate hydrolase from *Pseudomonas* sp. strain ADP was expressed and purified, and a tryptic digest fragment was subsequently identified, expressed and purified. This 50 kDa construct retained amidase activity and was crystallized. The crystals diffracted to 2.5 Å resolution and adopted space group $P2_1$, with unit-cell parameters $a = 82.4$, $b = 179.2$, $c = 112.6$ Å, $\beta = 106.6^\circ$.

1. Introduction

Allophanate hydrolase (EC 3.5.1.54) from *Pseudomonas* sp. strain ADP (AtzF) is a member of the amidase family of enzymes that possess a conserved serine- and glycine-rich motif, the so-called 'amidase signature' sequence (Shafir *et al.*, 2005). The amidases are serine hydrolases that catalyze the hydrolysis of an amide group to a carboxylic acid with the concomitant release of ammonia. These enzymes adopt a classic α - β - α fold and possess a Ser-cisSer-Lys catalytic triad that is conserved throughout the family (Shin *et al.*, 2002). The catalytic serine in AtzF (Ser189) has been identified experimentally (Shafir *et al.*, 2005). AtzF contains 605 amino acids with an estimated molecular weight of 64 kDa, while the estimated native molecular weight has been shown to be approximately 260 kDa (Shafir *et al.*, 2005), suggesting that the native functional enzyme is a tetramer. The substrate range of allophanate hydrolase is quite narrow, and to date allophanate, malonamate and biuret are the only known substrates of this enzyme (Shafir *et al.*, 2005, 2006). The native substrate of AtzF, allophanate, is generated by two biological processes: cyanuric acid mineralization and urea decomposition (Shafir *et al.*, 2005; Kanamori *et al.*, 2004).

In most ureolytic organisms, the nickel-containing metalloenzyme urease (EC 3.5.1.5) provides a single-step hydrolysis of urea that yields ammonia and carbamate (Blakeley *et al.*, 1969), which spontaneously decays into ammonia and carbon dioxide (Carter *et al.*, 2009). However, there is an alternate urea metabolism pathway (Fig. 1), the urea carboxylate pathway, which is used by some algae and fungi (Strope *et al.*, 2011). In this pathway, urea carboxylase first catalyzes the ATP-dependent carboxylation of urea to form allophanate, which is subsequently deaminated by allophanate hydrolase, leading to the production of ammonia and carbon dioxide (Kanamori *et al.*, 2004; Maitz *et al.*, 1982; Roon & Levenberg, 1972).

Allophanate hydrolase also participates in the cyanuric acid mineralization pathway (Fig. 1), in which the cyanuric acid ring is hydrolytically opened by cyanuric acid hydrolase (AtzD or TrzD; EC 3.5.2.15) forming the unstable metabolite carboxybiuret, which spontaneously decarboxylates to form biuret (Peat *et al.*, 2013; Udiković-Kolić *et al.*, 2012). Allophanate is produced from biuret by AtzE (biuret hydrolase; EC 3.5.1.84) via a single deamination (Martinez *et al.*, 2001). Hydrolysis of allophanate is then carried out by allophanate hydrolase (Martinez *et al.*, 2001).

Here, we report the expression, purification, crystallization and initial X-ray diffraction analysis of the allophanate hydrolase (AtzF) from *Pseudomonas* sp. strain ADP.



2. Materials and methods

2.1. Cloning

The *atzF* gene was obtained from GenScript (New Jersey, USA); the coding sequence used was identical to that described by Martinez *et al.* (2001) (accession No. U66917) and was provided as an *NdeI/BamHI* insert in pUC57 (supplied by GenScript; Table 1). The *NdeI/BamHI* restriction fragment containing the *atzF* gene was subcloned into the *NdeI* and *BamHI* sites of pETCC2 (Peat *et al.*, 2013).

For the construction of the gene encoding C-terminally truncated AtzF (AtzF₄₆₇), primers were designed to amplify the region corresponding to the first 1404 bp of the *atzF* gene (AtzF Rev Trunc; Table 1). Phusion DNA polymerase and dNTPs were purchased from New England Biolabs (Ipswich, USA). AtzF Fwd A and AtzF Rev Trunc primers (Table 1) were used to amplify the targeted region of *atzF*. The reaction conditions for PCR were 1× Phusion HF Buffer, 2 pg pETCC2-*atzF* DNA template, 0.5 μM primers, 200 μM dNTPs and one unit of Phusion DNA polymerase in a total reaction volume of 50 μl. The cycle conditions for the PCR reaction were a 30 s denaturation step at 98°C followed by annealing at 53°C for 20 s and then extension for 120 s at 72°C for 30 cycles. The amplicon was separated on 0.6% agarose gel by electrophoresis to confirm the size of the amplicon. The DNA band was excised from the gel and was purified using a NucleoSpin Gel and a PCR Clean-up kit (Macherey-Nagel). The purified DNA was double digested with *BamHI*-HF and *NdeI* enzymes (NEB) before it was cloned into pETCC2 vector at *BamHI* and *NdeI* restriction sites using T4 DNA ligase (NEB) as per the manufacturer's instructions.

2.2. Expression and purification of AtzF and AtzF₄₆₇

pETCC2 derivatives containing either *atzF* or *atzF₄₆₇* were used to transform electrocompetent *Escherichia coli* BL21 (λDE3) cells (Invitrogen) and were grown at 310 K on Luria-Bertani (LB), Miller (Miller, 1972) agar plates supplemented with 100 μg ml⁻¹ ampicillin. Overnight starter cultures of 50 ml were inoculated with 4–5 colonies. The overnight cultures were diluted 1:20 into 950 ml LB, Miller medium and shaken at 310 K and 200 rev min⁻¹ until an OD₆₀₀ of

0.6–0.8 was attained. Protein expression was initiated by the addition of 100 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). The induced cultures were kept at 310 K overnight whilst shaking at 200 rev min⁻¹.

The cells were harvested by centrifugation at 4000g for 10 min in an Avanti J-E centrifuge (Beckman Coulter, Indianapolis, USA) and then resuspended in lysis buffer (50 mM HEPES, 100 mM NaCl pH 7.5 for AtzF and 50 mM Tris, 100 mM NaCl pH 7.5 for AtzF₄₆₇) and lysed by passage through an Avestin C3 homogenizer three times at 124 MPa. Insoluble cellular debris was removed by centrifugation at 21 000g using an Avanti J-E centrifuge.

Wild-type AtzF was purified from the soluble cell-free extract in three steps: His-tag affinity chromatography using an Ni-NTA Superflow cartridge (Qiagen, Maryland, USA) with a gradient of 0–300 mM imidazole in 50 mM HEPES, 100 mM NaCl pH 7.5, anion exchange using a HiTrap Q HP column (GE Healthcare Life Sciences, Uppsala, Sweden) and a gradient of 100–1000 mM NaCl in 50 mM HEPES pH 7.5, and finally size-exclusion chromatography using a 130 ml column packed with Superdex 200 prep-grade resin (GE Healthcare Life Sciences) with a buffer comprised of 50 mM HEPES, 100 mM NaCl pH 7.5. 20 column volumes of buffer were used to achieve the gradients used for the Ni-NTA Superflow cartridge and HiTrap Q HP purification steps. The protein was concentrated in an Amicon Ultra-15 Centrifugal Filter Unit with an Ultracel-30 membrane (Millipore, Carrigtwohill, Ireland) to 5 mg ml⁻¹ and snap-frozen in liquid nitrogen in 100 ml aliquots. The final purity was estimated to be 98% from a Coomassie-stained gel, and typical yields were 18–20 mg purified protein from 11 LB medium.

AtzF₄₆₇ was purified from soluble cell-free extract in two steps: binding to an Ni-NTA column followed by a wash step with 50 mM Tris, 100 mM NaCl pH 7.5 buffer and then by two wash steps with 10 mM imidazole and 20 mM imidazole in 50 mM Tris, 100 mM NaCl pH 7.5 buffer. Finally, the protein was eluted with 250 mM imidazole. The protein was further purified by size-exclusion chromatography using the same column as used for AtzF but in 50 mM Tris, 100 mM NaCl pH 7.5 buffer. The final purity was estimated to be 98% from a Coomassie-stained gel, and typical yields were 70–75 mg purified

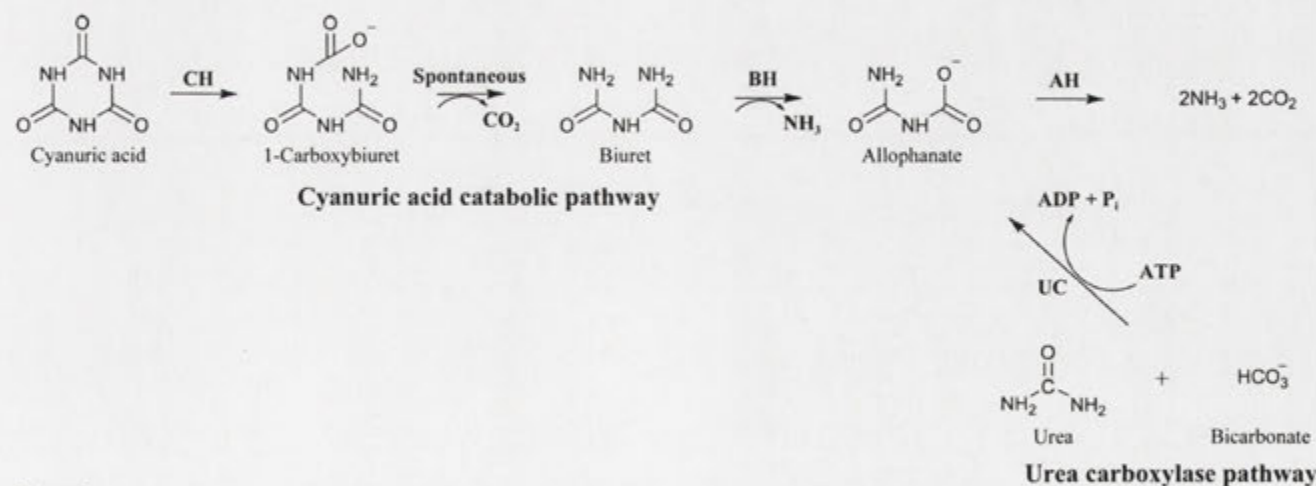


Figure 1

Catabolic processes involving allophanate hydrolase (AH). The cyanuric acid catabolic pathway involves ring opening of cyanuric acid by cyanuric acid hydrolase (CH), spontaneous decarboxylation of the product (1-carboxybiuret) and deamination of biuret by biuret hydrolase (BH) to produce allophanate. The urea catabolic pathway involves ATP-dependent urea carboxylation by urea carboxylase (UC) to produce allophanate. Both pathways then depend upon allophanate deamination by allophanate hydrolase to avoid spontaneous decarboxylation (and urea formation).

Table 1
DNA and protein used in this study.

The nucleotide and peptide sequences for the artificial His tag used for the purification of AtzF and AtzF₄₆₇ are shown in bold.

Source organism	<i>Pseudomonas</i> sp. strain ADP
DNA source	GenScript, based on accession No. U66917
Gene sequence encoding His-tagged full-length AtzF	atggcgacgacccatcatcatcatcaacagcgcgctgggcccgcggcgccat atgaatgaccgcgcgcccaacctgaaagatctgtcagtcacgccgatcaacctgaccgatctgctccatcaggtccctatgccgcggtaacagccgcgcgacgtcattcggacctgatgccgatcaaaagaacggcgaaatccgatctggattgacctgcccctggaaagcgcattggcgatgctggcgacgcgacgaacgaaggacaaggagaaagcgtggcctctggatccctcggcgtcaaggacaacatcagctcgcaggctcccgacgactggcggtgacgggttcgcgcgtacgcccgcacagcgcctctgctgacagcgcctgggacgctggcgccgatccgatcggaaaaacgaacctgatacaattccgaccgggtgaaacgcaactcgcacccggttggcattccgcgctggtgtcaacgagaactacgtatccggctccagcgtggcctccagcgtggcctccgacggcgacgttaccgllctcgcctgggacgcaactcgcgggttcgcgcgtatccctgctgcttcaacaactctggggcctgaaacgacaaagcctgttctcggcagtggaactgttcccgcggcggaagcctgactcactcagcctcgcaccatacctgatgacgccttgcggcgcacgcgctccgcgctacgatgatgacgcttttccgcaaggcggcgccgcacgacagaaaagagttggcctcgtcctcaattcgggtccacggcggaacatccgcaatlltccggacgcgaaagcggccttcaaaaagcggcttccgaaactgaaagagatgggtgacatcctgttgaclataccccctcaggcagcctgctgaactgctctacggcccttgggttgcgagcgcctggcgccatcgagaacttgcggacgagc-atcccaggtgctccaccggctgttctgacatcattctccggaagcgaatgagcagctegacgttcaacgttatctatcgcctggcacttgcacggctcagagagca-cttggaaaagatcgatgatgctgctgacggcgccgacacatcactgtagaagacatgctcgcgacatccgctacgctcaacagcaatcggcttctacagcaactctgta-acttgatggattgtccgalltctgtlcccgaggcttccgaaccaatggcctccatttggcctcaactcctggcggcgttccgacatcctcgctgcaatg
Gene sequence encoding His-tagged AtzF ₄₆₇	atggcgacgacccatcatcatcatcaacagcgcgctgggcccgcggcgccat atgaatgaccgcgcgcccaacctgaaagatctgtcagtcacgccgatcaacctgaccgatctgctccatcaggtccctatgccgcggtaacagccgcgcgacgtcattcggacctgatgccgatcaaaagaacggcgaaatccgatctggattgacctgcccctggaaagcgcattggcgatgctggcgacgcgacgaacgaaggacaaggagaaagcgtggcctctggatccctcggcgtcaaggacaacatcagctcgcaggctcccgacgactggcggtgacgggttcgcgcgtacgcccgcacagcgcctctgctgacagcgcctgggacgctggcgccgatccgatcggaaaaacgaacctgatacaattccgaccgggtgaaacgcaactcgcacccggttggcattccgcgctggtgtcaacgagaactacgtatccggctccagcgtggcctccgacggcgacgttaccgllctcgcctgggacgcaactcgcgggttcgcgcgtatccctgctgcttcaacaactctggggcctgaaacgacaaagcctgttctcggcagtggaactgttcccgcggcggaagcctgactcactcagcctcgcaccatacctgatgacgccttgcggcgcacgcgctccgcgctacgatgatgacgcttttccgcaaggcggcgccgcacgacagaaaagagttggcctcgtcctcaattcgggtccacggcggaacatccgcaatlltccggacgcgaaagcggccttcaaaaagcggcttccgaaactgaaagagatgggtgacatcctgttgaclataccccctcaggcagcctgctgaactgctctacggcccttgggttgcgagcgcctggcgccatcgagaacttgcggacgagc-atcccaggtgctccaccggctgttctgacatcattctccggaagcgaatgagcagctegacgttcaacgttatctatcgcctggcacttgcacggctcagagagca-cttggaaaagatcgatgatgctgctgacggcgccgacacatcactgtagaagacatgctcgcgacatccgctacgctcaacagcaatcggcttctacagcaactctgta-acttgatggattgtccgalltctgtlcccgaggcttccgaaccaatggcctccatttggcctcaactcctggcggcgttccgacatcctcgctgcaatg
Forward primer (AtzF Fwd A)	ATCATCACAGCAGCGCCTG
Reverse primer (AtzF Rev Trunc)	GTTTGGTGGATCCTCATTACTTGGCGAGGTCGTGCTCCACGAAAGC
Cloning vector	pUC57
Expression vector	pETCC2, a pET-14b derivative (Peat <i>et al.</i> , 2013)
Expression host	<i>E. coli</i> BL21 (ΔDE3), Invitrogen
Complete amino-acid sequence of AtzF	MGSSHHHHHSSGLVPRGSH mndrapbepersgrvtpdhldlasyaayaagtdaadvisdlyarikedgepniwisllplesalamladaqqrkdkgealplf-gi-pfgvkdnidvaglpittagctfartprqahfvvqrldvdaipigktnldqfatglntrtpfpgiprcvfnenysvsgssgsavavangtvpfsgtdtagsgripaafnlvlgkpt-kglfsglvpaaarsldcisviahvddalavarvaagydadafsrkagaaltekswprfnfvpaaehrfqgdaaealnkavrklemggtcisdytprqaaellyag-pwvaerlaaiesladehpevlhvvrrdiilskrmsavdftngjryladlvraaestwekidvmlptapiytvdmldpvrlnslngfytfnlmlsaiavpaftrnglp-gvtfgrafedgaiaslglkafvehdlakgnaataappkdtvaiaavghslsqplnhqltesgkkratrtapgyalyalrdatpakpmlrdqnavgsieviwdpvafgaf-vsepaplgitiledgshvkfbcpehaiaetaldithyggwraylaaq
Complete amino-acid sequence of AtzF ₄₆₇	MGSSHHHHHSSGLVPRGSH mndrapbepersgrvtpdhldlasyaayaagtdaadvisdlyarikedgepniwisllplesalamladaqqrkdkgealplf-gi-pfgvkdnidvaglpittagctfartprqahfvvqrldvdaipigktnldqfatglntrtpfpgiprcvfnenysvsgssgsavavangtvpfsgtdtagsgripaafnlvlgkpt-kglfsglvpaaarsldcisviahvddalavarvaagydadafsrkagaaltekswprfnfvpaaehrfqgdaaealnkavrklemggtcisdytprqaaellyag-pwvaerlaaiesladehpevlhvvrrdiilskrmsavdftngjryladlvraaestwekidvmlptapiytvdmldpvrlnslngfytfnlmlsaiavpaftrnglp-gvtfgrafedgaiaslglkafvehdlak

protein from 1 l LB medium. All protein chromatographic work was conducted using an ÄKTApurifier UPC 10 (GE Healthcare Life Sciences). The purified protein was concentrated to about 5 mg ml⁻¹ and snap-frozen (using liquid nitrogen) in 50–100 µl aliquots.

Both purified AtzF and AtzF₄₆₇ include 20-amino-acid N-terminal extensions that include a thrombin cleavage site and a hexahistidine tag (Table 1).

2.3. Proteolysis and N-terminal sequencing

Limited proteolysis of full-length AtzF (0.8 mg ml⁻¹) was carried out with 0.04 mg ml⁻¹ trypsin (MP Biomedicals, Australia) at 293 K for 48 h. Tryptic fragments were separated by SDS-PAGE (Shapiro *et al.*, 1967) using Tris-HEPES 10–20% precast gels (NuSep; Fig. 2). N-terminal sequencing of the proteolytic products was used to identify the position of the proteolysis site (Australian Proteome Analysis Facility, Sydney, Australia).

2.4. Activity test for full-length AtzF and AtzF₄₆₇

Biuret (Sigma–Aldrich, St Louis, USA) was used as a substrate to determine the activity of the wild-type and truncated enzymes. The hydrolytic production of ammonia from biuret by AtzF and truncated

AtzF was measured using an ammonia assay kit (Sigma–Aldrich). Briefly, L-glutamate dehydrogenase (GDH) reductively aminates α-ketoglutaric acid in an NADPH-dependent reaction (Bergmeyer, 1990). The consumption of NADPH results in a decrease in the UV absorbance at 340 nm. A SpectraMax M2 spectrophotometer (Molecular Devices, California, USA) was used to follow the decrease in absorbance in this assay. The assay was performed in 280 mM TAPS buffer pH 9; the concentrations and components of the reaction were 10 µg BSA, 0.5 mM NADPH, 10 mM α-ketoglutaric acid, 50 mM biuret, 191 pmol AtzF or AtzF₄₆₇, 0.875 U GDH.

An ammonia standard (150 µM) provided with the kit was used as a positive control. Enzyme-free negative controls contained 50 mM biuret in addition to the kit reagents and a substrate-free control contained 191 pmol AtzF or AtzF₄₆₇ along with the kit reagents.

2.5. Stability analysis

Heat-denaturation curves of the full-length AtzF protein, along with AtzF₄₆₇, were generated using differential scanning fluorimetry. The two samples were tested in a suite of different buffers and pHs in triplicate ('buffer screen 9'; Seabrook & Newman, 2013). The assay was performed in a CFX96 RT-PCR machine (Bio-Rad) with 19.6 µl

of each screening condition, 300 nl protein at 5–6 mg ml⁻¹ and 300 nl of a 1:10 (aqueous) dilution of SYPRO Orange dye (Sigma). Both proteins displayed curves indicative of a well folded protein: AtzF in the buffer used in the final chromatography step (50 mM HEPES, 100 mM NaCl pH 7.5) displayed a T_m of 50.4°C and AtzF₄₆₇ in 50 mM Tris pH 7.5, 100 mM NaCl displayed a melting temperature 2°C higher (Fig. 3)

2.6. Crystallization

All crystallization experiments were performed in 96-well SD-2 plates (IDEX, USA) against a 50 µl reservoir. Most often, the droplets consisted of 150 nl concentrated protein combined with 150 nl reservoir solution. Either a Phoenix (Art Robbins Instruments, USA) or a Mosquito (TTP LabTech, UK) robot was used to place the crystallization drops. Initial screening experiments with two commercial screens (JCSG+ and PACT; Newman *et al.*, 2005) and a 96-condition pH/salt gradient (PS gradient) screen at two temperatures, 20 and 8°C, showed no hits. The screening was extended to an 8 × 96 condition (768-condition) in-house screen (C3 screen; details of this and the PS gradient screen are available at <http://c6.csiro.au>), set up both in a conventional manner (droplets against a reservoir of crystallant solution) and against a reservoir of 1.5 M NaCl at 20°C. After 35 d, well shaped hexagonal crystals were observed in a drop equilibrated against 4 M sodium formate (Fig. 4). These crystals showed poor diffraction (6.5 Å maximum resolution) on the MX2 (microfocus) beamline of the Australian Synchrotron. These crystals could not be reproduced, even with seeding. *In situ* proteolysis screening experiments were set up using either trypsin or chymotrypsin (1:1000 ratio of protease:AtzF) in the JCSG+, PACT and PS

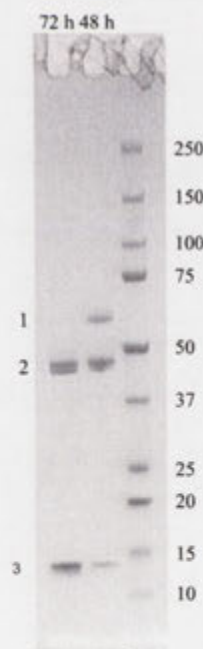


Figure 2
Fragmentation patterns from AtzF trypsinolysis. Tryptic digestion of AtzF resulted in the fragmentation of AtzF into two distinct bands on SDS-PAGE. The lane next to the ladder (right lane; labelled in kDa) shows a partial digest of AtzF 48 h after adding trypsin. The highest molecular weight band (labelled 1) represents undigested AtzF. The two bands adjacent to each other at nearly 48 kDa (labelled 2) possessed the same peptide sequence (as determined by N-terminal sequencing) and contain the N-terminus. The smallest visible band (labelled 3) is the 14 kDa C-terminal fragment of AtzF.

gradient screens at 20°C, and after a month some indications of crystallinity were observed in several ammonium sulfate-containing conditions of the PS gradient screen. The C3_6 screen, which focuses on ammonium sulfate-containing conditions, was set up at both 8 and 20°C with trypsin or chymotrypsin-treated protein, and an intergrown plate-shaped crystal was observed with the trypsin-treated protein in a condition consisting of 1 M ammonium sulfate, 1 M lithium sulfate, 0.1 M Tris chloride pH 8.5. Trypsin digests of the full-length protein, followed by mass-spectrometric analysis, revealed one major digestion site, resulting in a large N-terminal domain and a smaller C-terminal domain of molecular weights 48 and 14 kDa, respectively. Enzyme assays showed that the larger domain retained the amidase activity. An N-terminal His-tagged construct of the larger domain was produced, concentrated to 5.5 mg ml⁻¹ and set up against the JCSG+, PACT and PS gradient screens as above, with droplets consisting of 200 nl protein and 200 nl crystallant. The protein crystallized as intergrown plates from PACT condition B9 [0.2 M lithium chloride, 0.1 M MES pH 6.5, 20% (w/v) PEG 6K]. Crystals from this droplet were crushed and used to microseed further screening experiments, using in particular PEG-based screens, including PEG/Ion HT from Hampton Research. With seeding, the protein crystallized readily, producing a number of hits in PEG conditions. Crystals with a more robust morphology were found after two weeks in droplets from the PEG/Ion screen containing Tacsimate (Tacsimate is a mixture of organic acids produced by Hampton Research, which comprises 1.8305 M malonic acid, 0.25 M triammonium citrate, 0.12 M succinic acid, 0.3 M DL-malic acid, 0.4 M sodium acetate, 0.5 M sodium formate and 0.16 M diammonium tartrate) at pH 5 or lower and 12–16% (w/v) PEG 3350. Optimization around these conditions and microseeding with the improved crystals led to the production of optimized crystals in conditions consisting of 11–14% (w/v) PEG 3350, 2% Tacsimate pH 5 at 293 K (Fig. 4).

2.7. X-ray diffraction data collection

Optimized crystals were tested in a number of cryoprotectants on the MX2 beamline of the Australian Synchrotron. A crystal grown against a reservoir consisting of 11% (w/v) PEG 3350, 2% Tacsimate pH 5 was briefly introduced to a cryoprotectant consisting of 10% ethylene glycol, 10% glycerol, 80% reservoir solution and then flash-cooled in a stream of cold N₂ at the beamline. 360° of data (2 s exposure, 1° oscillation, wavelength of 0.9529 Å) were collected.

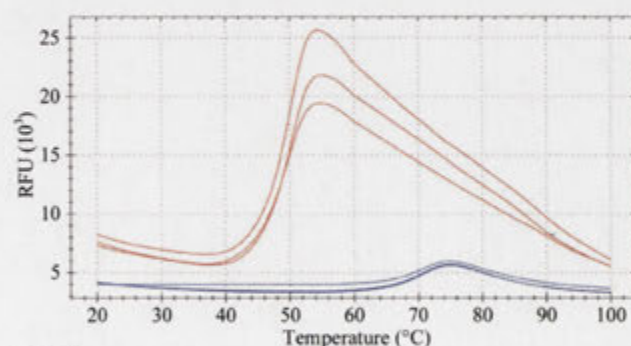


Figure 3
Melting curve. A temperature melting curve of AtzF was performed in triplicate, using differential scanning fluorimetry with the dye SYPRO Orange. The red curves are the AtzF protein (in triplicate) in the buffer used for crystallization (50 mM HEPES, 100 mM NaCl pH 7.5). The blue curves are a 0.1 mg ml⁻¹ lysozyme control.

crystallization communications

Data were indexed with *XDS* (Kabsch, 2010) and scaled using *SCALA* (Evans, 2006). Molecular replacement using *Phaser* (McCoy *et al.*, 2007) with PDB entry 2dqn (Nakamura *et al.*, 2006) as a search model revealed a clear solution with a log-likelihood gain of 757.

3. Results and discussion

AtzF from *Pseudomonas* sp. strain ADP was expressed heterologously in *E. coli* and purified to apparent homogeneity. The yield of purified AtzF varied from 18 to 20 mg per litre of culture. The specific activity of the purified enzyme was determined to be $15.4 \mu\text{mol min}^{-1} \text{pmol}^{-1}$ using 50 mM biuret as a substrate. Size-exclusion chromatography (Fig. 2) suggested that native AtzF was a homohexamer (~360 kDa), rather than a tetramer (~240 kDa) as suggested elsewhere (Shafir *et al.*, 2005).

Thermal stability studies of AtzF showed that the transition from the native to the unfolded state follows a first-order phase transition with a T_m of 48.3°C (Fig. 3); however, the protein was recalcitrant to crystallization (Fig. 4). *In situ* proteolysis crystallization trials resulted in the formation of needle-shaped crystallites (Fig. 4), which suggested that a fragment of the full-length enzyme produced by a tryptic digest may be more amenable to structural studies.

Scaled-up tryptic proteolysis of the full-length protein produced two major fragments of approximately 48 and 14 kDa (Fig. 2). N-terminal sequencing of the smaller 14 kDa band identified the location of the tryptic digestion as Lys467, whereas for the 48 kDa band the location of the tryptic cut was arginine at position 13. This suggested that larger 48 kDa fragment was the N-terminal fragment and the 14 kDa fragment was the C-terminal fragment. Moreover, the 48 kDa fragment was predicted by *pBLAST* to contain the entire amidase domain (*i.e.* the domain that is likely to be responsible for allophanate deamination).

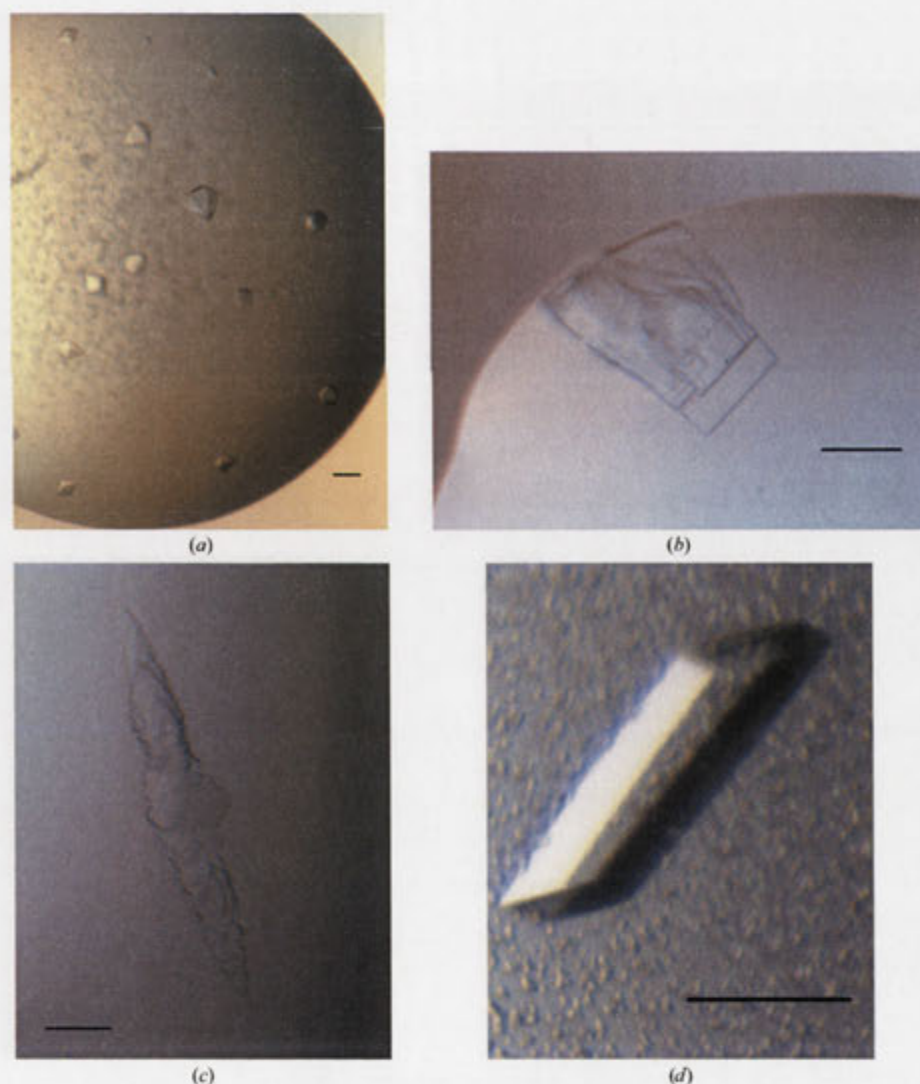


Figure 4

AtzF crystals. The scale bar represents 100 μm in each case. (a) Bipyramidal crystals of AtzF produced in 4 M sodium formate. These crystals diffracted X-rays to about 6.5 Å resolution and could not be reproduced. (b) Intergrown plates of trypsin-treated AtzF (*in situ* proteolysis) grown from 1 M ammonium sulfate, 1 M lithium sulfate, 0.1 M Tris chloride pH 8.5. (c) Intergrown plates of AtzF467 grown from 20% (w/v) PEG 6000, 0.1 M Na MES pH 6.5, 0.2 M calcium chloride. (d) Truncated AtzF crystals grown from 11% (w/v) PEG 3350, 2% Tacsimate pH 5 microseeded with the crystals shown in (c). These crystals diffracted X-rays to 2.5 Å resolution and were used for data collection.

A gene encoding the first 467 amino acids of AtzF (N-terminally histidine-tagged) was expressed and the product was purified to apparent homogeneity. The C-terminally truncated AtzF (AtzF₄₆₇) yielded 3.5–3.8 times more purified protein per litre of medium compared with full-length AtzF (*i.e.* 72–75 mg AtzF₄₆₇ per litre). The specific activity of AtzF₄₆₇ for biuret was 17.6 $\mu\text{mol min}^{-1} \text{pmol}^{-1}$ (using 50 mM biuret), demonstrating that it was as active as the full-length protein. However, unlike full-length AtzF, AtzF₄₆₇ was shown to be dimeric by size-exclusion chromatography (~ 110 kDa; data not shown), suggesting that the C-terminal 14 kDa fragment may play a role in organizing the oligomeric state of AtzF. Thermal melting analysis showed that AtzF₄₆₇ had a slightly higher T_m (50°C) than AtzF (Fig. 3).

The AtzF₄₆₇ construct produced large single crystals that were suitable for X-ray analysis in the presence of Tacsimate at pH 5 when seeded crystals were obtained from the Tacsimate-containing PEG/Ion screen. The crystals were of variable quality, despite looking similar, and tens of crystals were tested before finding one [grown in 11% (w/v) PEG 3350, 2% Tacsimate pH 5] that diffracted isotropically to 2.5 Å resolution, which was used for data collection (Fig. 4). The crystals adopted space group $P2_1$, with unit-cell parameters $a = 82.4$, $b = 179.2$, $c = 112.6$ Å, $\beta = 106.6^\circ$. In total, 86 96-well crystallization plates were set up to obtain the final well diffracting crystal.

While this manuscript was in preparation, the crystal structure of a full-length allophanate hydrolase from the ureolytic yeast *Kluyveromyces lactis* was published (Fan *et al.*, 2013). A comparison will be made between the published structure and that of AtzF in the near future.

We thank the beamline scientists of the Australian Synchrotron, and the CSIRO Collaborative Crystallization Centre (<http://www.csiro.au/C3>), Melbourne, Australia.

References

- Bergmeyer, H. U. (1990). Editor. *Methods of Enzymatic Analysis*, 3rd ed. Weinheim: VCH.
- Blakeley, R. L., Hinds, J. A., Kunze, H. E., Webb, E. C. & Zerner, B. (1969). *Biochemistry*, **8**, 1991–2000.
- Carter, E. L., Flugga, N., Boer, J. L., Mulrooney, S. B. & Hausinger, R. P. (2009). *Metallomics*, **1**, 207–221.
- Evans, P. (2006). *Acta Cryst. D* **62**, 72–82.
- Fan, C., Li, Z., Yin, H. & Xiang, S. (2013). *J. Biol. Chem.* **288**, 21422–21432.
- Kabsch, W. (2010). *Acta Cryst. D* **66**, 125–132.
- Kanamori, T., Kanou, N., Atomi, H. & Imanaka, T. (2004). *J. Bacteriol.* **186**, 2532–2539.
- Maitz, G. S., Haas, E. M. & Castric, P. A. (1982). *Biochem. Biophys. Acta*, **714**, 486–491.
- Martinez, B., Tomkins, J., Wackett, L. P., Wing, R. & Sadowsky, M. J. (2001). *J. Bacteriol.* **183**, 5684–5697.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Miller, J. H. (1972). *Experiments in Molecular Genetics*. New York: Cold Spring Harbor Laboratory Press.
- Nakamura, A., Yao, M., Chinnaronk, S., Sakai, N. & Tanaka, I. (2006). *Science*, **312**, 1954–1958.
- Newman, J., Egan, D., Walter, T. S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J. L., Stuart, D. I. & Perrakis, A. (2005). *Acta Cryst. D* **61**, 1426–1431.
- Peat, T. S., Balotra, S., Wilding, M., French, N. G., Briggs, L. J., Panjikar, S., Cowieson, N., Newman, J. & Scott, C. (2013). *Mol. Microbiol.* **88**, 1149–1163.
- Roon, R. J. & Levenberg, B. (1972). *J. Biol. Chem.* **247**, 4107–4113.
- Seabrook, S. A. & Newman, J. (2013). *ACS Comb. Sci.* **15**, 387–392.
- Shafir, N., Cheng, G., Sadowsky, M. J. & Wackett, L. P. (2006). *Appl. Environ. Microbiol.* **72**, 2491–2495.
- Shafir, N., Sadowsky, M. J. & Wackett, L. P. (2005). *J. Bacteriol.* **187**, 3731–3738.
- Shapiro, A. L., Viñuela, E. & Maizel, J. V. (1967). *Biochem. Biophys. Res. Commun.* **28**, 815–820.
- Shin, S., Lee, T.-H., Ha, N.-C., Koo, H. M., Kim, S., Lee, H.-S., Kim, Y. S. & Oh, B.-H. (2002). *EMBO J.* **21**, 2509–2516.
- Strope, P. K., Nickerson, K. W., Harris, S. D. & Moriyama, E. N. (2011). *BMC Evol. Biol.* **11**, 80.
- Uđiković-Kolić, N., Scott, C. & Martin-Laurent, F. (2012). *Appl. Microbiol. Biotechnol.* **96**, 1175–1189.

X-Ray Structure of the Amidase Domain of AtzF, the Allophanate Hydrolase from the Cyanuric Acid-Mineralizing Multienzyme Complex

Sahil Balotra,^{a,b} Janet Newman,^c Nathan P. Cowieson,^d Nigel G. French,^a Peter M. Campbell,^a Lyndall J. Briggs,^a Andrew C. Warden,^a Christopher J. Easton,^b Thomas S. Peat,^c Colin Scott^a

CSIRO Land and Water Flagship, Black Mountain, Canberra, ACT, Australia^a; Research School of Chemistry, Australian National University, Canberra, ACT, Australia^b; CSIRO Biomedical Manufacturing Program, Parkville, VIC, Australia^c; Australian Synchrotron, Clayton, VIC, Australia^d

The activity of the allophanate hydrolase from *Pseudomonas* sp. strain ADP, AtzF, provides the final hydrolytic step for the mineralization of *s*-triazines, such as atrazine and cyanuric acid. Indeed, the action of AtzF provides metabolic access to two of the three nitrogens in each triazine ring. The X-ray structure of the N-terminal amidase domain of AtzF reveals that it is highly homologous to allophanate hydrolases involved in a different catabolic process in other organisms (i.e., the mineralization of urea). The smaller C-terminal domain does not appear to have a physiologically relevant catalytic function, as reported for the allophanate hydrolase of *Kluyveromyces lactis*, when purified enzyme was tested *in vitro*. However, the C-terminal domain does have a function in coordinating the quaternary structure of AtzF. Interestingly, we also show that AtzF forms a large, ca. 660-kDa, multienzyme complex with AtzD and AtzE that is capable of mineralizing cyanuric acid. The function of this complex may be to channel substrates from one active site to the next, effectively protecting unstable metabolites, such as allophanate, from solvent-mediated decarboxylation to a dead-end metabolic product.

Atrazine (1-chloro-3-ethylamino-5-isopropylamino-2,4,6-triazine; Fig. 1) is one of the most heavily applied herbicides in the world and is registered for use in North and South America, Australia, Africa, Asia, and the Middle East. Atrazine is environmentally persistent (half-life, 4 to 57 weeks, depending on the location) and mobile, leading to the detection of atrazine in surface water, groundwater, and aquifers (1–3). Atrazine has been detected in the environment at concentrations of up to 4.6 μM in several countries (2, 3). It has been suggested that atrazine may be a carcinogen and an endocrine disrupter at such concentrations (4–6).

Since atrazine was introduced in the 1950s, bacteria have evolved highly efficient catabolic pathways that allow the use of atrazine as a sole nitrogen and carbon source (7–10). These pathways have provided valuable insights into the evolutionary processes that drive the establishment of new enzyme function and new catabolic pathways (11–15). In addition, these pathways and cognate enzymes provide a potential biotechnological solution to atrazine contamination (i.e., bioremediation) (16–19).

The most intensively studied atrazine catabolism pathway was discovered in *Pseudomonas* sp. strain ADP in the mid-1990s and is comprised of six hydrolases: atrazine chlorohydrolase (AtzA; EC 3.8.1.8) (20, 21), *N*-ethylaminohydrolase (AtzB; EC 3.5.99.3) (22, 23), *N*-isopropylammelide isopropylaminohydrolase (AtzC; EC 3.5.99.4) (24, 25), cyanuric acid amidohydrolase (AtzD; EC 3.5.2.15) (15, 26, 27), biuret amidohydrolase (AtzE; EC 3.5.1.84) (28), and allophanate hydrolase (AtzF; EC 3.5.1.54) (29–31). These hydrolases sequentially dechlorinate (AtzA) and remove the two *N*-alkyl side groups (AtzB and AtzC) to produce cyanuric acid, which is then further hydrolyzed to biuret, allophanate, and ammonia via AtzD, AtzE, and AtzF, respectively (Fig. 1).

In other bacterial genera, such as *Arthrobacter* and *Nocardioideis*, the function of AtzA is conducted by TrzN (32–36). Although TrzN is physiologically analogous to AtzA, it appears to

have evolved independently, as the two enzymes share low sequence identity and have significantly different reaction mechanisms and substrate ranges.

Recently, there have been concerted efforts to understand the structural biology of the hydrolases of the atrazine catabolic pathway, with the structure of TrzN being obtained in 2010 (33) and the structure of AtzD (15), which possesses a previously unreported Toblerone protein fold, being reported in 2013. A structure for AtzC has also been deposited in the Protein Data Bank (PDB accession number 2QT3). Structures for AtzB, AtzE, and AtzF have yet to be reported, although structures for the ureolytic allophanate hydrolases (AHs) from *Kluyveromyces lactis* (AH_{Kl}) and *Granulibacter bethesdaensis* (AH_{Gb}) were reported in 2013 (37, 38). Unlike the atrazine-degrading AHs, the ureolytic enzymes are found as multidomain enzymes along with a biotin- and ATP-dependent urea carboxylase, which is required to generate allophanate from urea (37, 39–42). The AH components of these complexes have considerable sequence conservation with each other and the triazine-related AHs (Fig. 2).

Received 24 August 2014 Accepted 18 October 2014

Accepted manuscript posted online 31 October 2014

Citation 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. doi:10.1128/AEM.02783-14.

Editor: R. E. Parales

Address correspondence to Colin Scott, colin.scott@csiro.au.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.02783-14>.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AEM.02783-14

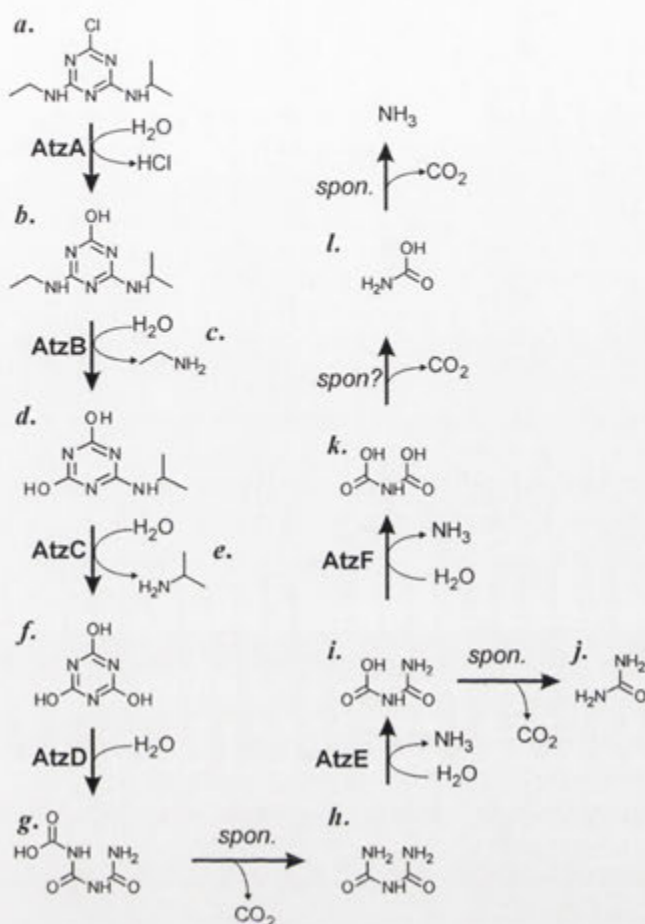


FIG 1 Mineralization of atrazine in *Pseudomonas* sp. strain ADP. Atrazine (a) is hydrolyzed by AtzA, releasing chloride. The product (b) is further hydrolyzed by AtzB, releasing ethylamine (c), and subsequent hydrolysis of the resulting product (d) releases isopropylamine (e), generating cyanuric acid (f). AtzD hydrolyzes cyanuric acid to produce the unstable 1-carboxybiuret (g), which decomposes to biuret (h). Biuret is deaminated by AtzE to yield allophanate (i), which can spontaneously decompose to urea (j) or can be deaminated by AtzF to produce the unstable *N*-carboxycarbamate (k), which spontaneously decomposes first to carbamate (l) and then to ammonia. spon., spontaneous.

AH is a member of the amidase signature family, characterized by an ~130-amino-acid-long region that is rich in serine and glycine residues and contains a characteristic Ser–*cis*–Ser–Lys catalytic triad (43, 44). Uniquely among this family of amidases, the AHs possess a conserved ca. 15-kDa extension at the C terminus of the amidase domain, of uncertain function, that is present in both eukaryotic and prokaryotic AHs (38). In a report from Fan et al., the authors proposed that the C-terminal domain of AH_{K7} is involved in the hydrolysis of *N*-carboxycarbamate, the unstable product of AH-mediated allophanate deamination (Fig. 1) (37). Moreover, *in silico* substrate docking suggested that a C-terminal histidine residue (His492) plays an important catalytic role (37).

Herein, we present the X-ray structure of the N-terminal amidase domain of AtzF from *Pseudomonas* sp. strain ADP and compare it with the recently determined AH structures. Our *in vitro* biochemical data were unable to support a physiologically relevant catalytic function for the C-terminal domain of AtzF. How-

ever, we present evidence for a functional role for the ~15-kDa C-terminal domain in coordinating the quaternary structure of the protein; AtzF forms a large multienzyme complex with AtzD and AtzE that mineralizes cyanuric acid. The function of this complex may be to protect unstable metabolites, such as allophanate, from solvent-mediated decarboxylation to a dead-end metabolic product.

MATERIALS AND METHODS

DNA manipulation. The cloning of *atzF* and the truncated gene encoding the amidase (N-terminal) domain of AtzF (*atzF*₄₆₇) is described elsewhere (45). A mutant of *atzF* that encoded the AtzF H488A variant was produced by the overlapping PCR method of Ho et al. (46), using *atzF* as the template DNA and primers *atzF* H488A fwd (5'-ACCAGCCCTTGAATGCTCAGCTCACGGAG-3') and *atzF* H488A rev (5'-TCTCCGTGAGCTGAGCATTCAAGGGCTGG-3'). PCRs with Phusion DNA polymerase (New England Biolabs [NEB], Ipswich, MA, USA) were conducted under the following conditions: 30 s of denaturation (98°C) and 30 s of annealing (55°C), followed by a 2-min extension (72°C).

The *atzF* H488A amplicon was digested with NdeI and BamHI (NEB, Ipswich, MA, USA) and cloned into the pETCC2 expression vector (15). The resultant plasmid was used to transform electrocompetent *Escherichia coli* BL21 λ(DE3) cells (Invitrogen, CA) following the manufacturer's instructions.

Protein expression and purification. Expression, postexpression harvest, and lysis of cells for the AtzF H488A variant were carried out under the same conditions used for AtzF and AtzF₄₆₇ expression reported elsewhere (45). Purification of AtzF H488A from soluble cell extract was carried out by metal ion affinity chromatography using an Ni-nitrilotriacetic acid Superflow cartridge (Qiagen, MD). After loading the soluble fraction onto the column, a step gradient of 12.5 mM imidazole was applied for six column volumes, followed by a seven-column-volume wash with 50 mM imidazole, after which protein was eluted in seven column volumes of 250 mM imidazole.

Crystallization and structure solution. Crystallization and data collection were done as previously described (45). Briefly, crystals grew from a reservoir containing 11 to 14% (wt/vol) polyethylene glycol 3350 and 2% Tacsimate reagent (pH 5) at 293 K and were used to collect X-ray data at the MX-2 beamline of the Australian Synchrotron using a wavelength of 0.9529 Å (13,011 eV). The structure was solved using molecular replacement (Phaser software) (47) of the structure with PDB accession number 2DQN (48), and a clear solution with six protomers in the asymmetric unit was found. The space group was found to be P2₁, and the resolution of the data extended to 2.5 Å (Table 1). The model was rebuilt by hand using Coot software (49) and refined using the Refmac program (50), with the noncrystallographic symmetry (NCS) restraints option being used during all stages of refinement. Four of the molecules (chains A to D) had a significantly better electron density than the other two molecules (chains E and F).

The crystals were grown in the Tacsimate reagent (Hampton Research), which is predominantly malonic acid, and there was an unambiguous density for malonic acid in the four protomers with a good electron density. Despite the poor density for two of the molecules in the asymmetric unit, the model refined to *R*_{work} and *R*_{free} values of 22.4% and 25.9%, respectively (Table 1). The Ramachandran plot (from Coot software) shows that 94.3% of the residues were in the most favorable region, 4.3% were in the allowed region, and were 1.5% in the outlier region.

SAXS. AtzF and AtzF₄₆₇ were dialyzed overnight into 50 mM Tris buffer, pH 7.5, 100 mM NaCl. The same buffer used as the buffer standard during data collection was used. A dilution series of AtzF (from 6.8 to 0.2 mg/ml) and AtzF₄₆₇ (from 13.1 to 0.4 mg/ml) was prepared, and scattering data were collected for 1 s using a Pilatus 1 M photon-counting detector (Dektris) with a sample-to-detector distance of 1.6 m. Ten replicate images were collected for each sample and averaged, with outlier rejection, to control for radiation damage. Data were measured in a Q (scat-

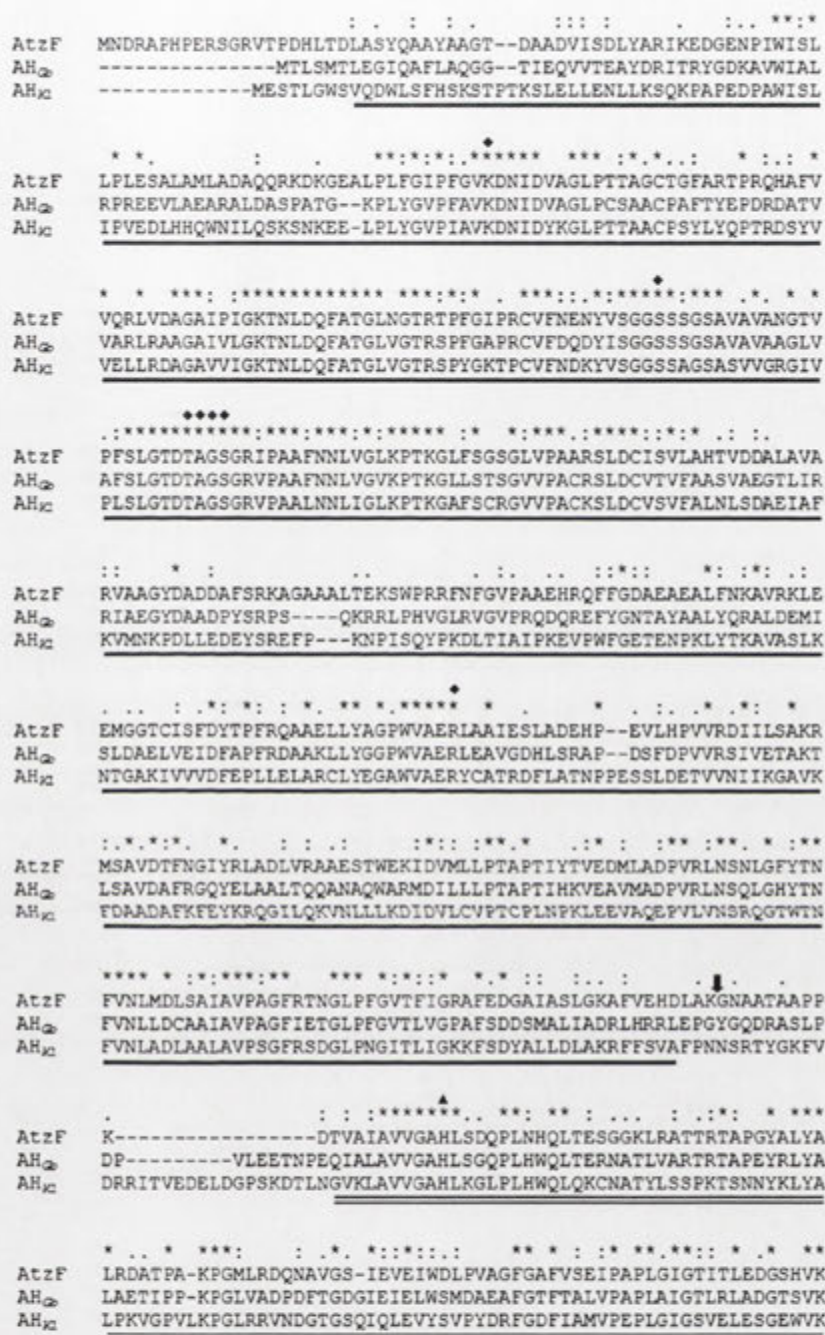


FIG 2 Alignment of the allophanate hydrolases from *Pseudomonas* sp. strain ADP (AtzF), *Kluyveromyces lactis* (AH₃₂), and *Granulibacter bethedensis* (AH₂₆). Single underline, amidase domain; double underline, C-terminal domain; diamonds, active-site residues of the amidase domain; triangle, proposed catalytic histidine residue in the C-terminal domain; arrow, position where AtzF was truncated to produce AtzF468; asterisks, identical residues; colons, highly similar residues; periods, similar.

tering vector) range from 0.01 to 0.5 Å⁻¹, and at the highest protein concentration, the scattering remained above the noise threshold to the edge of the detector.

AtzF (75 μl, 6 mg/ml) and AtzF₄₆₇ (75 μl, 6 mg/ml) were injected onto a size-exclusion column (Wyatt silica bead column; pore size, 300 Å) that had been preequilibrated with the PO₄-NaCl buffer. The column was developed at 0.2 ml/min, and a single peak eluted. The small-angle X-ray scattering (SAXS) showed no change over the peak (unpublished observations).

The radius of gyration (R_g) and total forward scatter [$I(0)$] for each concentration in the dilution series was calculated using the autorg routine in the PRIMUS system (51). Fits to atomic models were performed using the CRYSOLO program (52), and dummy atom models were calculated using the DAMMIF program (53) and superimposed on the high-resolution models using the SUPCOMB program (54). Molecular weight was calculated from $I(0)$ (55), together with protein concentration measurements and the scattering length density and partial specific volume calculated from the protein sequences, using the web application MULCh (56).

TABLE 1 Data collection and refinement statistics

Parameter ^a	Value for structure with PDB accession no. 4CP8 ^b
Data collection statistics	
Space group	P2 ₁
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	82.45, 179.23, 112.61
α , β , γ (°)	90, 106.63, 90
Resolution (Å)	2.50 (2.64–2.50)
<i>R</i> _{merge}	0.150 (0.810)
<i>I</i> / σ <i>I</i>	12.6 (2.8)
Completeness (%)	100 (100)
Redundancy	7.7 (7.7)
Refinement statistics	
Resolution (Å)	40.0–2.50
No. of reflections	102, 548
<i>R</i> _{work} / <i>R</i> _{free} (%)	22.4/25.9
No. of:	
Atoms	20, 462
Proteins	20, 132
Ligands/ions	28
Water molecules	312
<i>B</i> -factors (Å ²) for:	
Proteins	44.2
Ligands/ions	30.3
Water	24.2
RMSD	
Bond lengths (Å)	0.010
Bond angles (°)	1.367

^a *R*_{merge} = $\sum |I - \langle I \rangle| / \sum I \times 100$, where *I* is the intensity of a reflection and $\langle I \rangle$ is the mean intensity.

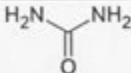

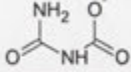
^b Values in parentheses are for the highest-resolution shell.

Substrate preparation and kinetic assays. The potassium salt of allophanate was prepared by hydrolyzing ethyl allophanate (Sigma-Aldrich, St. Louis, MO) with a 5 M excess of 1 M potassium hydroxide at 40°C for 3 h (42). Potassium allophanate was precipitated overnight in an ice-cold mixture of 5 volumes of ethanol and 1 volume of diethyl ether. The precipitate was separated by filtration and subsequently dried and stored in a desiccator.

The synthesis of allophanate was confirmed by ¹³C nuclear magnetic resonance (NMR) (see Fig. S1 in the supplemental material) using a Varian Inova-500 NMR spectrometer. The substrate was dissolved in 1 M KOH prepared in D₂O, and 2,048 scans were performed in total. The purity of the potassium allophanate produced was assessed using a glutamate dehydrogenase (GDH)-coupled assay. Complete AtzF-mediated hydrolysis of allophanate was performed, and the GDH-based ammonia assay was used to determine the purity of the allophanate prepared. The GDH-dependent decrease in the absorbance at 340 nm was measured using a SpectraMax M2 spectrophotometer (Molecular Devices, CA) to follow the consumption of NADH during the reductive amination of α -ketoglutaric acid to form L-glutamate. The hydrolysis of allophanate by AtzF was assumed to have reached completion, as a high concentration of AtzF (~1 μ M) was added, the reaction components of the GDH-linked ammonia assay were added well in excess of the allophanate concentration, and the decrease in NADH absorbance plateaued before all of the NADH was consumed. Allophanate prepared by the method described above was found to be 86% pure (Table 2).

To study the pH dependence of AtzF activity, kinetic assays were performed in the pH range of 7 to 9.5 using allophanate as a substrate, and *k*_{cat}/*K*_m values were plotted against the pH (see Fig. S3 in the supplemental material). The kinetic assays were carried out in buffer comprised of 100 mM HEPES and 100 mM CHES (*N*-cyclohexyl-2-aminoethanesulfonic

TABLE 2 ¹³C NMR shifts of allophanate and possible impurities

Substrate	Structure	¹³ C NMR shift
Urea		162.89 to 163.45
Bicarbonate		165.23 to 169.10
Allophanate		159.36 to 158.87, 158.37 to 157.97

acid) adjusted to the desired pH. The combined buffer system was used to keep the conditions similar throughout the wide pH range used in these assays. Kinetic assays were performed for AtzF, AtzF₄₆₇, and AtzF H488A using allophanate as the substrate at pH values ranging from 7.0 to 9.0 and temperatures of 28°C and 4°C. To account for the temperature-dependent decrease in AtzF activity, the reaction mixtures in assays conducted at 28°C contained 27 nM enzyme, and the reaction mixtures in assays carried out at 4°C contained 53 nM enzyme. In addition to the enzyme and substrate, each reaction mixture was comprised of 0.5 mM NADH (Sigma-Aldrich), 10 mM α -ketoglutaric acid (Sigma-Aldrich), 1.4 μ M bovine serum albumin (Sigma-Aldrich), and 12 μ M glutamate dehydrogenase (Sigma-Aldrich). *V*_{max} and the Michaelis-Menten constant were obtained by nonlinear regression analysis and applying a robust fit in GraphPad Prism (version 6) software (GraphPad Software, Inc., CA).

The substrate (allophanate) and product (*N*-carboxycarbamate) were also analyzed by liquid chromatography-mass spectrometry (LC-MS). Enzyme reaction mixtures were comprised of 53.5 mM potassium allophanate in 100 mM ammonium acetate buffer, pH 9.0. Reaction mixtures contained one of the following: no enzyme, 0.48 μ M AtzF, 0.48 μ M AtzF₄₆₈, or 0.96 μ M AtzF₄₆₈. The reactions were carried out on ice. LC-MS analyses were carried out on an Agilent 6550 iFunnel quadrupole time of flight LC-MS system, with samples being introduced by direct injection. The mobile phase was comprised of 10 mM ammonium acetate, pH 9.0, and acetonitrile. Samples were analyzed in negative ion mode, and *m/z* values ranging from 50 to 150 were scanned.

The effect of pH on enzyme stability was determined by obtaining residual activities after heating AtzF, AtzF₄₆₇, and AtzF H488A in either pH 7 or pH 9 buffer for 5 min in an Eppendorf Mastercycler EP apparatus (Eppendorf, Hamburg, Germany). The reaction components and method were the same as those used in the kinetic assays, and 2.2 pmol of each protein was used in 100 μ l of reaction mixture.

Identification of cyanuric acid mineralization complex. A culture of *Pseudomonas* sp. strain ADP was grown at 28°C with cyanuric acid as a sole nitrogen source. The composition of the medium used is as follows: 26.1 mM Na₂HPO₄·7H₂O, 22 mM KH₂PO₄, 8.5 mM NaCl, 200 μ M MgSO₄, 2.9 mM sucrose, 3.4 mM trisodium citrate, 4.4 μ M CaCl₂, 10 mM cyanuric acid, 20 ml vitamin stock, and 1 ml of 1,000 \times stock solution of trace elements per liter of culture medium. The 1,000 \times trace elements solution comprised 34.8 μ M ZnSO₄·7H₂O, 15.2 μ M MnCl₂·4H₂O, 485 μ M H₃BO₃, 84 μ M CoCl₂·2H₂O, 5.9 μ M CuCl₂·2H₂O, 12.4 μ M Na₂MoO₄·2H₂O, and 8.4 μ M NiCl₂·6H₂O, and the pH of the medium was adjusted to 7.3 by addition of sodium hydroxide. The composition of the vitamin stock solution was 5 mg of thiamine HCl, 2 mg of biotin, 2 mg of folic acid, 10 mg of nicotinamide, and 10 mg of pyridoxine HCl per liter of deionized water, and 20 ml of vitamin stock solution was added per liter of culture medium. The culture medium was made sterile by filtering through 0.2- μ m-pore-size filters (VacuCap 60; Pall Life Sciences, NY, USA) before the medium was inoculated with a single colony of *Pseudomonas* sp. strain ADP that had been grown on nutrient agar plates.

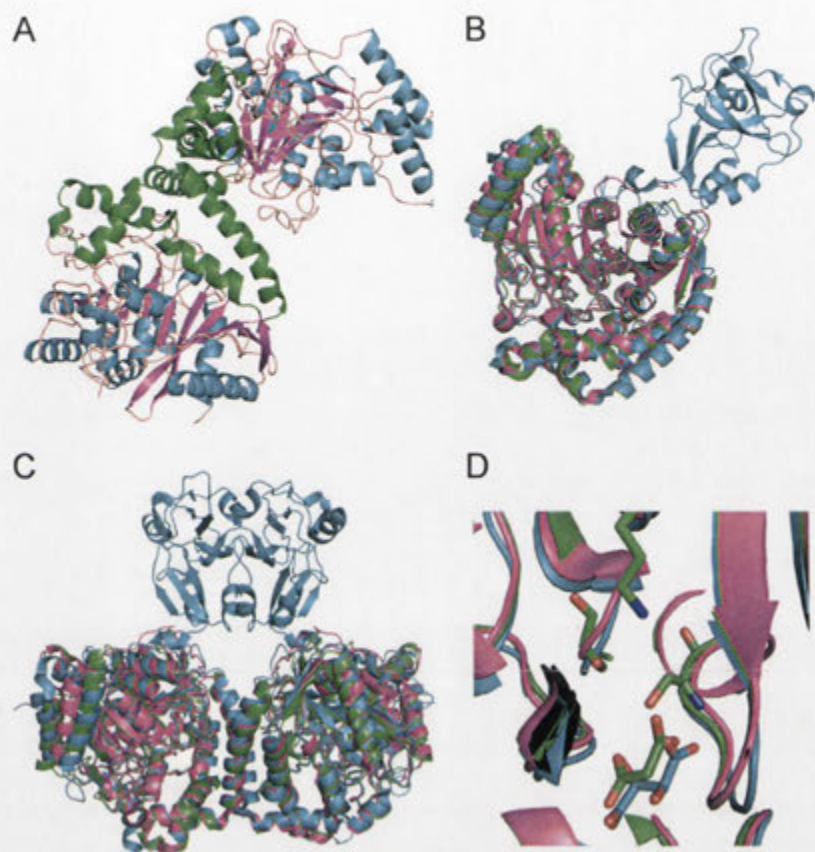


FIG 3 The AtzF structure and comparison to related structures. (A) A cartoon diagram of the dimer of the AtzF structure with the secondary structure colored by magenta for the beta sheet, cyan for alpha helices, and orange for loop structures. AtzF has two main domains: the catalytic domain and a second all-alpha-helical domain that forms the dimer interface. This has been highlighted by coloring these helices in green. (B) The AtzF structure superposed with structures with PDB accession numbers 4ISS and 4GYS. The structure with PDB accession number 4ISS is colored in cyan and has an extra domain which extends away from the rest of the molecule; AtzF is colored green, and the structure with PDB accession number 4GYS is colored magenta. The structures superpose well, despite limited sequence identity, with RMSD values of 1.3 to 1.6 Å. (C) Comparison of the AtzF dimer and the dimers with PDB accession numbers 4ISS and 4GYS. The figure shows how the dimers are similar and how the dimer with the extra domain (PDB accession number 4ISS) helps the dimer formation for this protein. (D) The catalytic site of these proteins. In two cases we see substrate mimetics bound into the catalytic site: in the case of AtzF we see clear density for malonate, whereas for the structure with PDB accession number 4ISS, there is tartrate.

The culture was grown until it reached an optical density at 600 nm of 0.6 to 0.7, after which the cells were harvested by centrifugation at $8,000 \times g$ for 15 min. Harvested cells were resuspended in 50 mM HEPES, 100 mM NaCl (pH 7.5) and lysed using BugBuster protein extraction reagent (Novagen, Darmstadt, Germany) per the manufacturer's instructions. After cell lysis, the soluble fraction was collected by spinning the lysate at $21,000 \times g$ for 15 min. The cell extract was passed through 0.45- μm -pore-size syringe-driven filters (Millex HV, Cork, Ireland), before it was loaded onto a 130-ml size-exclusion column (Superdex 200; GE, Uppsala, Sweden). The activities of AtzD, AtzE, and AtzF in each fraction were assessed using 68 μM cyanuric acid and biuret and 560 μM allophanate. The rates of background hydrolysis were subtracted in each case.

The fraction in which the catalytic activities of AtzD, AtzE, and AtzF were detected was resolved by SDS-PAGE (4 to 20% Tris-HEPES-SDS gels; Thermo Scientific, Rockford, IL, USA) (see Fig. S2 in the supplemental material). Certain zones corresponding to the expected locations of AtzD, AtzE, and AtzF were excised from the gel. In-gel tryptic digestion (with the inclusion of ProteaseMAX surfactant [Promega, Madison, WI, USA]) and tandem mass spectral analysis were performed as previously described using an Agilent Chip Cube system coupled to an Agilent XCD ion trap mass spectrometer (57). Mass spectra corresponding to common contaminants, such as the added trypsin and keratin, were identified be-

fore the remaining mass spectral data were used to search against a database containing all nonredundant UniProtKB/Swiss-Prot protein sequences from *Pseudomonas* and the Atz sequences using SpectrumMill software (version Rev A.03.03.084 SR4; Agilent) with its stringent default autovalidation settings.

Protein structure accession number. The structure identified in this study has been deposited in the Protein Data Bank under accession number 4CP8.

RESULTS AND DISCUSSION

Structure of amidase domain of AtzF. The X-ray crystal structure for AtzF₄₆₇ is dimeric (Fig. 3A), consistent with observations from previous studies (30). There are three independent dimers found in the asymmetric unit of the crystal: dimers A-B, C-D, and E-F. The A-B and C-D dimers are significantly more ordered, with average temperature factors (*B*-factors) of 30 to 33 Å². The E-F dimer has a weaker density than the other dimers and has an average *B*-factor about double the average *B*-factors of the other dimers, 65 to 75 Å².

The overall structure is similar to the structures of the *K. lactis* allophanate hydrolase (AH_K; PDB accession number 4ISS) and

the N terminus of the *G. betshedenensis* allophanate hydrolase (AH_{Gb}; PDB accession number 4GYS) and to the structures of other Ser-cis-Ser-Lys hydrolases, such as the amidase subunit of the *Staphylococcus aureus* glutamine amidotransferase (PDB accession number 2DQN) (48) used for molecular replacement in this study. The structure with PDB accession number 2DQN has a root mean square deviation (RMSD) of 1.8 Å with 16 gaps, and its sequence has less than 30% identity to the AtzF sequence. The dimer interface buries approximately 1,150 Å² of surface per monomer (Fig. 2). The dimer interface is essentially the same as that found in AH_{Kl} (PDB accession number 4ISS) and AH_{Gb}, except that the AtzF structure is missing the C-terminal domain, which forms a separate interface between the two monomers in AH_{Kl} (Fig. 3A).

Superposition of the AtzF and AH_{Kl} dimers (Fig. 3B and C) gives an RMSD of 1.6 Å for 851 residues (out of 895 and 1,226 residues for the AtzF and AH_{Kl} dimers, respectively) with 21 gaps (superpositioning was done with the secondary-structure matching (SSM) algorithm implemented in Coot software). A comparison with the structure of AH_{Gb} gives an RMSD of 1.3 Å for 857 residues (out of 895 and 921 residues for the AtzF and AH_{Gb} dimers, respectively) with 15 gaps (Fig. 3B and C). The two structures are very similar, with only a slight shift in helices 2 and 3 at the N terminus and a C-terminal extension of residues 444 to 462 in AH_{Gb} that is not seen in the AtzF structure due to the difference in the position of the truncations of each of the proteins. The only other region of significant deviation between the structures is in a mobile loop region between residues 255 and 268 of AtzF (residues 238 to 247 of AH_{Gb}). Notably, the structure used as the molecular replacement model, the structure with PDB accession number 2DQN, also possessed differences from AtzF similar to those shown by AH_{Gb}, in addition to an extended loop (residues 322 to 348 in the structure with PDB accession number 2DQN and residues 333 to 340 in AtzF). The active sites of AtzF, AH_{Gb}, and AH_{Kl} are also essentially identical to the positions of the amino acids essential for catalysis (Ser165, Ser189, and Lys91) and substrate binding (Tyr320 and Arg328), conserved between the three enzymes (Fig. 3D).

SAXS data (Fig. 4A to F) indicate that purified, full-length AtzF is a homotetramer in solution (Fig. 4A), rather than a hexamer, as size-exclusion chromatography (SEC) had previously suggested (31, 45). This is evident both from the fit of the data to a model of tetrameric AtzF and from the molecular weight of the complex calculated from the SAXS data (see Table S1 in the supplemental material). While there is a reasonable agreement overall between the tetrameric model of AtzF and the SAXS data, it can be seen that there is some divergence between the two at a low angle and around an oscillation in the data at about 0.1 Å⁻¹. This is most likely explained by some conformational flexibility in the C-terminal domains of the complex, though modeling of this structural feature is beyond the scope of this study. The tetramer itself is a cylinder of sorts that has dimensions of about 180 by 115 by 60 Å. The length of the long axis of the tetramer (180 Å) is longer than might be expected for a tetramer with the molecular weight of AtzF, and this is possibly the cause of this inconsistency between SEC and the more accurate observations from SAXS.

Previously, we had shown that removal of the C terminus of AtzF resulted in a 48-kDa protein (AtzF₄₆₇) which was kinetically indistinguishable from AtzF (45). However, the removal of the C-terminal domain altered the quaternary structure of AtzF, as

indicated by the reduction in apparent molecular mass of the enzyme during SEC from ~360 kDa for the wild-type protein to ~110 kDa (dimeric) for the truncated protein (45). SAXS of the amidase domain of AtzF (AtzF₄₆₇) confirmed that the truncated variant is indeed a dimer in solution (Fig. 4B). These observations suggest that the C terminus of AtzF is required for homotetramer formation but not homodimer formation. In contrast to the data presented by Fan et al. for the AH_{Kl} protein (37), which they showed was a monomer in solution when the C-terminal domain was removed (or disrupted by mutations), we found that the truncated AtzF migrates as a dimer in solution, as shown by both SAXS and chromatographic (SEC) data. As the dimeric interfaces are similar between the two proteins, it is not clear why the truncated form of AH_{Kl} would behave differently from the truncated form of AtzF.

It is interesting that the kinetic parameters for AtzF characterized here are different from those reported by Shapir et al. in 2005 (35); the k_{cat} and K_m values for AtzF reported here are ~4 to 13 s⁻¹ and ~120 to 300 μM, respectively (see Table S2 in the supplemental material), whereas they were 16 s⁻¹ and 1,500 μM, respectively, in the previous study (35). While it is unclear why there is such a large difference in K_m , it is noteworthy that the lowest substrate concentration tested in the study of Shapir et al. (34) was 200 μM (i.e., close to the K_m determined in the work reported here). It is also of note that the purity of the allophanate used in the work by Shapir et al. (35) was not reported, and so the reported value of 150 μM for the K_m reflects the highest possible K_m (assuming 100% substrate purity).

There is no catalytic advantage conferred by the C terminus of AtzF *in vitro*. The N-terminal domain of allophanate hydrolase deaminates allophanate to produce ammonia and *N*-carboxycarbamate (Fig. 1). The role of the smaller C-terminal domain is less clear, however. Fan et al. (37) suggested that the C-terminal domain of AH_{Kl} is catalytic, decarboxylating *N*-carboxycarbamate to form carbamate and carbon dioxide. On the basis of *in silico* substrate docking and mutagenesis studies, it was proposed that a histidine residue in the C-terminal domain (His492; Fig. 2) acts as a catalytic residue. If this is also true for AtzF, removal of the C-terminal domain or replacement of the catalytic histidine by alanine would be expected to reduce the rate of ammonia production by AtzF and the accumulation of the intermediate (*N*-carboxycarbamate) during the reaction.

We were unable to detect *N*-carboxycarbamate accumulation in reactions catalyzed by either the full-length AtzF or AtzF₄₆₇ by LC-MS analysis (see Fig. S2 in the supplemental material), which may indicate that the C terminus of AtzF is not catalytic. However, *N*-carboxycarbamate is highly unstable in an aqueous environment, decaying by spontaneous decarboxylation to form carbamate and carbon dioxide. As the stability of carbamates is known to be temperature and pH dependent (58), the pH of the AtzF-catalyzed reactions was raised to 9 and the temperature was reduced to 4°C. Despite the conditions favoring the accumulation of *N*-carboxycarbamate, we were unable to detect this product in reactions catalyzed by either full-length AtzF or AtzF₄₆₇. Although the absence of the accumulation of *N*-carboxycarbamate in reactions is not direct evidence for a lack of the hypothesized catalytic activity in the C terminus, it does suggest that such a function would have little physiological relevance.

To be sure that our inability to detect *N*-carboxycarbamate in the reactions was due to the absence of this reaction intermediate

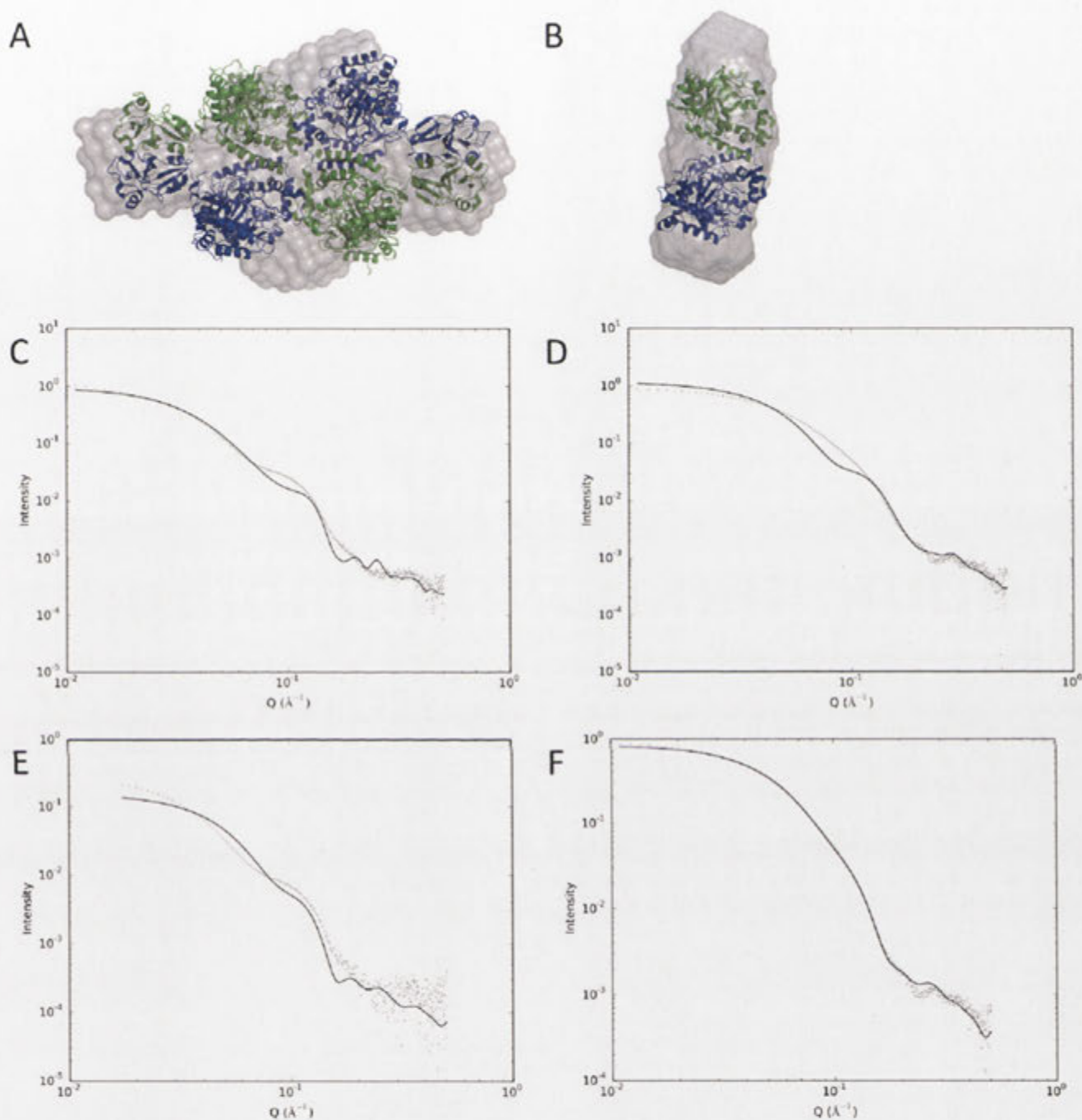


FIG 4 Analysis of the multimerization of AtzF and AtzF₄₆₇ by SAXS. (A) High-resolution model of an AtzF tetramer (green and blue) superimposed on a dummy atom model derived from the SAXS data (gray density). (B) High-resolution model of dimeric AtzF₄₆₇ (green and blue) superimposed on a dummy atom model derived from the SAXS data (gray density). (C) Fit between the SAXS data and the AtzF tetramer model shown in panel A calculated using CRYSOLOG. (D) Fit between the measured AtzF₄₆₇ protein and a model of tetrameric AtzF₄₆₇ based on the AtzF tetramer model calculated using CRYSOLOG. (E) Fit between dimeric AtzF and the measured AtzF SAXS data calculated using CRYSOLOG. (F) Fit between the dimeric AtzF₄₆₇ model shown in panel B and the measured AtzF₄₆₇ SAXS data.

rather than a flaw in our detection method, we also assessed the rate of ammonia production by full-length AtzF or AtzF₄₆₇ and by an AtzF variant in which the catalytic histidine residue had been replaced by an alanine residue. The enzymes were tested at a range

of pHs and temperatures, to account for the difference in *N*-carboxycarbamate stability at different temperatures and pHs.

For reactions at 28°C, the full-length AtzF was indistinguishable from the two variants at pH values of 8 and below. However,

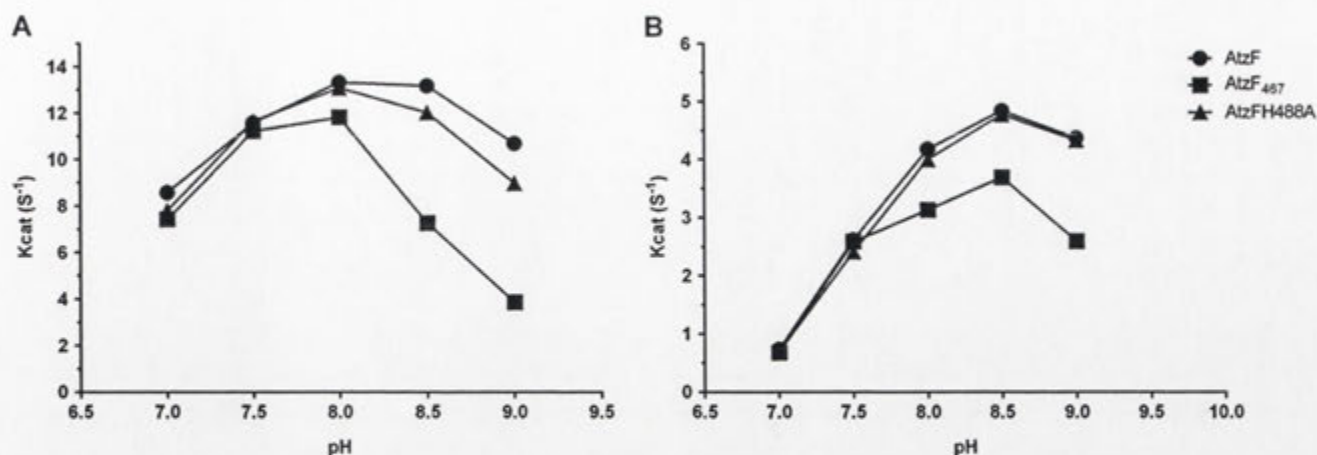


FIG 5 Activity of AtzF (and fragments) at different pHs and temperatures. k_{cat} values at different pH levels obtained at 28°C (A) and 4°C (B) for AtzF, AtzF₄₆₇, and AtzF H488A were compared. All the data points plotted had standard errors of less than 5.3%, and standard errors are presented in Table S2 in the supplemental material.

at pH values above 8, the full-length enzyme had k_{cat} values greater than those for either AtzF₄₆₇ or AtzF H488A. The difference for AtzF H488A was slight, but the reduction in activity for AtzF₄₆₇ was considerable (Fig. 5A). For reactions at 4°C, the k_{cat} value of full-length AtzF was indistinguishable from those of the two variants at pH values of 7.5 and below; however, at pH values above 7.5, the full-length enzymes had k_{cat} values greater than those for AtzF₄₆₇, albeit the effect on the catalytic rate was less extensive than that at 28°C. At 4°C, there was no difference in the k_{cat} values for AtzF and the AtzF H488A variant (Fig. 5B).

If the differences in apparent k_{cat} values were solely a result of a reduction in catalytic efficiency, it would be expected that the AtzF H488A and AtzF₄₆₇ variants would be indistinguishable from each other, which they are not. Moreover, the effect on the catalytic rate in the variants is reduced at low temperatures, where the stabilization of *N*-carboxycarbamate should make the effects more pronounced and not less so. Indeed, the k_{cat} values of AtzF H488A were essentially identical to those of the wild-type AtzF protein at 4°C.

As the pH-dependent differences between full-length AtzF and its variants were more pronounced at room temperature than at 4°C, it is possible that they are a result of differences in the thermal stability of the three enzymes. The residual activities of AtzF,

AtzF₄₆₇, and AtzF H488A were therefore tested after incubation for 5 min at temperatures between 30°C and 70°C and were assayed at pH 7, where no difference in the apparent k_{cat} values for the three enzymes had been observed, and pH 9, where differences in apparent k_{cat} values for the three enzymes had been observed (Fig. 6A and B).

At pH 7, there was no significant difference in the residual activities recovered for AtzF, AtzF₄₆₇, and AtzF H488A, but at pH 9, the apparent melting temperature (T_m) for AtzF₄₆₇ was reduced from ~44°C to ~41°C, and although the apparent T_m for AtzF H488A was ~45°C, it was lower than that for AtzF by nearly 1.5°C. It is therefore plausible that at least a component of the pH-dependent difference in apparent k_{cat} values was a result of a pH-dependent reduction in stability for AtzF when the C terminus was removed or H488 was mutated to an alanine.

Regardless of the rate differences observed at pHs higher than the physiological pH (i.e., 6.5 to 7.5), we were unable to observe a difference in the rate of ammonia release between the wild-type and variant AtzF enzymes tested *in vitro* under physiological pH conditions.

AtzF appears to form a large complex with AtzD and AtzE. Allophanate is unstable and has been reported to decompose rapidly to form urea and carbon dioxide under physiological condi-

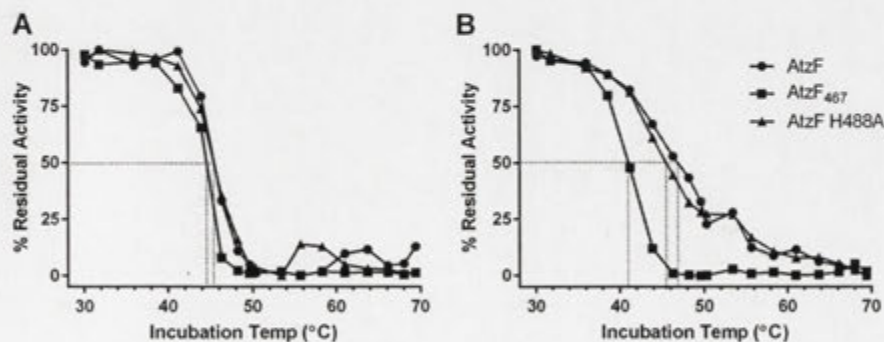


FIG 6 Effect of pH on residual activity of AtzF and its variants. Differences in the residual activity of AtzF, AtzF₄₆₇, and AtzF H488A at pH 7 (A) and pH 9 (B) were compared. Enzymes were incubated in pH 7 or pH 9 buffers for 5 min at temperatures ranging from 30 to 70°C before testing for residual activity.

tions (Fig. 1) (30). For both the ureolytic pathway and triazine degradation pathway, the spontaneous decomposition of allophanate is undesirable, causing an ATP-consuming futile cycle for the ureolytic enzymes and forming a dead-end product in triazine hydrolysis that results in the metabolic loss of two of the three nitrogens of the heterocycle. It is likely that the formation of a urea carboxylase/allophanate hydrolase complex in the ureolytic system protects allophanate from decarboxylation by channeling the substrate from one active site to another. It is reasonable to assume that a similar relationship may exist between AtzF and the biuret hydrolase, AtzE, which prevents the production of urea. As AtzF is a relatively inefficient enzyme, with k_{cat} and K_m values of 8 s^{-1} and $120 \mu\text{M}$, respectively (see Table S2 in the supplemental material), it is likely that the nitrogen from the degradation of the triazine ring is lost to the competing abiotic decarboxylation rather than channeled through allophanate hydrolase in the absence of the formation of a substrate-channeling complex.

SEC was used to fractionate a *Pseudomonas* sp. strain ADP cell extract obtained from a culture that had been induced for AtzD, AtzE, and AtzF production by incubation in a medium containing cyanuric acid as the sole carbon source. AtzD, AtzE, and AtzF activities were found to coelute in a fraction with an estimated molecular mass of $\sim 660 \text{ kDa}$. The activities of these three enzymes were not found in other fractions. This observation is consistent with a complex that consists of four AtzD monomers, four AtzF monomers, and two AtzE monomers, albeit deriving the stoichiometry of the three enzymes from SEC alone may be misleading.

The SEC fraction with AtzD, AtzE, and AtzF activities was resolved by SDS-PAGE and with gel sections taken from where the three proteins should migrate (on the basis of molecular mass; see Fig. S3 in the supplemental material). LC-MS of the tryptic digestion products from these sections showed that peptides that were consistent with a tryptic fragmentation pattern of AtzD, AtzE, and AtzF were present. The sequence coverage for AtzD, AtzE, and AtzF was 21% by 9 unique peptides, 25% by 9 unique peptides, and 3% by 2 unique peptides, respectively. Unsurprisingly, other proteins identified in the samples were abundant cellular components of *Pseudomonas*, such as translation elongation factors (GI 169761912, GI 169761911, and GI 169760553) and heat shock proteins (GI 169760669). When considered along with the enzyme activity data from SEC, these data support the presence of a large protein complex consisting of AtzD, AtzE, and AtzF.

Production of ammonia was observed upon addition of cyanuric acid to this fraction, with 3 mol of ammonia being produced for every mole of cyanuric acid added, suggesting that the complex mineralized cyanuric acid. Interestingly, the rate of ammonia production was much greater upon addition of cyanuric acid than it was when biuret (the AtzE substrate) was added (Table 3), which may imply that access to the AtzE active site from solvent may be restricted and passage of substrate from the AtzD active site to the AtzE active site may occur via substrate tunneling or channeling. If the AtzD product (1-carboxybiuret; Fig. 1) is protected from extended contact with the solvent, it is possible that AtzE is a 1-carboxybiuret hydrolase, rather than a biuret hydrolase. Unfortunately, despite considerable effort, we were unable to produce soluble AtzE in a quantity sufficient to reconstruct the complex *in vitro* for further study and so were unable to probe this hypothesis further.

Conclusion. AtzF is a component of a large ($\sim 0.7\text{-MDa}$) cyanuric acid-mineralizing complex, comprised of AtzD, AtzE, and

TABLE 3 Rates of ammonia release by the SEC fraction containing AtzD, AtzE, and AtzF activities^a

Substrate	Substrate/ammonia ratio	Rate of ammonia release (nM/s)
Cyanuric acid	3.2	45.4
Biuret	ND	1
Allophanate	1.1	180

^a Cyanuric acid, biuret, and allophanate were added as substrates. The SEC fraction was estimated to contain proteins with a molecular mass of $\sim 660 \text{ kDa}$. ND, the stoichiometry of substrate to ammonia was not determined for biuret, because of the low rate of ammonia production.

AtzF. We postulate that this complex protects the unstable intermediate allophanate from spontaneous decarboxylation under physiological conditions by substrate channeling and/or by precluding water from reaching the reactive centers, strategies used in other enzymes that catalyze water-sensitive reactions, such as carbamoyl phosphate synthase and the GatCAB glutamine amidotransferase (48). Interestingly, ureolytic allophanate hydrolases also form complexes with urea carboxylase. This suggests either that allophanate hydrolases have evolved to form multiprotein complexes more than once or that they have switched partners over the course of their evolution.

The biochemical data that we have presented (accumulation of *N*-carboxycarbamate and the rate of ammonia production by AtzF and its variants) fail to support a physiologically relevant catalytic function for the C terminus of AtzF. It is possible that the proposed *N*-carboxycarbamate decarboxylase activity is relevant under conditions that stabilize *N*-carboxycarbamate, such as a water-limited environment. These conditions could be present in the cyanuric acid-mineralizing complex, albeit further work will be required to elucidate structure-function relationships within the complex.

ACKNOWLEDGMENTS

We thank the beamline scientists of the Australian Synchrotron and the CSIRO Collaborative Crystallisation Centre, Melbourne, Australia. We also thank Chris Blake (Research School of Chemistry, Australian National University) for his help with the NMR studies and Carol Hartley and Robyn Russell (CSIRO) for helpful discussions.

We do not have a conflict of interest to declare.

REFERENCES

- Tomlin C (ed). 2006. The pesticide manual: a world compendium, 14th ed. BCPC, Alton, United Kingdom.
- Thurman EM, Meyer MT. 1996. Herbicide metabolites in surface water and groundwater: introduction and overview, p 1–15. In Meyer MT, Thurman EM (ed), Herbicide metabolites in surface water and groundwater, vol 630. American Chemical Society, Washington, DC.
- Gavrillescu M. 2005. Fate of pesticides in the environment and its bioremediation. *Eng Life Sci* 5:497–526. <http://dx.doi.org/10.1002/elsc.200520098>.
- Hayes TB, Khoury V, Narayan A, Nazir M, Park A, Brown T, Adame L, Chan E, Buchholz D, Stueve T, Gallipeau S. 2010. Atrazine induces complete feminization and chemical castration in male African clawed frogs (*Xenopus laevis*). *Proc Natl Acad Sci U S A* 107:4612–4617. <http://dx.doi.org/10.1073/pnas.0909519107>.
- Hayes T, Haston K, Tsui M, Hoang A, Haeffele C, Vonk A. 2003. Atrazine-induced hermaphroditism at 0.1 ppb in American leopard frogs (*Rana pipiens*): laboratory and field evidence. *Environ Health Perspect* 111:568–575.
- Hayes TB, Collins A, Lee M, Mendoza M, Noriega N, Stuart AA, Vonk A. 2002. Hermaphroditic, demasculinized frogs after exposure to the her-

- bicide atrazine at low ecologically relevant doses. *Proc Natl Acad Sci U S A* 99:5476–5480. <http://dx.doi.org/10.1073/pnas.082121499>.
7. Udikovic-Kolic N, Scott C, Martin-Laurent F. 2012. Evolution of atrazine-degrading capabilities in the environment. *Appl Microbiol Biotechnol* 96:1175–1189. <http://dx.doi.org/10.1007/s00253-012-4495-0>.
 8. Russell RJ, Scott C, Jackson CJ, Pandey R, Pandey G, Taylor MC, Coppin CW, Liu JW, Oakshott JG. 2011. The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects. *Evol Appl* 4:225–248. <http://dx.doi.org/10.1111/j.1752-4571.2010.00175.x>.
 9. 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. <http://dx.doi.org/10.1128/JB.01257-06>.
 10. Wackett LP. 2004. Evolution of enzymes for the metabolism of new chemical inputs into the environment. *J Biol Chem* 279:41259–41262. <http://dx.doi.org/10.1074/jbc.R400014200>.
 11. Noor S, Changey F, Oakshott JG, Scott C, Martin-Laurent F. 2014. Ongoing functional evolution of the bacterial atrazine chlorohydrolase AtzA. *Biodegradation* 25:21–30. <http://dx.doi.org/10.1007/s10532-013-9637-2>.
 12. Noor S, Taylor MC, Russell RJ, Jermini LS, Jackson CJ, Oakshott JG, Scott C. 2012. Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS One* 7:e39822. <http://dx.doi.org/10.1371/journal.pone.0039822>.
 13. Raillard S, Krebber A, Chen YC, Ness JE, Bermudez E, Trinidad R, Fullem R, Davis C, Welch M, Seffernick J, Wackett LP, Stemmer WPC, Minshull J. 2001. Novel enzyme activities and functional plasticity revealed by recombining highly homologous enzymes. *Chem Biol* 8:891–898. [http://dx.doi.org/10.1016/S1074-5521\(01\)00061-8](http://dx.doi.org/10.1016/S1074-5521(01)00061-8).
 14. Seffernick JL, de Souza ML, Sadowsky MJ, Wackett LP. 2001. Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *J Bacteriol* 183:2405–2410. <http://dx.doi.org/10.1128/JB.183.8.2405-2410.2001>.
 15. 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. <http://dx.doi.org/10.1111/mmi.12249>.
 16. Lima D, Viana P, Andre S, Chelinho S, Costa C, Ribeiro R, Sousa JP, Fialho AM, Viegas CA. 2009. Evaluating a bioremediation tool for atrazine contaminated soils in open soil microcosms: the effectiveness of bioaugmentation and biostimulation approaches. *Chemosphere* 74:187–192. <http://dx.doi.org/10.1016/j.chemosphere.2008.09.083>.
 17. Scott C, Lewis SE, Milla R, Taylor MC, Rodgers AJW, Dumsday G, Brodie JE, Oakshott JG, Russell RJ. 2010. A free-enzyme catalyst for the bioremediation of environmental atrazine contamination. *J Environ Manage* 91:2075–2078. <http://dx.doi.org/10.1016/j.jenvman.2010.05.007>.
 18. Scott C, Pandey G, Hartley CJ, Jackson CJ, Cheesman MJ, Taylor MC, Pandey R, Khurana JL, Teese M, Coppin CW, Weir KM, Jain RK, Lal R, Russell RJ, Oakshott JG. 2008. The enzymatic basis for pesticide bioremediation. *Indian J Microbiol* 48:65–79. <http://dx.doi.org/10.1007/s12088-008-0007-4>.
 19. Topp E, Martin-Laurent F, Hartmann A, Soulas G. 2004. Bioremediation of atrazine-contaminated soil, p 141–154. *In* Gan JJ, Zhu PC, Aust SD, Lemley AT (ed), *Pesticide decontamination and detoxification*, vol 863. American Chemical Society, Washington, DC.
 20. de Souza 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.
 21. de Souza 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.
 22. Boundy-Mills 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.
 23. 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. <http://dx.doi.org/10.1128/JB.00630-07>.
 24. 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.
 25. 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. <http://dx.doi.org/10.1128/JB.184.19.5376-5384.2002>.
 26. 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. <http://dx.doi.org/10.1128/AEM.69.6.3653-3657.2003>.
 27. 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)/barbiturate protein family. *J Bacteriol* 194:4579–4588. <http://dx.doi.org/10.1128/JB.00791-12>.
 28. Cameron SM, Durchschein K, Richman JE, Sadowsky MJ, Wackett LP. 2011. New family of biuret hydrolases involved in *s*-triazine ring metabolism. *ACS Catal* 1:1075–1082. <http://dx.doi.org/10.1021/cs200295n>.
 29. 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. <http://dx.doi.org/10.1128/AEM.72.4.2491-2495.2006>.
 30. 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. <http://dx.doi.org/10.1128/AEM.71.8.4437-4445.2005>.
 31. Shapir N, Sadowsky MJ, Wackett LP. 2005. Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp. strain ADP. *J Bacteriol* 187:3731–3738. <http://dx.doi.org/10.1128/JB.187.11.3731-3738.2005>.
 32. Jackson CJ, Coppin CW, Carr PD, Aleksandrov A, Wilding M, Sugrue E, Uebels J, Paks M, Newman J, Peat TS, Russell RJ, Field M, Weik M, Oakshott 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. <http://dx.doi.org/10.1128/AEM.00916-14>.
 33. 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. <http://dx.doi.org/10.1074/jbc.M110.138677>.
 34. Shapir N, Pedersen C, Gil O, Strong L, Seffernick J, Sadowsky MJ, Wackett LP. 2006. TrzN from *Arthrobacter aureus* TC1 is a zinc amidohydrolase. *J Bacteriol* 188:5859–5864. <http://dx.doi.org/10.1128/JB.00517-06>.
 35. Shapir N, Rosendahl C, Johnson G, Andreina M, Sadowsky MJ, Wackett LP. 2005. Substrate specificity and colorimetric assay for recombinant TrzN derived from *Arthrobacter aureus* TC1. *Appl Environ Microbiol* 71:2214–2220. <http://dx.doi.org/10.1128/AEM.71.5.2214-2220.2005>.
 36. Mulbry WW, Zhu H, Nour SM, Topp E. 2002. The triazine hydrolase gene *trzN* from *Nocardioideis* sp. strain C190: cloning and construction of gene-specific primers. *FEMS Microbiol Lett* 206:75–79. <http://dx.doi.org/10.1111/j.1574-6968.2002.tb10989.x>.
 37. Fan C, Li Z, Yin HY, Xiang S. 2013. Structure and function of allophanate hydrolase. *J Biol Chem* 288:21422–21432. <http://dx.doi.org/10.1074/jbc.M113.453837>.
 38. 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. <http://dx.doi.org/10.1021/bi301242m>.
 39. Whitney PA, Cooper TG, Magasani B. 1973. Induction of urea carboxylase and allophanate hydrolase in *Saccharomyces cerevisiae*. *J Biol Chem* 248:6203–6209.
 40. Whitney PA, Cooper TG. 1972. Urea carboxylase and allophanate hydrolase—2 components of a multienzyme complex in *Saccharomyces cerevisiae*. *Biochem Biophys Res Commun* 49:45–51. [http://dx.doi.org/10.1016/0006-291X\(72\)90007-1](http://dx.doi.org/10.1016/0006-291X(72)90007-1).
 41. Whitney PA, Cooper TF. 1972. Urea carboxylase and allophanate hydrolase—2 components of adenosine triphosphate-urea amidohydrolase in *Saccharomyces cerevisiae*. *J Biol Chem* 247:1349–1353.
 42. Kanamori T, Kanou N, Kusakabe S, Atomi H, Imanaka T. 2005. Allophanate hydrolase of *Oleomonas sagaranensis* involved in an ATP-dependent degradation pathway specific to urea. *FEMS Microbiol Lett* 245:61–65. <http://dx.doi.org/10.1016/j.femsle.2005.02.023>.

43. Shin S, Yun YS, Koo HM, Kim YS, Choi KY, Oh BH. 2003. Characterization of a novel Ser-*cis*Ser-Lys catalytic triad in comparison with the classical Ser-His-Asp triad. *J Biol Chem* 278:24937–24943. <http://dx.doi.org/10.1074/jbc.M302156200>.
44. 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 J* 21:2509–2516. <http://dx.doi.org/10.1093/emboj/21.11.2509>.
45. 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 Cryst Commun* 70:310–315. <http://dx.doi.org/10.1107/S2053230X13034705>.
46. Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain-reaction. *Gene* 77:51–59. [http://dx.doi.org/10.1016/0378-1119\(89\)90358-2](http://dx.doi.org/10.1016/0378-1119(89)90358-2).
47. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* 40:658–674. <http://dx.doi.org/10.1107/S0021889807021206>.
48. Nakamura A, Yao M, Chimnaronk S, Sakai N, Tanaka I. 2006. Ammonia channel couples glutaminase with transamidase reactions in GatCAB. *Science* 312:1954–1958. <http://dx.doi.org/10.1126/science.1127156>.
49. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501. <http://dx.doi.org/10.1107/S0907444910007493>.
50. Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA. 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):355–367. <http://dx.doi.org/10.1107/S0907444911001314>.
51. Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. PRIMUM: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* 36:1277–1282. <http://dx.doi.org/10.1107/S0021889803012779>.
52. Svergun D, Barberato C, Koch MHJ. 1995. CRYSOLE—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* 28:768–773. <http://dx.doi.org/10.1107/S0021889895007047>.
53. Franke D, Svergun DI. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* 42:342–346. <http://dx.doi.org/10.1107/S0021889809000338>.
54. Kozin MB, Svergun DI. 2001. Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* 34:33–41. <http://dx.doi.org/10.1107/S0021889800014126>.
55. Mylonas E, Svergun DI. 2007. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J Appl Crystallogr* 40:S245–S249. <http://dx.doi.org/10.1107/S002188980700252X>.
56. Whitten AE, Cai SZ, Trehwella J. 2008. MULCh: modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies. *J Appl Crystallogr* 41:222–226. <http://dx.doi.org/10.1107/S0021889807055136>.
57. Campbell PM, Trueman HE, Zhang Q, Kojima K, Kameda T, Sutherland TD. 2014. Cross-linking in the silks of bees, ants and hornets. *Insect Biochem Mol Biol* 48:40–50. <http://dx.doi.org/10.1016/j.ibmb.2014.02.009>.
58. Frahn JL, Mills JA. 1964. Paper ionophoresis of amino compounds: formation of carbamates and related reactions. *Aust J Chem* 17:256–260. <http://dx.doi.org/10.1071/CH9640256>.

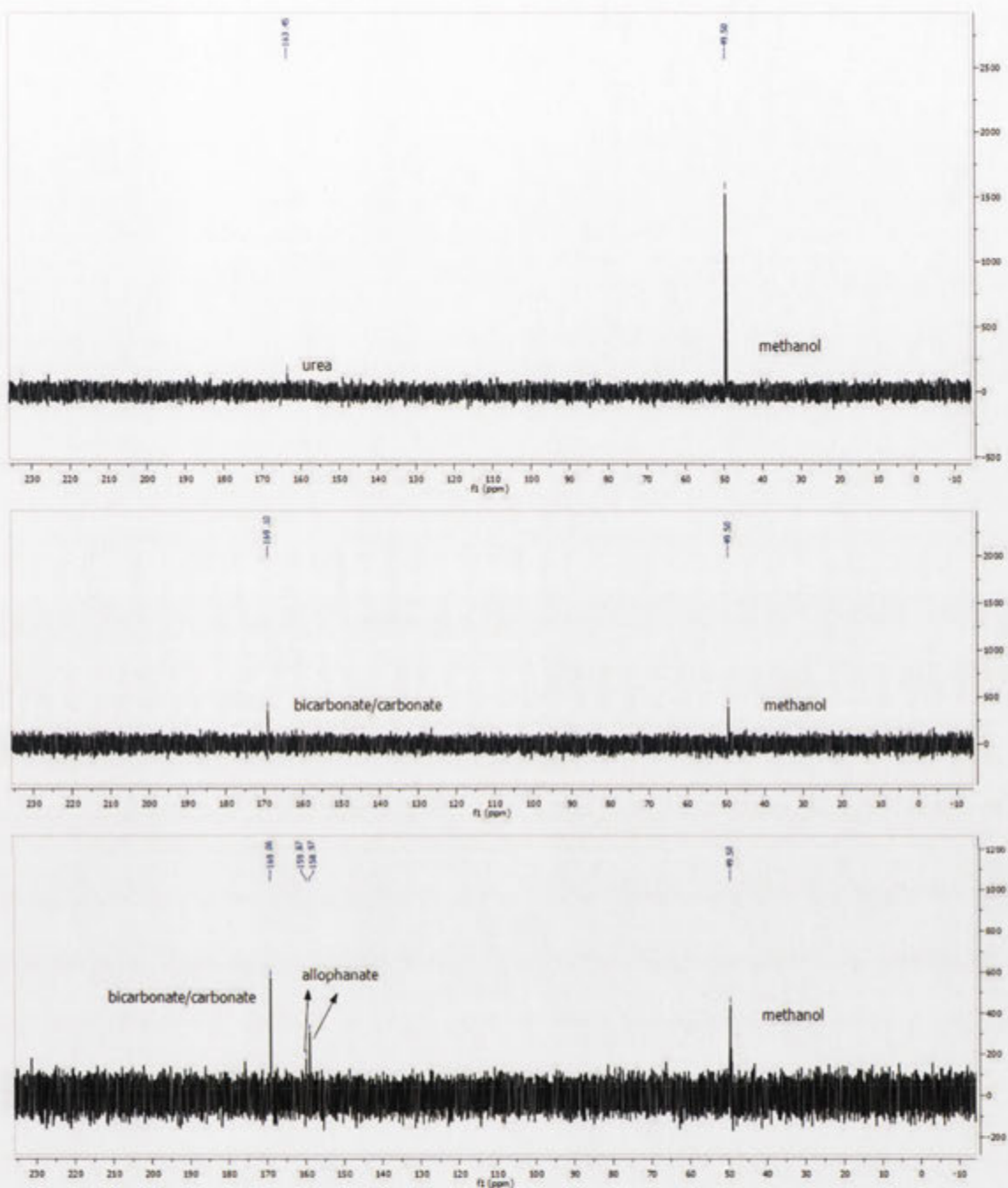
Supplementary Table 1. SAXS parameters across the AtzF and AtzF₄₆₇ concentration series.

AtzF				
Conc (mg/ml)	I(0)	MW (Kda)	Rg	No. subunits*
0.4	0.065	249.3	50.6	3.8
0.7	0.132	253.2	51.1	3.8
1.5	0.267	256.1	51.2	3.9
2.9	0.521	249.8	51.2	3.8
5.8	0.993	238.1	51.4	3.6

AtzF₄₆₇				
protein conc (mg/ml)	I(0)	Mol. Weight (Kda)	Rg (Å)	No. subunits*
0.8	0.071	119.8	37.7	2.3
1.6	0.132	111.4	37.2	2.2
3.3	0.268	113.1	37.0	2.2
6.6	0.506	106.7	35.8	2.1
13.1	0.919	96.9	36.5	1.9

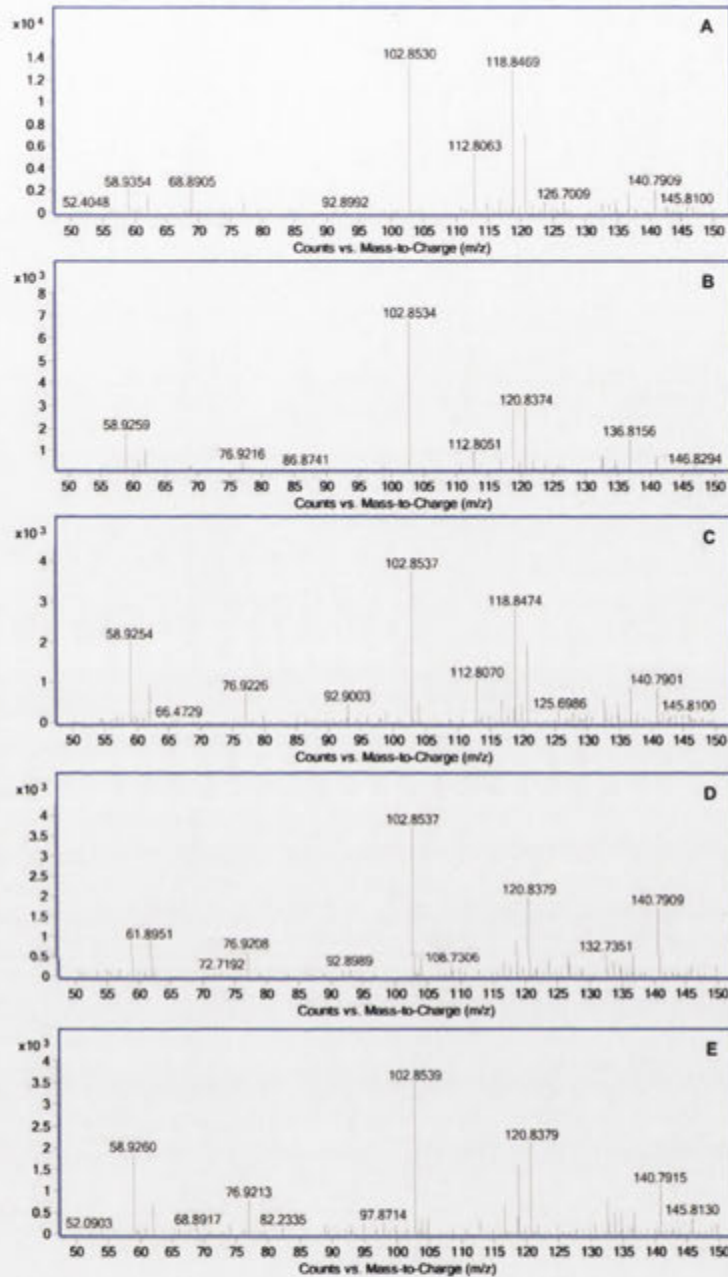
Supplementary Table 2. Effect of temperature and pH on the catalytic properties of AtzF, AtzF₄₆₇, and AtzFH488A.

Enzyme	Temp (°C)	pH	K_M (μM)	k_{cat} (s ⁻¹)	k_{cat}/K_M (M ⁻¹ .s ⁻¹) x10 ⁴
AtzF	28	7.0	120.0 ± 21.4	8.6 ± 0.4	7.1 ± 1.3
AtzF ₄₆₇			114.6 ± 17.2	7.4 ± 0.3	6.5 ± 0.1
AtzF H488A			127.6 ± 13.6	7.8 ± 0.2	6.1 ± 0.7
AtzF	28	7.5	149.5 ± 15.7	11.6 ± 0.3	7.7 ± 0.8
AtzF ₄₆₇			168.6 ± 21.9	11.2 ± 0.4	6.7 ± 0.9
AtzF H488A			156.1 ± 11.2	11.6 ± 0.2	7.5 ± 0.5
AtzF	28	8.0	115.7 ± 8.8	13.3 ± 0.2	11.5 ± 0.9
AtzF ₄₆₇			109.3 ± 4.2	11.8 ± 0.1	10.8 ± 0.4
AtzF H488A			113.9 ± 4.7	13.1 ± 0.1	11.5 ± 0.5
AtzF	28	8.5	78.6 ± 4.6	13.2 ± 0.2	16.8 ± 0.1
AtzF ₄₆₇			71.0 ± 4.7	7.3 ± 0.1	10.2 ± 0.7
AtzF H488A			72.6 ± 5.0	12.0 ± 0.2	16.6 ± 1.0
AtzF	28	9.0	74.0 ± 4.6	10.7 ± 0.1	14.4 ± 0.9
AtzF ₄₆₇			174.9 ± 26.0	3.9 ± 0.2	2.2 ± 0.3
AtzF H488A			82.2 ± 5.9	9.0 ± 0.1	10.9 ± 0.8
AtzF	4	7.0	380.2 ± 62.0	0.7 ± 0.04	0.19±0.03
AtzF ₄₆₇			363.3 ± 61.0	0.7 ± 0.04	0.19±0.07
AtzF H488A			305.5 ± 46.9	0.7 ± 0.03	0.2 ± 0.01
AtzF	4	7.5	146.3 ± 5.4	2.6 ± 0.02	1.8 ± 0.1
AtzF ₄₆₇			164.0 ± 12.7	2.6 ± 0.05	1.6 ± 0.1
AtzF H488A			124.1 ± 10.9	2.4 ± 0.06	1.9 ± 0.2
AtzF	4	8.0	44.6 ± 3.3	4.2 ± 0.05	9.4 ± 0.7
AtzF ₄₆₇			57.4 ± 4.8	3.1 ± 0.05	5.4 ± 0.5
AtzF H488A			40.4 ± 2.5	4.0 ± 0.04	9.9 ± 0.6
AtzF	4	8.5	50.8 ± 3.9	4.8 ± 0.06	9.5 ± 0.7
AtzF ₄₆₇			54.6 ± 4.3	3.7 ± 0.05	6.8 ± 0.5
AtzF H488A			47.4 ± 3.7	4.8 ± 0.06	10.1 ± 0.7
AtzF	4	9.0	59.6 ± 4.9	4.4 ± 0.07	7.3 ± 0.6
AtzF ₄₆₇			66.0 ± 5.2	2.6 ± 0.04	3.9 ± 0.3
AtzF H488A			54.5 ± 3.2	4.3 ± 0.05	8.0 ± 0.5

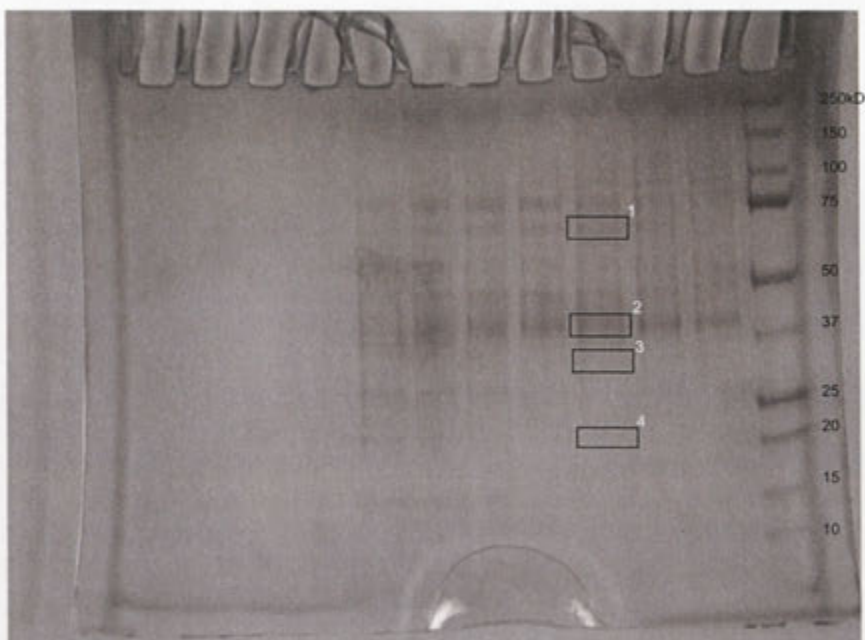


Supplementary figure 1. ^{13}C NMR of allophanate, urea and sodium bicarbonate.

^{13}C NMR shifts of urea, sodium bicarbonate and potassium allophanate in 1M potassium hydroxide prepared in D_2O . First two spectra are of pure urea and bicarbonate/carbonate, which were expected to be present as in allophanate sample as a result of autodecarboxylation, whereas, the third spectra correspond to allophanate sample.



Supplementary Figure 2. Allophanate and *N*-carboxycarbamate in AtzF catalyzed allophanate deamination reactions. Reactions were conducted at pH 9.0 and 4 °C to promote stabilization of *N*-carboxycarbamate (mass ≈ 104). An ion consistent with the mass of allophanate (≈ 103) is observed in all reactions. A) Allophanate only, 0 minutes incubation, B) allophanate only at 45 minutes, C) allophanate with AtzF₄₆₈ (0.96 μM) after 45 minutes, D) allophanate with AtzF₄₆₈ (0.48 μM) after 45 minutes, and E) allophanate with AtzF (0.48 μM) after 45 minutes.



Supplementary figure 3. Gel fragments used in mass spectroscopy

Gel fragments used for mass spectroscopy of ~660kD complex that showed cyanuric acid hydrolase, biuret hydrolase and allophanate hydrolase activity and peptides.

Chapter 7: Discussion

7.1 Studies with the enzymes of atrazine mineralisation pathway

Five of six of the atrazine mineralization pathway enzymes from *Pseudomonas* sp. strain ADP were expressed and purified in *E. coli*. The X-ray structures of four of the enzymes (AtzA, AtzC, AtzD and AtzF) were obtained. Detailed structure-function studies with the four structures obtained provided plausible reaction mechanisms that were supported by biochemical and mutagenesis studies [1-4]. Moreover, AtzD, AtzE and AtzF were shown to form a large (~700 kDa) complex, possibly facilitating substrate channelling of water-sensitive metabolites [1].

The structures of AtzB and AtzE remain unsolved. AtzB was purified and the his-tag has been removed. Further crystallography studies will be conducted, including co-crystallization studies with inhibitors and substrates. Elucidating this structure will be of interest as AtzB accepts a broad range of substrates, including chlorinated-s-triazines. Understanding how the promiscuous chlorohydrolase activity of AtzB functions may shed further light on how 'non-natural' catalytic functions can be acquired. Moreover, *in vitro* evolution of AtzB towards its promiscuous chlorohydrolase activity will provide an enlightening companion study to that already conducted with AtzA and TriA [5, 6], particularly if the evolutionary trajectories closely resemble those seen in the Noor *et al.* (2012) study [6].

Obtaining the AtzE structure is a more daunting prospect; expression of AtzE on its own and in concert with AtzD and AtzF in *E. coli* resulted in insoluble protein (personal correspondence, Lyndall Briggs). Other attempts to improve soluble expression of AtzE in *E. coli* by varying incubation and induction conditions have also failed. It is possible that AtzE cannot fold correctly except during the formation of the AtzDEF complex. Attempts are underway to purify this complex from *Pseudomonas* sp. strain ADP, to be followed by structural studies using X-ray crystallography, small-angle X-ray scattering and EM-tomography. Such studies will elucidate the functional interactions between the components of the complex, and will allow testing of the substrate channelling hypothesis proposed in Balotra *et al.* (2015) [1].

Notwithstanding the two structures yet to be obtained, the work presented here provides examples of naturally evolved enzymes that illustrate some of the molecular mechanisms and constraints that govern the evolution of new physiological and biochemical functionality in enzymes.

7.2 AtzA and the limits of evolutionary innovation

In Chapter 2 [4], it was suggested that the active site of AtzA is poorly adapted for binding atrazine, as demonstrated by: 1) the high K_M of AtzA for atrazine, 2) *in silico* docking studies, and 3) the failure to obtain an AtzA:atrazine complex in crystallization studies. Additionally, several active-site amino acid residues are held in strained conformations in the crystal structure [4] and the catalytic cation (Fe^{2+}) was shown to be co-ordinated *via* some

unusually distant amino acid residues (with bond-lengths of up to 3.14 Å). These observations suggest that AtzA may be a poorly adapted enzyme. Indeed, this is particularly apparent when AtzA is compared with the alternative atrazine chlorohydrolase, TrzN, which has far better complementarity between its active site and substrate (resulting in a lower K_M value TrzN *cf.* AtzA for atrazine).

Approximately two decades have elapsed since AtzA was first isolated; despite the potentially enormous number of generations of bacteria that separate this event from the present (assuming a 30 minute to 1 week generation time this could represent between ~1,000 and $\sim 3 \times 10^5$ generations) there is little deviation between the historical AtzA sequence and AtzA sequences sourced from contemporary soils [7, 8]. Indeed, even where differences to the amino acid sequence in more modern AtzAs have been identified, they have little effect on the catalytic properties of the enzyme [8]. *In vitro* evolution studies seem to indicate that little improvement can be made to the catalytic properties of AtzA *via* step-wise accumulation of single amino acid substitutions. A thorough attempt at directed evolution by random mutagenesis by Wang *et al.* (2013) resulted in minor improvements in the catalytic efficiency of AtzA (~ 1.5 fold for k_{cat} and ~ 4 fold for k_{cat}/K_M). Large improvements to the k_{cat}/K_M of AtzA (>20-fold) have been achieved in the laboratory, but only by randomizing five amino acids simultaneously [9] (note: this strategy was adopted because random mutagenesis failed to produce detectable improvements in atrazine chlorohydrolase activity; personal correspondence, Colin Scott).

It appears that AtzA may be an 'evolutionary dead-end' with respect to atrazine chlorohydrolase activity; it is poorly adapted to its substrate, but has not improved catalytically *via* natural selection in the last two decades and cannot be improved catalytically by step-wise accumulation of single amino acid substitutions (i.e. directed evolution) in the laboratory. However, as multiple simultaneous amino acid substitutions can improve the atrazine chlorohydrolase activity, it appears that AtzA occupies a local fitness optimum and that fitness optima and that greater catalytic activities may only be accessible by traversing adaptive valleys, i.e. *via* less fit intermediate forms of the enzyme that would be out competed by the parental AtzA.

It is possible that AtzA will eventually be outcompeted by the catalytically superior TrzN. Indeed, a study by Arbeli *et al.* (2009) in Columbia found that the *trzN* gene was far more prevalent in isolated atrazine-degrading bacteria, with 83 of 84 isolates carrying the *trzN* gene and only one containing the *atzA* gene [10]. In the 1990s, the atrazine-degrading soil microbiota at the different locations had been dominated by *atzA* containing bacteria [10, 11]. The shift in the composition of the microbial community could reflect a fitness advantage conferred by TrzN, albeit other factors promoting a change in the composition of the microbial community at these locations could have co-incidentally enriched for *trzN*-containing bacteria. Unfortunately, there are no temporally longitudinal studies available with which to test this hypothesis.

7.3 Partner swapping and domain recruitment as engines of innovation

Both AtzD and AtzF present examples of evolutionary innovation *via* domain recruitment, a relatively common and well understood mechanism by which enzymes innovate their biochemical and physiological function (extensively reviewed by Kim *et al.* and Nagano *et al.* [12, 13]). In the case of AtzD, three YjgF-family monomers have fused to form a new fold. This ancestral 'Toblerone'-fold then adapted to fulfil at least two new functions: barbituric acid hydrolysis and cyanuric acid hydrolysis to a new function. Phylogenetic analysis suggested that there are two 'Toblerone'-fold clades that appeared to be functionally distinct from the clades that contained enzymes that have been shown to have cyanuric acid or barbituric acid hydrolase activities. It is possible that these clades possess as yet uncharacterised functions, perhaps also as ring opening amidases.

The allophanate hydrolase has engaged in partner swapping, rather than domain swapping *per se*, to fulfil two discrete physiological functions: cyanuric acid catabolism and urea catabolism [1, 14]. In both catabolic pathways, allophanate hydrolase is part of a multidomain enzyme along with urea carboxylase (for urea catabolism) or cyanuric acid hydrolase and biuret hydrolase (for cyanuric acid catabolism) [1, 14]. The allophanate hydrolases from both pathways are biochemically indistinguishable, and it is only through their protein:protein interactions that they fulfil diverse physiological roles. The structural determinants that are required for allophanate hydrolase to participate two multiprotein complexes are, as yet, unknown. It is also uncertain whether the complexes evolved independently of each other from an ancestral non-complexed form of allophanate hydrolase or whether one complex evolved in the presence of the other, 'co-opting' allophanate hydrolase into a new physiological role by forming new protein:protein interactions.

7.4 The role of metal-specificity in the evolution of new enzyme function

The metal centres of both TrzN and AtzA are atypical when compared with the majority of amidohydrolase family proteins; AtzA has an unusually 'loose' binding site (a high K_D probably resulting from long bond lengths between the metal and its ligands [4]) and the typical aspartate ligand has been replaced by a threonine residue in TrzN [15]. Unusual metal-binding centres are also observed in other members of the amidohydrolase family that have evolved recently in response to the recent increase in abundance of anthropogenic contaminants (e.g. the molanate and phenyl urea hydrolases, paraoxonase-1, MolA and PuhB, respectively [16]). Why, then, are newly evolved enzymes particularly prone to possessing unusual metal centres?

As discussed in the Introduction, new enzymatic functions evolve from the non-physiological, promiscuous catalytic functions present in pre-existing enzymes [17-19]. Some metalloenzymes have promiscuous activities that can be accessed by altering the identity of the cation in the metal centre. For example, replacing the physiologically relevant Mg^{2+} from dihydroxyacetone kinase [20] with Mn^{2+} alters its activity from that of a dihydroxyacetone kinase to that of a flavin adenine dinucleotide (FAD) cyclase (producing riboflavin 4',5'-cyclic phosphate) [20]. Similarly, the *N*-succinyl-L, L-diaminopimelate desuccinylase from *Salmonella enterica* is a Zn^{2+} -dependent enzyme, but shows promiscuous aspartyl dipeptidase activity if the Zn^{2+} metal is replaced with Mn^{2+} [21]. In recent work from Dr. Nobuhiko Tokuriki's laboratory, the metal content of five functionally distinct members of the metallo- β -lactamase family was altered, replacing the native metal

with Cd^{2+} , Co^{2+} , Fe^{2+} , Mn^{2+} , Ni^{2+} or Zn^{2+} . For each of the proteins, substituting a 'native' (or preferred) metal for a non-preferred metal altered the profile of catalytic activities (both physiological and promiscuous) [22]. Moreover, in some cases, cryptic promiscuous activities were uncovered [22].

There are number of mechanisms that can be proposed for the altered substrate specificities for enzyme isoforms. Different metals will have different coordination geometries with their ligands, for example Zn^{2+} is known to have predominantly tetrahedral coordination geometry whereas Co^{2+} and Fe^{2+} have octahedral geometry. These different coordination geometries may result in permit the binding of non-physiological substrates *via* altered substrate:metal interactions. Additionally, altered binding co-ordination of physiological substrates may alter the position of the substrate relative to the catalytic centre, altering the reaction catalysed. The pK_a of metal or co-ordinated water differs for different metal centres too, for example the pK_a of water bound to Fe^{2+} is 8.4, whereas for Zn^{2+} it is 9.6 [23]. Such changes in the pK_a of the reactive centre will also promote shifts in substrate specificity upon altering the metal centre of the enzyme.

It was hypothesized that subpopulations of enzyme 'isoforms' may be present in the cell, with the preferred metal presented in the largest subpopulation and less preferred metals in smaller subpopulations. Non-physiological, and otherwise cryptic, promiscuous activities would be accessible in such a mixed population. Moreover, selection could operate on these activities, perhaps selecting for mutations that promote a greater abundance of an enzyme isoform that confer a fitness advantage through the associated 'cryptic' activity. It is

plausible that mutations that enhance the abundance of particular enzyme isoforms would deviate from the 'canonical' metal-binding centre. If so, this provides a mechanism that explains the presence of unusual metal-binding centres in metalloenzymes with newly evolved catalytic activities. This hypothetical mechanism should be amenable to interrogation by empirical experimentation. Interestingly, during the purification of AtzB, mononuclear Fe^{2+} and Zn^{2+} centres were found to be present in this enzyme in equal abundance [24]. AtzB is known to perform a promiscuous dechlorination reaction in addition to its physiological deamination reaction [24]. It is yet to be determined whether the two abundant metalloisoforms of AtzB possess differing catalytic profiles. If so, AtzB may be an ideal candidate for probing the evolutionary significance of metalloisoforms as an enabling the acquisition of new enzymatic activities.

References

1. 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., X-ray structure of the amidase domain of atzf, the allophanate hydrolase from the cyanuric acid-mineralizing multienzyme complex. *Applied and Environmental Microbiology*, 2015. **81**: p. 470-480.
2. Balotra, S., Newman, J., French, N.G., Briggs, L.J., Peat, T.S., and Scott, C., Crystallization and preliminary x-ray diffraction analysis of the amidase domain of allophanate hydrolase from *pseudomonas* sp strain adp. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*, 2014. **70**: p. 310-315.
3. Peat, T.S., Balotra, S., Wilding, M., French, N.G., Briggs, L.J., Panjikar, S., Cowieson, N., Newman, J., and Scott, C., Cyanuric acid hydrolase: Evolutionary innovation by structural concatenation. *Molecular Microbiology*, 2013. **88**: p. 1149-1163.
4. Peat, T.S., Newman, J., Balotra, S., Lucent, D., Warden, A.C., and Scott, C., The structure of the hexameric atrazine chlorohydrolase atza. *Acta Crystallogr D Biol Crystallogr*, 2015. **71**: p. 710-720.
5. Renata, H., Wang, Z.J., and Arnold, F.H., Expanding the enzyme universe: Accessing non-natural reactions by mechanism-guided directed evolution. *Angew Chem Int Ed Engl*, 2015. **54**: p. 3351-3367.
6. Noor, S., Taylor, M.C., Russell, R.J., Jermiin, L.S., Jackson, C.J., Oakeshott, J.G., and Scott, C., Intramolecular epistasis and the evolution of a new enzymatic function. *PLoS One*, 2012. **7**.

7. Siripattanakul, S., Wirojanagud, W., McEvoy, J., Limpiyakorn, T., and Khan, E., Atrazine degradation by stable mixed cultures enriched from agricultural soil and their characterization. *Journal of Applied Microbiology*, 2009. **106**: p. 986-992.
8. Noor, S., Changey, F., Oakeshott, J.G., Scott, C., and Martin-Laurent, F., Ongoing functional evolution of the bacterial atrazine chlorohydrolase atza. *Biodegradation*, 2014. **25**: p. 21-30.
9. Scott, C., Jackson, C.J., Coppin, C.W., Mourant, R.G., Hilton, M.E., Sutherland, T.D., Russell, R.J., and Oakeshott, J.G., Catalytic improvement and evolution of atrazine chlorohydrolase. *Applied and Environmental Microbiology*, 2009. **75**: p. 2184-2191.
10. Arbeli, Z. and Fuentes, C., Prevalence of the gene *trzn* and biogeographic patterns among atrazine-degrading bacteria isolated from 13 colombian agricultural soils. *Fems Microbiology Ecology*, 2010. **73**: p. 611-623.
11. Udikovic-Kolic, N., Scott, C., and Martin-Laurent, F., Evolution of atrazine-degrading capabilities in the environment. *Applied Microbiology and Biotechnology*, 2012. **96**: p. 1175-1189.
12. Kim, H.S., Mittenenthal, J.E., and Caetano-Anolles, G., Widespread recruitment of ancient domain structures in modern enzymes during metabolic evolution. *J Integr Bioinform*, 2013. **10**.
13. Nagano, N., Orengo, C.A., and Thornton, J.M., One fold with many functions: The evolutionary relationships between tim barrel families based on their sequences, structures and functions. *Journal of Molecular Biology*, 2002. **321**: p. 741-765.
14. Whitney, P.A. and Cooper, T.F., Urea carboxylase and allophanate hydrolase - 2 components of adenosine triphosphate urea amidolyase in *saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 1972. **247**: p. 1349-1353.
15. Seffernick, J.L., Reynolds, E., Fedorov, A.A., Fedorov, E., Almo, S.C., Sadowsky, M.J., and Wackett, L.P., X-ray structure and mutational analysis of the atrazine chlorohydrolase trzn. *Journal of Biological Chemistry*, 2010. **285**: p. 30606-30614.
16. Sugrue, E., Fraser, N.J., Hopkins, D.H., Carr, P.D., Khurana, J.L., Oakeshott, J.G., Scott, C., and Jackson, C.J., Evolutionary expansion of the amidohydrolase superfamily in bacteria in response to the synthetic compounds molinate and diuron. *Appl Environ Microbiol*, 2015. **81**: p. 2612-2624.
17. Copley, S.D., An evolutionary biochemist's perspective on promiscuity. *Trends Biochem Sci*, 2015. **40**: p. 72-78.
18. Khersonsky, O. and Tawfik, D.S., Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu Rev Biochem*, 2010. **79**: p. 471-505.
19. Verdel-Aranda, K., Lopez-Cortina, S.T., Hodgson, D.A., and Barona-Gomez, F., Molecular annotation of ketol-acid reductoisomerases from *streptomyces* reveals a novel amino acid biosynthesis interlock mediated by enzyme promiscuity. *Microbial Biotechnology*, 2015. **8**: p. 239-252.
20. Sanchez-Moreno, I., Iturrate, L., Martin-Hoyos, R., Jimeno, M.L., Mena, M., Bastida, A., and Garcia-Junceda, E., From kinase to cyclase: An unusual example of catalytic promiscuity modulated by metal switching. *ChemBiochem*, 2009. **10**: p. 225-229.
21. Broder, D.H. and Miller, C.G., Dape can function as an aspartyl peptidase in the presence of Mn^{2+} . *Journal of Bacteriology*, 2003. **185**: p. 4748-4754.
22. Baier, F., Chen, J., Solomonson, M., Strynadka, N., C.J., and Tokuriki, N., Distinct metal isoforms underlie promiscuous activity profiles of metalloenzymes. *ACS chemical biology*, 2015: p. in press.
23. Frey, P.A. and Hegeman, A.D., *Enzymatic reaction mechanisms*. 2007, Oxford ; New York: Oxford University Press.
24. Seffernick, J.L., Aleem, A., Osborne, J.P., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Hydroxyatrazine n-ethylaminohydrolase (atzb): An amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *Journal of Bacteriology*, 2007. **189**: p. 6989-6997.

8. Appendices

8.1 Sequences of genes used in this study (3' → 5')

Atrazine chlorohydrolase (*atzA*)

ATGCAAACGCTCAGCATCCAGCACGGTACCCTCGTCACGATGGATCAGTACCGCAGAGTCCTTGGG
GATAGCTGGGTTACGTGCAGGATGGACGGATCGTCGCGCTCGGAGTGCACGCCGAGTCGGTGCCT
CCGCCAGCGGATCGGGTGATCGATGCACGCGGCAAGGTCGTGTTACCCGGTTTCATCAATGCCAC
ACCCATGTGAACCAGATCCTCCTGCGCGGAGGGCCCTCGCACGGGCGTCAATTCTATGACTGGCTGT
TCAACGTTGTGTATCCGGGACAAAAGGCGATGAGACCGGAGGACGTAGCGGTGGCGGTGAGGTTG
TATTGTGCGGAAGCTGTGCGCAGCGGGATTACGACGATCAACGAAAACGCCGATTCGGCCATCTAC
CCAGGCAACATCGAGGCCGCGATGGCGGTCTATGGTGAGGTGGGTGTGAGGGTCGTCTACGCCG
CATGTTCTTTGATCGGATGGACGGGCGCATTCAAGGGTATGTGGACGCCTTGAAGGCTCGCTCTCCC
CAAGTCGAACTGTGCTCGATCATGGAGGAAACGGCTGTGGCCAAAGATCGGATCACAGCCCTGTCA
GATCAGTATCATGGCACGGCAGGAGGTCGTATATCAGTTTGGCCCGCTCCTGCCACTACCACGGCG
GTGACAGTTGAAGGAATGCGATGGGCACAAGCCTTCGCCCCTGATCGGGCGGTAATGTGGACGCTT
CACATGGCGGAGAGCGATCATGATGAGCGGATTCATGGGATGAGTCCCGCCGAGTACATGGAGTG
TTACGGACTCTTGATGAGCGTCTGCAGGTCGCGCATTGCGTGTACTTTGACCGGAAGGATGTTCCG
CTGCTGCACCGCCACAATGTGAAGGTCGCGTCGCAGGTTGTGAGCAATGCCTACCTCGGCTCAGGG
GTGGCCCCCGTGCCAGAGATGGTGGAGCGCGGCATGGCCGTGGGCATTGGAACAGATAACGGGAA
TAGTAATGACTCCGTAAACATGATCGGAGACATGAAGTTTATGGCCCATATTCACCGCGCGGTGCAT
CGGGATGCGGACGTGCTGACCCAGAGAAGATTCTTGAAATGGCGACGATCGATGGGGCGCGTTC
GTTGGGAATGGACCACGAGATTGGTTCCATCGAAACCGGCAAGCGCGCGGACCTTATCCTGCTTGA
CCTGCGTCACCCTCAGACGACTCCTCACCATCATTTGGCGGCCACGATCGTGTTTCAGGCTTACGGCA
ATGAGGTGGACACTGTCCTGATTGACGGAAACGTTGTGATGGAGAACCGCCGCTTGAGCTTTCTCC

CCCTGAACGTGAGTTGGCGTTCCTTGAGGAAGCGCAGAGCCGCGCCACAGCTATTTTGCAGCGGGC
GAACATGGTGGCTAACCCAGCTTGGCGCAGCCTCTAG

Hydroxyatrazine hydrolase (*atzB*)

ATGACCACCACTCTTTACACCGGGTTTCACCAGCTGGTCACCGGGGATGTCGCGGGCACGGTGCTTA
ATGGCGTCGATATTCTGGTGAGAGACGGGGAGATTATTGGCCTTGGCCCCGATCTGCCCCGAACTC
TGGCCCCAATTGGCGTTGGCCAGGAACAGGGTGTGAGGTGGTGAAGTCCCGGGGGCTGACCGCC
TATCCGGGGCTGATCAACACCCATCACCATTTTTTTTCAGGCCTTTGTCCGCAATCTTGCCCCCTCGAC
TGGACCAGCTTGATGTGCTGGCCTGGCTGCGCAAGATCTACCCGGTGTGGCCTGGTGGATGAA
GACTGCATCTACCACAGCACAGTGGTGTCCATGGCGGAGCTGATCAAGCACGGCTGCACCACCGCG
TTTGATCACCAGTACAACACTACAGCCGCCGGGGTGGACCCTTTCTGGTGGACCGCCAGTTTGATGCCG
CTAACCTGCTGGGGCTGCGTTTTACGCCGGGCGCGGCTGCATCACCTGCCGATGGCGGAGGGCA
GCACCATTCCCGATGCCATGCGGGAGAGCACCGATACTTTTCTTGCCGATTGCGAACGCCTGGTGAG
CCGCTTTCATGACCCGCGGCCCTTCGCCATGCAGCGGGTGGTGGTGGCGCCCAGCTCGCCGGTGATT
GCCTACCCGGAAACCTTCGTGAGTCTGCCCGTCTGGCCCGCCACCTGGGGGTGAGCCTGCACACCC
ATCTGGGGGAGGGGGAAACCCCGCCATGGTGGCCCGTTTTGGCGAGCGCAGCCTCGACTGGTGC
GAAAACCGGGGCTTTGTGCGGGCCGGATGTGTGGCTCGCCACGGCTGGGAATTCACCGCAGCGGA
CATCGCCCGCCTTGCCGCCACCGGCACCGGTGTTGCCATTGCCCGCGCCGGTGTCTGGTGGGT
GCCGAAGTCACCGACATTCCCGCCATGGCGGCGGCAGGTGTGCGGGTGGGCTTTGGGGTGGATGG
CCACGCCTCCAACGACAGCTCGAACCTCGCCGAGTGCATCCGCCTTGCTACCTGCTCCAGTGCCTG
AAGGCCAGTGAGCGCCAGCATCCGGTGCCAGCTCCCTATGACTTTCTGCGCATGGCCACCCAGGGT
GGTGC GGATTGCCTCAACCGGCCTGATCTCGGCGCCCTCGCCGTGGGCAGGGCTGCGGATTTCTTTG
CCGTCGATCTGAACCGGATCGAATACATCGGCGCCAACCACGATCCACGGAGCCTGCCCCGAAAAG

TGGGTTTTTCCGGCCCCGTGGACATGACCGTGATCAACGGCAAGGTGGTGTGGCGCAACGGTGAAT
TTCCGGGCCTTGATGAAATGGAGCTGGCCCCGGGCTGCGGACGGGGTGTTCGCCGGGTGATTACG
GCGATCCTTTGGTGGCAGCCCTTCGACGGGGCACCGGTGTCACACCATGCTGA

***N*-isopropylammelide isopropyl amidohydrolase (*atzC*)**

ATGAGTAAAGATTTTGATTTAATCATTAGAAACGCCTATCTAAGTGAAAAAGACAGTGTATATGATA
TTGGGATTGTTGGTGACAGAATAATCAAATAGAAGCTAAAATTGAAGGAACCGTAAAAGACGAAA
TTGATGCAAAGGGTAACCTTGTGTCTCCCGGATTTGTCGATGCACATACCCATATGGATAAGTCATT
ACGAGCACAGGTGAAAGATTACCGAAGTTTTGGAGCAGACCTTATACAAGGGATGCTGCCATCGAG
GATGGCTTGAAATATTATAAAAATGCTACCCACGAAGAAATAAAAAGACATGTGATAGAACATGCTC
ACATGCAGGTACTIONCATGGGACTTTATACACCCGGACCCATGTAGATGTAGATTGATTGCTAAAAC
AAAAGCAGTGAAGCAGTTTTAGAAGCCAAGGAAGAGTTAAAGGATCTTATCGATATAAAGTCGT
AGCCTTTGCACAGAGTGGATTTTTCGTTGATTTGGAATCTGAATCATTGATTAGAAAATCCTTGGATA
TGGGCTGTGATTTAGTTGGGGGAGTTGATCCTGCTACGCGGGAAAATAATGTTGAGGGTTCTTTAG
ACCTATGCTTTAAATTAGCAAAGGAATACGATGTTGATATCGACTATCACATACATGATATTGGAAC
GTTGGAGTATATTCGATAAATCGTCTTGCCCAAAGACAATTGAAAATGGGTATAAGGGTAGAGTA
ACTACGAGTCATGCCTGGTGTGTTTTGCAGATGCTCCGTCCGAATGGCTCGATGAGGCAATCCCATTGT
ACAAGGATTCGGGTATGAAATTTGTTACCTGTTTTAGTAGTACACCGCCTACTATGCCGGTGATAAA
GCTGCTTGAAGCTGGCATCAATCTTGGCTGTGCTTCGGACAATATCAGAGATTTTTGGGGTCCCTTTG
GCAACGGTGATATGGTACAAGGGGCTCTGATCGAAACTCAGAGATTAGAGTTAAAGACAAACAGA
GATTTGGGACTAATTTGGAAAATGATAACGTCAGAGGGTGCTAGAGTTTTAGGAATTGAAAAGAAC
TATGGGATAGAAGTTGGTAAAAAGGCCGATCTTGTTGATTAATTCGTTGTCACCACAATGGGCAA

TAATCGACCAAGCAAAAAGACTATGCGTAATTA AAAATGGACGTATCATTGTGAAGGATGAGGTTA
TAGTTGCCTAA

Cyanuric acid hydrolase (*atzD*)

ATGTATCACATCGACGTTTTCCGAATCCCTTGCCACAGCCCTGGTGATACATCGGGTCTCGAGGATT
GATTGAAACAGGCCGCGTTGCCCCGCCGACATCGTCGCGGTAATGGGCAAGACCGAGGGCAATG
GCTGCGTCAACGATTACACGCGTGAATACGCCACCGCCATGCTTGCTGCGTGCCTGGGCGTCATT
GCAACTCCCACCCATGAGGTGGAAAAGCGGGTCGCGTTTGTGATGTCAGGTGGGACGGAAGGCG
TGCTGTCCCCCACCACACGGTATTGCAAGACGTCCGGCAATCGACGCGCATCGTCCCGCTGGCAA
ACGTCTCACGCTTGGAATCGCCTTCACGCGTGATTTTCTGCCGGAGGAAATTGGCCGCCACGCTCAG
ATAACGGAGACAGCCGGCGCCGTCAAACGCGCAATGCGAGATGCCGGGATCGCTTCGATTGACGAT
CTGCATTTTGTGCAGGTGAAGTGTCCGCTGCTGACACCAGCAAAGATCGCCTCGGCGCGATCACGC
GGATGCGCTCCAGTCACGACGGATACGTATGAATCGATGGGCTATTCGCGCGGCGCTTCGGCCCTG
GGCATCGCTCTCGCTACAGAAGAGGTGCCCTCCTCGATGCTCGTAGACGAATCAGTGCTGAATGACT
GGAGTCTCTCATCGTCACTGGCGTCGGCGTCTGCAGGCATCGAACTGGAGCACAACGTGGTGATCG
CTATTGGCATGAGCGAGCAGGCCACCAGTGA ACTGGTCATTGCCACGGCGTGATGAGCGACGCGA
TCGACGCGGCCTCGGTGCGGCGAACGATTGAATCGCTGGGCATACGTAGCGATGACGAGATGGAT
CGCATCGTCAACGTATTCGCCAAAGCGGAGGCGAGCCCGGACGGGGTTGTACGAGGTATGCGGCA
CACGATGCTAAGTGACTCCGACATTAATTCGACCCGCCATGCGCGGGCGGTCACCGGCGCGGCCAT
TGCCTCGGTAGTTGGGCATGGCATGGTGTATGTGTCCGGTGGCGCCGAGCATCAGGGACCTGCCGG
CGGCGGCCCTTTTGCAGTCATTGCCCGCGCTTAA

Allophanate hydrolase (*atzF*)

ATGAATGACCGCGCGCCCCACCCTGAAAGATCTGGTCGAGTCACGCCGGATCACCTGACCGATCTG
GCTTCCTATCAGGCTGCCTATGCCGCCGGTACAGACGCCGCCGACGTCATTCGGACCTGTATGCC
GTATCAAAGAAGACGGCGAAAATCCGATCTGGATTAGCCTGTTGCCCTTGAAAGCGCATTGGCGA
TGCTGGCCGACGCGCAGCAACGCAAGGACAAGGGAGAAGCGTTGCCGCTCTTTGGCATCCCCTTCG
GCGTCAAGGACAACATCGACGTCGAGGCCTTCGACGACTGCCGGGTGTACGGGGTTCGCGCGTA
CGCCCCGACAGCACGCCTTCGTCTACAGCGCCTGGTGGACGCTGGCGCGATCCCGATCGGAAAA
CGAACCTCGATCAATTCGCGACCGGGTTGAACGGCACTCGCACGCCGTTTGGCATTCCGCGCTGCGT
GTTCAACGAGAACTACGTATCCGGCGGCTCCAGCAGTGGCTCCGCAGTGGCCGTCGCCAACGGCAC
GGTACCGTTCTCGCTCGGGACGGACACTGCCGGTTCGGGCCATTCTGCTGCGTTCAACAATCTG
GTGGGCTTGAAACCGACCAAAGGCCTGTTCTCGGGCAGTGGACTGGTTCGCGGGCGCGAAGCCTT
GACTGCATCAGCGTCCTCGCCATACCGTAGATGACGCCCTTGGGTGCGCACGCGTCGCCGCCGGCT
ACGATGCTGATGACGCTTTTTCGCGCAAGGCGGGCGCCGCCACTGACAGAAAAGAGTTGGCCTC
GTCGCTTCAATTCGGGGTCCCAGCGGCGGAACATCGCCAGTTTTTCGGTGACGCGGAAGCCGAGG
CGTTTTCAATAAAGCGGTCGCAAGCTTGAAGAGATGGGTGGCACCTGCATCTCGTTTACTATAAC
CCCCTTCAGGCAGGCTGCTGAACTGCTCTACGCCGGCCCTGGGTGCGGAGCGCCTGGCGGCCATC
GAGAGCCTTGGGACGAGCATCCCGAGGTGCTCCACCCGTCGTTCTGTGACATCATCTTGTCCGCGA
AGCGAATGAGCGCAGTCGACACGTTCAACGGTATCTATCGCTGGCCGACCTTGTGAGGGCTGCAG
AGAGCACTTGGGAAAAGATCGATGTGATGCTGCTGCCGACGGCGCCGACCATCTACACTGTAGAAG
ACATGCTCGCCGATCCGGTACGCCTCAACAGCAATCTGGGCTTCTACACGAACTTCGTGAACTTGAT
GGATTTGTCCGCGATTGCTGTTCCCGCAGGCTTCCGAACCAATGGCCTGCCATTTGGCGTCACTTTCA
TCGGTCGGGCGTTCGAAGATGGGGCGATCGCAAGCTTGGGAAAAGCTTTCGTGGAGCACGACCTC
GCCAAGGGCAACGCGGCCACGGCGGCCACCCAAGGATACCGTCGCAATCGCCGTGGTAGGTGC
ACATCTCTCCGACCAGCCCTTGAATCATCAGCTCACGGAGAGCGGCCGAAAGCTACGGGCAACAAC

CGTACTGCGCCGGGATATGCCTTGTACGCACTCCGTGATGCGACGCCGGCTAAGCCTGGAATGTT
GCGCGACCAGAATGCGGTCGGGAGCATCGAAGTGGAATCTGGGATCTGCCGGTCGCCGGGTTCG
GTGCGTTTGTAAGTGAAATTCCGGCGCCGTTGGGTATCGGGACAATAACACTCGAAGACGGCAGCC
ATGTGAAAGGCTTTCTGTGCGAGCCACATGCCATCGAGACGGCGCTCGACATCACTCACTACGGCG
GCTGGCGAGCATACCTCGCGGCTCAATAG

8.2 Amino acid sequences of the enzymes investigated in this study

Atrazine chlorohydrolase (AtzA)

MQTLSIQHGTLVTMQYRRVLGDSWVHVQDGRIVALGVHAESVPPPADRVIDARGKVVLPGFINAHT
HVNQILLRGGPSHGRQFYDWLFNVVYPGQKAMRPEDVAVAVRLYCAEAVRSGITTINENADSAIYPGNI
EAAMAVYGEVGVVYARMFFDRMDGRIQGYVDALKARSPQVELCSIMEETAVAKDRITALSDQYHGT
AGGRISVWPAPATTTAVTVEGMRWAQAFARDRAVMWTLHMAESDHDERIHGMSPAEMYECYGLLD
ERLQVAHCVFYDRKDVRLLRHNVKVASQVVSNAYLGSVAPVPEMVERGMAVGIGTDNGNSNDSV
NMIGDMKFMMAHIHRAVHRDADVLTPKILEMATIDGARSLGMDHEIGSIETGKRADLILLDLRHPQTP
HHHLAATIVFQAYGNEVDTVLIDGNVVMENRRLSFLPPERELAFLEEAQSRATAILQRANMVANPAWRS
L

Hydroxyatrazine hydrolase (AtzB)

MTTTLTYGFHQLVTGDVAGTVLNGVDILVRDGEIIGLGPDLRPTLAPIGVGQEQGVEVVNCRGLTAYPGL
INTHHHFFQAFVRNLAPLDWTQLDVLAWLRKIYPVFALVDEDCIYHSTVVSMALIKHGCTTAFDHQYN
YSRRGGPFLVDRQFDAANLLGLRFHAGRGCITLPAEGSTIPDAMRESTDTFLADCERLVSFRHDPFPFA
MQRVVVAPSSPVIAYPETFVESARLARHLGVSLSLHHLGEGETPAMVARFGERSLDWCENRGFVGPDVW
LAHGWEFTAADIARLAATGTGVAHCPAPVFLVGAEVTDIPAMAAAGVVRVGFVGDGHASNDSSNLAECI
RLAYLLQCLKASERQHPVPAPYDFLRMATQGGADCLNRPDLGALAVGRAADFFAVDLNRIEYIGANHDP
RSLPAKVGFSGPVDMTVINGKVVWRNGEFPLDEMELARAADGVFRRVIYGDPLVAALRRGTGVTPC

***N*-isopropylammelide isopropyl amidohydrolase (AtzC)**

MSKDFDLIRNAYLSEKDSVYDIGIVGDRIIKIEAKIEGTVKDEIDAKGNLVSPGFVDAHATHMDKSFTST
GERLPKFWSRPYTRDAAIEDGLKYYKNATHEEIKRHVIEHAHMQVLHGTLYTRTHVDVDSVAKTKAVEA
VLEAKEELKDLIDIQVVAFAQSGFFVDLESESLIRKSLDMGCDLVGGVDPATRENNVEGSLDLCFKLAKEY
DVIDIDYHIHDIGTVGVYSINRLAQKTIENGYKGRVTTSHAWCFADAPSEWLDEAIPLYKDSGMKFVTCFS
STPPTMPVIKLEAGINLGCASDNIRDFWVPPFGNGDMVQGALLETQRLELKTNRDLGLIWKMITSEGAR
VLGIEKNYGIEVGKKADLVVLSLSPQWAIIDQAKRLCVIKNGRIIVKDEVIVA

Cyanuric acid hydrolase (AtzD)

MYHIDVFRIPCHSPGDTSGLEDLIETGRVAPADIVAVMGKTEGNGCVNDYTREYATAMLAACLGRHLQL
PPHEVEKRVAFFVMSGGTEGVLSPHHTVFARRPAIDHRPAGKRLTLGIAFTRDFLPEEIGRHAQITETAGA
VKRAMRDAGIASIDDLHFVQVKPLLTPAKIASARSRGCAPVTTDTYESMGYSRGASALGIALATEEVPSS
MLVDESVLNDWSLSSSLASASAGIELEHNVVIAIGMSEQATSELVIAHGVMSDAIDAASVRRRTIESLGIRS
DDEMDRIVNVFAKAEASPDGVVRGMRHTMLSDSDINSTRHARAVTGAAIASVVGHRGMVYVSGGAEH
QGPAGGGPFAVIARA

Allophanate hydrolase (AtzF)

MNDRAPHPERSONGRVTPDHLTDLASYQAAYAAGTDAADVISDLYARIKEDGENPIWISLLPLESALAMLAD
AQRKDKGEALPLFGIPFGVKDNIDVAGLPTTAGCTGFARTPRQHAFVVQRLVDAGAIPIGKTNLDQFA
TGLNGTRTPFGIPRCVFNENYVSGGSSSGSAVAVANGTVPFSLGTDTAGSGRIPAAFNNLVGLKPTKGLF
SGSGLVPAARSLDCISVLAHTVDDALAVARVAAGYDADDAFSRKAGAAALTEKSWPRRFNFGVPAAEHR

QFFGDAEAEALFNKAVRKLEEMGGTCISFDYTPFRQAAELLYAGPWVAERLAAIESLADEHPEVLHPVVR
DIILSAKRMSAVDTFNNGIYRLADLVRAAESTWEKIDVMLLPTAPTIIYTVEDMLADPVRNLNSNLGFYTNFVN
LMDLSAIAVPAGFRTNGLPFGVTFIGRAFEDGAIASLGKAFVEHDLAKGNAATAAPPKDTVAIAVVG AHL
SDQPLNHQLTESGGKLRATTRTAPGYALYALRDATPAKPGMLRDQNAVGSIEVEIWDLPVAGFGAFVSE
IPAPLGIGTITLEDGSHVKGFLCEPHAIETALDITHYGGWRAYLAAQ

8.3 List of PDBe entries from this study

Enzyme	PDBe numbers	Description
Atrazine chlorohydrolase (AtzA)	4V1X,	2.2 Å resolution structure with six Fe ³⁺ and six di(hydroxyethyl)ether molecules bound as ligands.
	4V1Y	2.8 Å resolution structure with six Fe ³⁺ , one Cl ⁻ , six 1,2-ethanediol and three molecules of di(hydroxyethyl)ether.
<i>N</i> -isopropylammelide isopropylaminohydrolase (AtzC)	4CQB,	1.84 Å resolution structure with two Zn ²⁺ , two Cl ⁻ , three molecules of dimethyl sulfoxide and three molecules of malonate ion.
	4CQC,	2.2 Å resolution structure with two Zn ²⁺ , two Cl ⁻ , and two molecules of 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid.
	4CQD,	2.25 Å resolution structure with two Zn ²⁺ , two Cl ⁻ , two molecules of 2-(<i>N</i> -morpholino)-ethanesulphonic acid and two copies of malonate ion.
	5AKQ	2.6 Å resolution structure with two Zn ²⁺ and two Cl ⁻ .
Cyanuric acid hydrolase (AtzD)	4BVQ,	1.9 Å resolution structure with four PO ₄ ³⁻ , four Mg ²⁺ , and eight molecules of Di(hydroxyethyl)ether.
	4BVR,	2.58 Å resolution structure with four Mg ²⁺ , four molecules of cyanuric acid, and four copies of di(hydroxyethyl)ether.
	4BVS,	2.6 Å resolution structure with four Mg ²⁺ and four copies of melamine.
	4BVT	3.1 Å resolution structure with four Mg ²⁺ , four molecules of barbituric acid and four copies of di(hydroxyethyl)ether.

<p>Amidase domain of allophanate hydrolase (AtzF)</p>	<p>4CP8</p>	<p>2.5 Å resolution structure with two copies of malonate ion.</p>
-------------------------------------------------------	-------------	--------------------------------------------------------------------

8.4 List of known inhibitors and substrates for the enzymes of the atrazine catabolic pathway

Enzyme	Inhibitors	Substrates	Source
AtzA	1,10-phenanthroline, aminotriazine, Cu ²⁺ , Ni ²⁺ , Zn ²⁺ , Fe ³⁺ , oxalic acid.	atrazine, deethylatrazine, deisopropylatrazine, simazine, terbuthylazine	[1-3]
AtzB	1,10-phenanthroline, 8-hydroxyquinoline-5-sulfonic acid, dipicolinic acid, chelex-100 resin.	2-(<i>N</i> -ethylamino)-4-hydroxy-6-(<i>N</i> -isopropylamino)-1,3,5-triazine , 2,4-dimethylamino-6-hydroxy-1,3,5-triazine , 2,4-diethylamino-6-hydroxy-1,3,5-triazine , 2,4-diisopropylamino-6-hydroxy-1,3,5-triazine , 2,4-di(<i>N</i> -isobutylamino)-6-hydroxy-1,3,5-triazine , 2,4-di(<i>N</i> -secbutylamino)-6-hydroxy-1,3,5-triazine , 2,4-di(<i>N</i> -terbutylamino)-6-hydroxy-1,3,5-triazine , 4-(<i>N</i> -ethylamino)-6-hydroxy-2-(<i>N</i> -methylamino)-1,3,5-triazine , 6-hydroxy-4-(<i>N</i> -isopropylamino)-2-(<i>N</i> -methylamino)-1,3,5-triazine , 2,4-diamino-6-hydroxy-1,3,5-triazine , 2-amino-6-(<i>N</i> -ethylamino)-4-hydroxy-1,3,5-triazine , 2-amino-4-hydroxy-6-(<i>N</i> -isopropylamino)-1,3,5-triazine , 2-hydroxy-4,6-di(<i>N</i> -hydroxyethylamino)-1,3,5-triazine , 2-(<i>N</i> -ethylamino)-4-hydroxy-6-(<i>N</i> -hydroxyethylamino)-1,3,5-triazine , 2-amino-4-hydroxy-6-(<i>N</i> -hydroxyethylamino)-1,3,5-triazine , 2-hydroxy-4-(<i>N</i> -isopropylamino)-6- <i>N</i> -(3-methoxypropylamino)-1,3,5-triazine , 2-(<i>N</i> -ethyl- <i>N</i> -methylamino)-4-ethylamino-6-hydroxy-1,3,5-triazine , 2-amino-4-chloro-6-hydroxy-	[4]

		1,3,5-triazine , 2-chloro-4-(<i>N</i> -ethylamino)-6-hydroxy-1,3,5-triazine , 2-chloro-4-hydroxy-6-(<i>N</i> -isopropylamino)-1,3,5-triazine	
AtzC	1,10-phenanthroline, 8-hydroxyquinoline-5-sulfonic acid, EDTA.	ammelide, <i>N</i> -dimethylammelide <i>N</i> -methylammelide, <i>N</i> -ethylammelide, <i>N</i> -hydroxyethylammelide, <i>N</i> -isopropylammelide, <i>N</i> -cyclopropylammelide, <i>N</i> - <i>t</i> -butylammelide,	[5]
AtzD	barbituric acid, <i>O</i> -phenanthroline, PMSF.	cyanuric acid, <i>N</i> -methylisocyanuric acid	[6-8]
AtzF	phenyl phosphorodiamidate	allophanate, malonamate, biuret	[9, 10]

References

1. Seffernick, J.L., McTavish, H., Osborne, J.P., de Souza, M.L., Sadowsky, M.J., and Wackett, L.P., Atrazine chlorohydrolase from *Pseudomonas* sp strain ADP is a metalloenzyme. *Biochemistry*, 2002. **41**: 14430-14437.
2. Seffernick, J.L., de Souza, M.L., Sadowsky, M.J., and Wackett, L.P., Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *Journal of Bacteriology*, 2001. **183**: 2405-2410.
3. Seffernick, J.L., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Substrate specificity of atrazine chlorohydrolase and atrazine-catabolizing bacteria. *Applied and Environmental Microbiology*, 2000. **66**: 4247-4252.
4. Seffernick, J.L., Aleem, A., Osborne, J.P., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Hydroxyatrazine *N*-ethylaminohydrolase (AtzB): An amidohydrolase superfamily enzyme catalyzing deamination and dechlorination. *Journal of Bacteriology*, 2007. **189**: 6989-6997.
5. Shapir, N., Osborne, J.P., Johnson, G., Sadowsky, M.J., and Wackett, L.P., Purification, substrate range, and metal center of AtzC: The *N*-isopropylammelide aminohydrolase involved in bacterial atrazine metabolism. *Journal of Bacteriology*, 2002. **184**: 5376-5384.
6. Karns, J.S., Gene sequence and properties of an *s*-triazine ring-cleavage enzyme from *Pseudomonas* sp strain NRRLB-12227. *Applied and Environmental Microbiology*, 1999. **65**: 3512-3517.
7. Peat, T.S., Balotra, S., Wilding, M., French, N.G., Briggs, L.J., Panjikar, S., Cowieson, N., Newman, J., and Scott, C., Cyanuric acid hydrolase: Evolutionary innovation by structural concatenation. *Molecular Microbiology*, 2013. **88**: 1149-1163.
8. Fruchey, I., Shapir, N., Sadowsky, M.J., and Wackett, L.P., On the origins of cyanuric acid hydrolase: Purification, substrates, and prevalence of AtzD from *Pseudomonas* sp strain ADP. *Applied and Environmental Microbiology*, 2003. **69**: 3653-3657.
9. Shapir, N., Sadowsky, M.J., and Wackett, L.P., Purification and characterization of allophanate hydrolase (AtzF) from *Pseudomonas* sp strain ADP. *Journal of Bacteriology*, 2005. **187**: 3731-3738.
10. Balotra, S., Newman, J., French, N.G., Briggs, L.J., Peat, T.S., and Scott, C., Crystallization and preliminary X-ray diffraction analysis of the amidase domain of allophanate hydrolase from *Pseudomonas* sp strain ADP. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications*, 2014. **70**: 310-315.