Structural and mechanistic studies of cyanuric acid hydrolase and biodegradation of 1,3,5-triazine

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF UNIVERSITY OF MINNESOTA

BY

Seunghee Cho

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

ADVISER: Lawrence P. Wackett

October, 2015

Acknowledgements

My graduate career has been possible because there is the gracious and kind support from my mentors, colleagues, and family. Firstly, I wish to thank my adviser, Lawrence P. Wackett, for his guidance in my research. With the best advices during graduate studies, he trained me in multiple fields covering enzymology, analytical methods, X-ray crystallography, and microbiology. I would like to thank Professor Carrie Wilmot. She kindly served as a committee chair for my thesis. Whenever I visited her office, she gives me thoughtful advices for resolving the problems. Professor Hideki Aihara is thanked for serving as a committee member and giving me great advice in collaboration for the crystallography works. He and his lab members kindly permit me to take advantage of all of the resources relating to the crystallography works. I also thank Professor Daniel Bonds for his insights about my research and serving as a committee member. I am very grateful for his helpful direction in writing my thesis. I am greatly thankful to Professor Arkady Khodursky who kindly served as a committee member for my thesis and reviewed my thesis. He always welcomes me and gives me helpful advices regarding problems I encountered during my graduate studies. I would like to thank Dr. Tony Dodge for teaching me nearly all of the experimental techniques from cloning to protein characterization with deep scientific insights. I also would like to thank Dr. Jennifer Seffernick for her insights on the whole research I have done. I am very grateful and thankful to Dr Jack Richman who synthesized essential chemicals for the experiments and analyzed samples with his great scientific insights. I have also appreciated Dr. Asim Bera for his advices on X-ray crystallography and teaching the techniques. I am grateful to Dr. Kelly Aukema for providing valuable input on my experiments. I am also thankful to all members of the Wackett group. Dr. Ke Shi also deserves special note for valuable insights and experimental support on structure determinations.

I also have appreciated Darlene Toedter, Sarah Dittrich, Sue Knoblauch, and Ann Johnson for consistent support during my graduate studies.

Dedication

This work is dedicated to my parents, all of my friends.

Abstract

Microorganisms have developed the greatest diversity of enzymatic pathways to allow their successful proliferation in the natural world. Those enzymatic pathways are the result of interactions between microorganisms and their habitat leading to adaptations by the microorganisms. As knowledge about the enzymatic pathways accumulates, increasingly those pathways can be used to degrade recalcitrant pollutants. The biodegradation pathway of *s*-triazine compounds is an example. Cyanuric acid hydrolase (CAH) is the one of the enzymes in the *s*-triazine degradation pathway. CAH is of particular interest because it hydrolyses cyanuric acid that is the by-product of tri-chlorocyanuric acid decomposition which has been used in swimming pool disinfection.

The first three chapters will deal with the mechanism of CAH mainly using X-ray crystallography. Chapter 2 describes the structural characteristics of the active site of CAH with emphasis on emerging evidence that a serine in domain B is the nucleophilic serine in the catalytic cycle. Evidence from mutational studies, Burgi-Dunitz angle analysis, and sequence alignments all support the hypothesis that the serine in domain B is the nucleophilic serine. Direct evidence is presented in chapter 3 for this serine acting as the nucleophile by our solving an X-ray structure showing a reaction intermediate covalently attached to this serine (the so-called “acyl intermediate”). The X-ray structure also shows 24 monomers in the crystal’s asymmetric unit and different active sites within the asymmetric unit contain different stages of the reaction such as the unreacted substrate, the acyl intermediate, carboxybiuret, biuret, and an empty active that has expelled the product.

The thesis concludes with chapter 5 that provides the novel biodegradation pathway of unsubstituted triazine (1,3,5-triazine) by *Acinetobacter* sp. Trz. Although there have been a lot of studies on substituted triazines biodegradation, no study on unsubstituted triazine (1,3,5-triazine) has been done previously. The chapter 5 describes the isolation of *Acinetobacter* sp. Trz that utilizes 1,3,5-triazine as its sole nitrogen source and the novel biodegradation pathway of 1,3,5-triazine. The ^{1}H-NMR and HPLC analysis shows that 1,3,5-triazine first transforms to formamidine, formamide and ammonia non-

enzymatically and then enzymes likely transform formamidine to formamide and formamide to formate and ammonia.

Table of Contents

Acknowledgement ……...i

Dedication……....ii

Abstract……...iii

List of Tables…….vii

List of Figures……...viii

Chapter 1: Introduction

1.1 Use of *s*-Triazine compounds in industry and agriculture…….1

1.2 *s*-Triazines in the environment and possible toxicity……2

1.3 *s*-Triazine degradation pathways……...4

1.4 Cyanuric acid hydrolases……...6

1.5 The application of CAH in degrading cyanuric acid in water treatment…….12

Chapter 2: Crystallization and preliminary X-ray diffraction studies of cyanuric acid hydrolase from *Azorhizobium caulinodans*

2.1 Summary……..……13

2.2 Introduction……..14

2.3 Experimental design…….15

2.4 Result and discussion……...18

Chapter 3: Cyanuric acid hydrolase from *Azorhizobium caulinodans* ORS 571: Crystal structure and insights into a new class of Ser-Lys dyad proteins

3.1 Summary..……23

3.2 Introduction……..24

3.3 Materials and Methods…….……26

3.4 Result…… ……...31

3.5 Discussion……46

Chapter 4: Redesign and X-ray structure of a commercial, thermophilic cyanuric acid hydrolase from *Moorella thermoacetica* : enzyme-substrate intermediate trapped in action

4.1 Summary..……50

4.2 Introduction……..50

4.3 Materials and Methods……52

4.4 Result…… ……...54

4.5 Discussion……63

Chapter 5: Biodegradation of 1,3,5-triazine by *Acinetobacter* sp. Trz

5.1 Summary..……66

5.2 Introduction……..66

5.3 Materials and Methods....……67

5.4 Result…….…… .....……...67

5.5 Discussion……76

References cited……78

List of Tables

Table 2.1 Data collection statistics for CAH…… ……21

Table 3.1 Data collection and refinement statistics…… ……...29

Table 3.2 Specific activity of wild type and mutant CAH on cyanuric acid as determined by the biuret hydrolase-coupled ammonia detection assay……39

Table 4.1 Data collection and Refinement Statistics…… .56

Table 4.2 Burgi-Dunitz angles and their corresponding bond distances… ……...62

Table 5.1 HPLC result for foramidine and formamide degradation analysis… ……74

Table 5.2 The O.D values at 600 nm from *Acinetobacter sp.* Trz culture with different N sources……76

List of Figures

Figure 1-1: Structure of *s*-triazine compounds……2

Figure 1-2: Comparison of the hydrolytic degradation pathways of melamine and atrazine……..3

Figure 1-3: Physical circular map of the catabolic plasmid pADP-1 from Pseudomonas sp. strain ADP……4

Figure 1-4: Metabolic pathways showing the reactions catalyzed by cyanuric acid hydrolase and barbiturase……7

Figure 1-5: Maximum-likelihood phylogenetic tree for the AtzD/barbiturase family ……8

Figure 1-6: Structural alignments between RMCAH and ACAH or AtzD…… …….9

Figure 1-7: Monomer of CAH…… ....…...10

Figure 1-8: Trimeric conglomeration of CAH active site……..11

Figure 1-9: Sequence alignment showing conservation of three Ser/Lys dyads found within subgroups of the CAH/barbiturase family……..12

Figure 2-1: Ring opening reaction of cyanuric acid by cyanuric acid hydrolase…… …..15

Figure 2-2: An optimized crystal of CAH-barbituric acid complex obtained from the crystallization condition containing 1.7 M magnesium sulphate and 0.1 M Tris-HCl pH 7.0……...17

Figure 2-3: Diffraction image from the CAH-barbituric acid complex crystal…… …….18

Figure 2-4: Self-rotation function showing the presence of three-fold rotational symmetry ($\kappa = 120°$ section) in the CAH-barbituric acid complex crystal…… ……...20

Figure 3-1: CAH catalytic reaction and inhibitor structure…… ……...26

Figure 3-2: Monomer of CAH……...31

Figure 3-3: Structural comparison between CAH and AtzD…….34

Figure 3-4: Comparison between CAH fold and the homotrimeric proteins in the YjgF-like SCOP structural superfamily…… ……..36

Figure 3-5: Trimeric conglomeration of CAH active site……..38

Figure 3-6: Sequence alignment showing conservation of three Ser/Lys dyads found within subgroups of the CAH/barbiturase family…….....40

Figure 3-7: Hydrogen bonding and electrostatic interactions network around three lysines from each domain in CAH…….42

Figure 3-8: Proposed mechanism for CAH..…… ……44

Figure 3-9: Front and back channels around the active site of CAH…….45

Figure 3-S1: The oligomerization of CAH…… ……32

Figure 3-S2 :Interactions in tetramer interface .....……..33

Figure 4-1: The scheme for the mutations of loops for enhancing crystallization of MCAH…….……...55

Figure 4-2: Oligomerization state of RMCAH……..57

Figure 4-3: Structural alignments between RMCAH and ACAH or AtzD……...58

Figure 4-4: Active site environment around the water molecule for deacylation step…..60

Figure 4-5: *Fo-Fc* omit electron density map contoured at 3.0 σ level…….61

Figure 4-6: The entire catalytic cycle of CYA hydrolysis…….64

Figure 5-1: Phylogenetic tree for Actinetobacter sp. Trz based on 16S rRNA gene sequencing……..……70

Figure 5-2: Degradation rate of 1,3,5-triazine in different pH…….……..71

Figure 5-3: The growth yields of *Acinetobacter* sp. Trz at different concentrations of nitrogen sources …… ……72

Figure 5-4: Formamidine degradation activity by the cells ..……...……73

Figure 5-5: Degradation pathway of 1,3,5-triazine via abiotic and biotic pathway……...77

Chapter1

Introduction

1.1 Use of *s*-Triazine compounds in industry and agriculture

Anthropogenic chemicals may persist in nature until organisms have time to evolve the specific enzymes to metabolize these newly introduced chemicals (1). Depending on the structure of the chemical, the evolution of new enzymes can take months or many years to occur.

As one class of the anthropogenic chemicals, *s*-Triazine ring compounds such as atrazine [2-chloro-4-(ethylamino)-6-(isopropylamino)-1,3,5-triazine], melamine and other triazines have been applied as herbicides, pesticides, resin intermediates, and dyes (Figure 1-1). Atrazine is used for the control of unwanted broad-leaf weeds for crops such as maize, sugarcane, sorghum, and pine over the past 50 years (2-4). In 2014 alone, 76 million pounds of atrazine were applied to crops and was credited for enhancing crop yields by 1% to 20% (5-7). Atrazine is synthesized from cyanuric chloride that is reacted with ethylamine and isopropyl amine to install the side chains characteristic of this herbicide (8). Atrazine is known to act on the plastoquinone-binding protein in photosystem II, causing the collapse of the electron transport system, causing oxidative damage and starvation (9). In the 1980's, the Environmental Protections Agency (EPA) reported that the half-life of atrazine in different soils ranged from 13 to 261 days although atrazine in some soils can persisted more than 4 years (10, 11). As a nonselective herbicide, simazine is also used to control annual grasses although now its use is banned by the European Union. Melamine is a major polymer intermediate, with on the order of one billion pounds used industrially. Cyanuric acid [1,3,5-triazine-2,4,6-triol] is an industrially versatile compound used as a precursor for the manufacture of herbicides, disinfectants, and non-protein nitrogen additives. There is evidence that cyanuric acid formed crystals with melamine that caused kidney failure and death when both were present in pet foods (12).

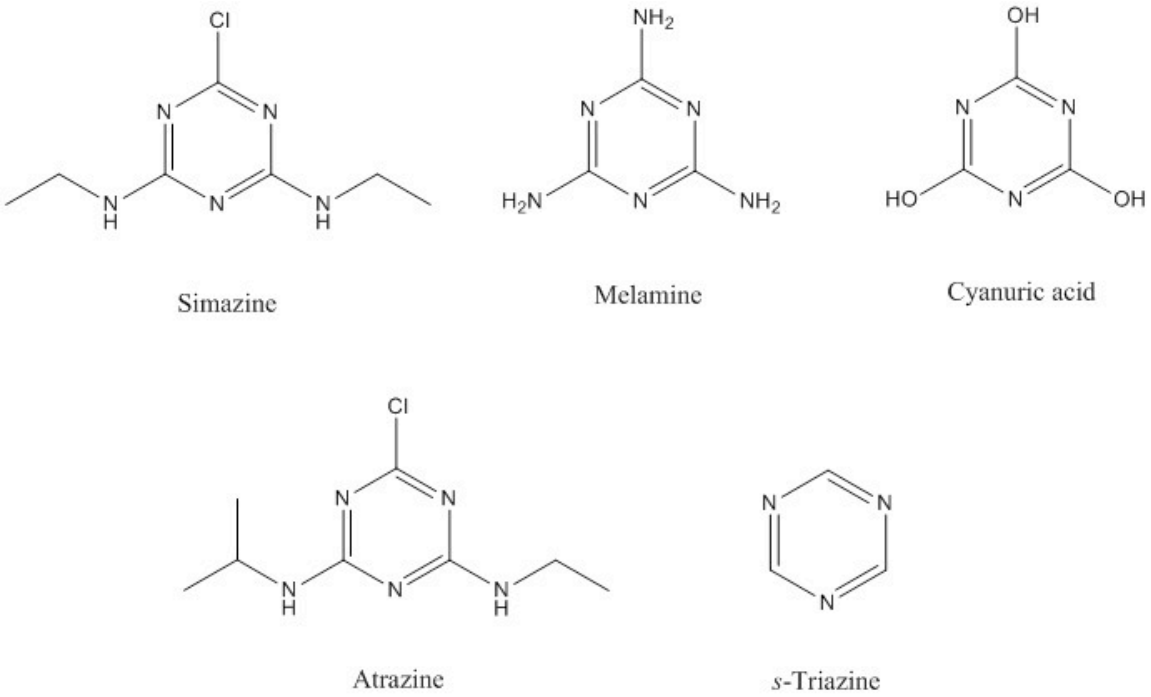


Figure 1-1. Structure of *s*-triazine compounds

1.2. *s*-Triazines in the environment and possible toxicity

Due to its long history of application, the atrazine level has sometimes exceeded the 3 ppb allowable level in groundwater as established by the Environment Protection Agency (13). Atrazine and its derivatives are detected in groundwater because it binds poorly to soil particles and so it can wash off farm fields in heavy rains. Therefore atrazine and its derivatives are mobile in soil with its mobility leading to the contamination of water resources (2, 14-17). Recent EPA reports showed that atrazine levels sometimes increased beyond the maximum level set by EPA in middle-western regions such as Missouri and Nebraska. There are numerous reports that atrazine or simazine residues in soil cause injury in corn or other plants by acting on young seedlings (18-20). Therefore, atrazine use was banned in the European Union in 2004 because of its potential environmental effects (5).

A number of bacteria that metabolize *s*-triazine compounds such as melamine, cyanuric acid, deethylsimazine as their nitrogen source were obtained out of soil by enrichment culture in the 1980s (21-24). In the 1990s, a number of bacteria which catabolized atrazine were discovered by different research groups (14, 25-28). Those bacteria could convert atrazine first to hydroxyatrazine that does not have any known negative impact on the environment or human health, and so is a very desirable initial metabolic reaction (Figure 1-2).

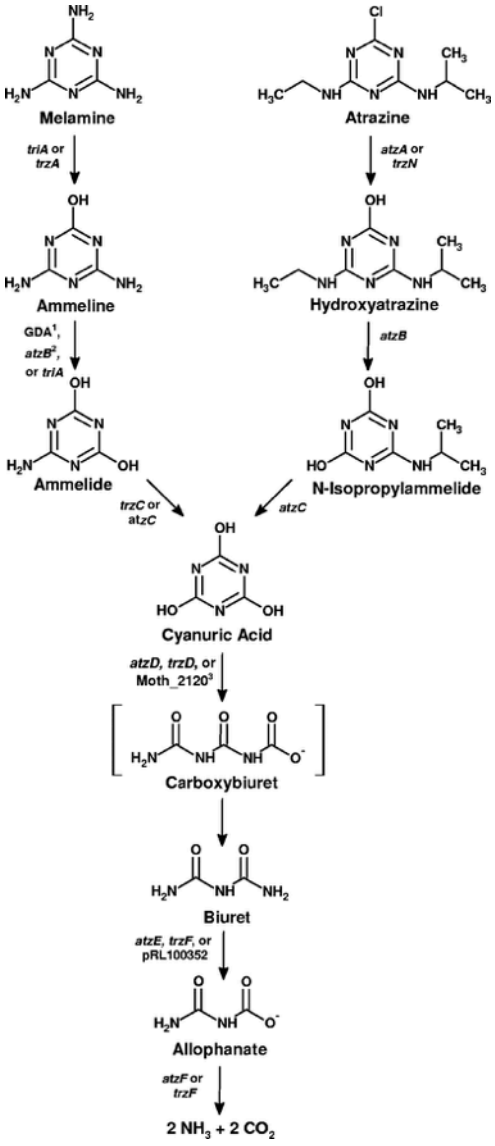


Figure 1-2. Comparison of the hydrolytic degradation pathways of melamine and atrazine.

(Anthony G. Dodge et al. Appl. Environ. Microbiol. 2012;78:1397-1403, Reproduced with Permission of ASM Press)

1.3. *s*-Triazine degradation pathways

Bacterial consortia (two or more bacteria) that carried out the complete mineralization of atrazine were identified in the 1980s (29). Those experiments showed that Rhodococci could completely degrade deisopropylatrazine that is the final metabolite in the atrazine catabolic pathway by a *Pseudomonas* species. After unsuccessful attempts at isolating single bacteria that metabolize atrazine in the 1980's, many different research groups subsequently isolated microorganisms that metabolize atrazine in pure culture (14, 25, 30-34). Interestingly, biodegradation abilities for a couple of recalcitrant chemicals are encoded by catabolic plasmids. Atrazine is also metabolized by enzymes encoded by genes on a catabolic plasmid (Figure 1-3) and the plasmid has been designated as pADP-1 (14, 35).

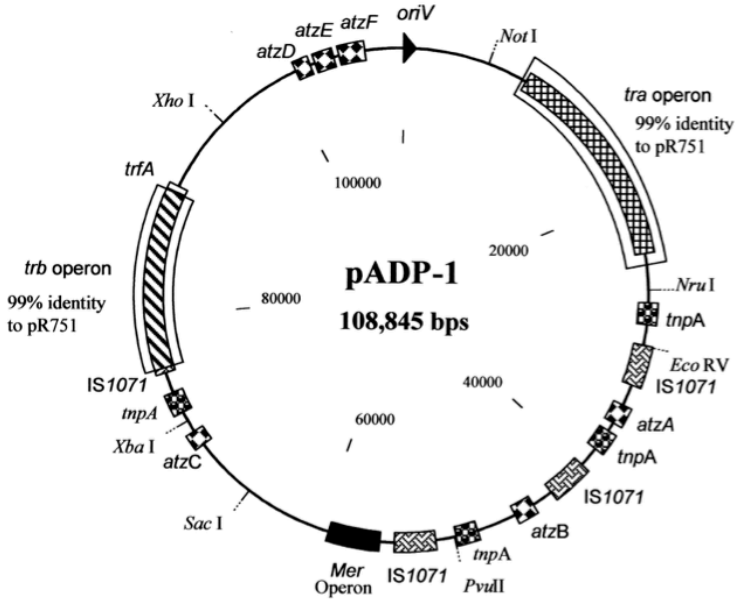


Figure 1-3. Physical circular map of the catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. The map positions of selected restriction sites, genes, and operons on pADP-1 are indicated
(Betsy Martinez et al. J. Bacteriol. 2001;183:5684-5697, Reproduced with Permission of ASM Press)

The complete genome sequence of plasmid pADP-1 revealed that the regions for maintaining, replicating and transferring the plasmid in the environment are

homologous to gene regions of IncPβ plasmid from *Enterobacter aerogenes*. Plasmid pADP-1 is broadly transmissible as shown by its transfer to *Escherichia coli* and other bacteria (36). Although the *atzDEF* genes are clustered on plasmid pADP-1, the *atzABC genes* are more distributed and each are flanked by the insertion sequence IS1071 elements. The insertion sequence has more than 95% sequence identity with those of *Alcaligenes* and other bacteria that also contain catabolic plasmids. Hence it is believed that transposases probably contribute to the building of the atrazine metabolic pathway encoded on the pADP-1 plasmid. The plasmid contains 104 identified genes including mercury resistance genes.

Our laboratory had characterized the genes and enzymes that are responsible for the first three reactions mediated by separate enzyme in the atrazine degradation pathway of *Pseudomonas* sp. ADP (31, 32, 37, 38). The first enzyme, atrazine chlorohydrolase or AtzA, hydrolytically displaces the chlorine atom bonded to the atrazine ring structure, yielding the product hydroxyatrazine. In Gram positive bacteria such as *Arthrobacter* and *Nocardioides* that degrade atrazine, the dechlorinating reaction is catalyzed by TrzN, an alternative chlorohydrolase (39-41). The dechlorination by AtzA or TrzN has received special attention because those enzymes convert atrazine to hydroxyatrazine, a safer compound towards the environment. Currently, the crystal structures of both AtzA and TrzN are available (41, 42). AtzB is the second enzyme that converts hydroxyatrazine to *N*-isopropylammelide through deamidation step. AtzC is the enzyme that transforms *N*-isopropylammelide to cyanuric acid. Cyanuric acid is also an intermediate formed during the biodegradation of atrazine, and other *s*-triazine compounds such as melamine (43) (Figure 1-2).

Cyanuric acid is further hydrolyzed by Cyanuric Acid Hydrolase (CAH) to carboxybiuret that spontaneously decarboxylate to biret. Biuret is converted to allophanate via AtzE and allophonate is transformed to $2NH_3$ and $2CO_2$ by AtzF. Recently the crystal structure of AtzF was solved and it is proposed to form a multienzyme complex with AtzD and AtzE (44).

As stated in this chapter, substituted *s*-triazine compounds such as melamine and atrazine undergo mineralization through dechlorination and deamidation to form

cyanuric acid, followed by further degradation to CO_2 and NH_3. Mineralization absolutely requires ring cleavage initiated by the enzyme CAH. However there have not been enough studies regarding the degradation mechanism of the unsubstituted *s*-triazine (1,3,5-triazine). Chapter 5 in the dissertation reports how 1,3,5-triazine is degraded into formate and NH_3 via abiotic and biotic reactions.

1.4. Cyanuric acid hydrolases

After the identification of the pADP plasmid that controls *s*-triazine degradation (35, 45) in *Pseudomonas* strain ADP, the genes designated *atzD* and *trzD* encoding cyanuric acid hydrolases were discovered in different microorganisms (46, 47). With help of genome sequencing information and sophisticated bioinformatics tools such as BLAST, Shotgun and HMMER, a number of the newly identified protein sequences were categorized into what has been designated the Cyanuric Acid Hydrolase/Barbiturase (CAH/Bar) protein family (48-50). The proteins of the CAH/Bar family consist of 340 to 400 amino acid having molecular weights ranging from 36 to 43 kDa. The sequences within the CAH/Bar family do not have any apparent functional or structural linkage to other known protein families (51).

CAH hydrolytically cleaves one of the C-N bonds in cyanuric acid to produce carboxybiuret. Carboxybiuret, a direct product from the ring opening reaction by CAH subsequently undergoes decarboxylation to form biuret (51, 52).

As a homologous enzyme with CAH, barbiturase is known to hydrolyze barbituric acid to produce ureidomalonic acid. Barbiturase is involved in one step in the oxidative pyrimidine catabolic pathway (Figure 1-4) in *Rhodococcus erythropolis* JCM3132 (53). It is particularly interesting that the activity of CAH is inhibited by barbituric acid in a competitive manner (46, 47) although the structures are very similar. The only difference between cyanuric acid and barbituric acid is that one nitrogen in cynauric acid is replaced by a carbon in barbituric acid. Despite that, cyanuric acid is not hydrolyzed by barbiturase but inhibits barbiturase in a competitive manner.

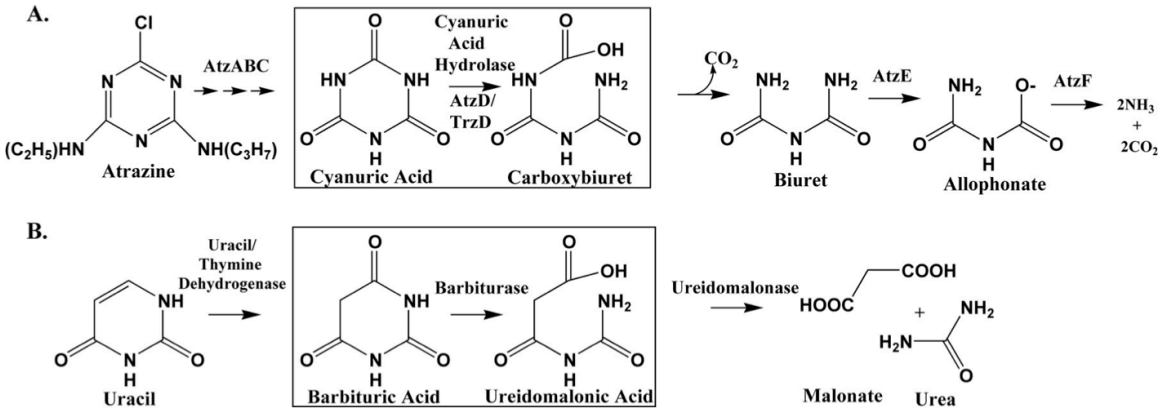


Figure 1-4. Metabolic pathways showing the reactions catalyzed by cyanuric acid hydrolase (A) and barbiturase (B). (Jennifer L. Seffernick et al. J. Bacteriol. 2012;194:4579-4588)

Based on the recent phylogenetic analysis with the potential CAH protein sequences, the predicted CAH sequences were placed into two clades (51). The sequences from two clades further were subdivided into five groups (51) (Figure 5). The phylogenetic tree indicates that CAH sequences in clade 1 have higher sequence identities each other compared to within sequences of clade 2. It is interesting that sequences in clade 1 have sequence identities ranging from 40 to 65% with functionally characterized CAHs while sequences in clade 2 have more than 72% of sequence identities with functionally characterized barbiturase.

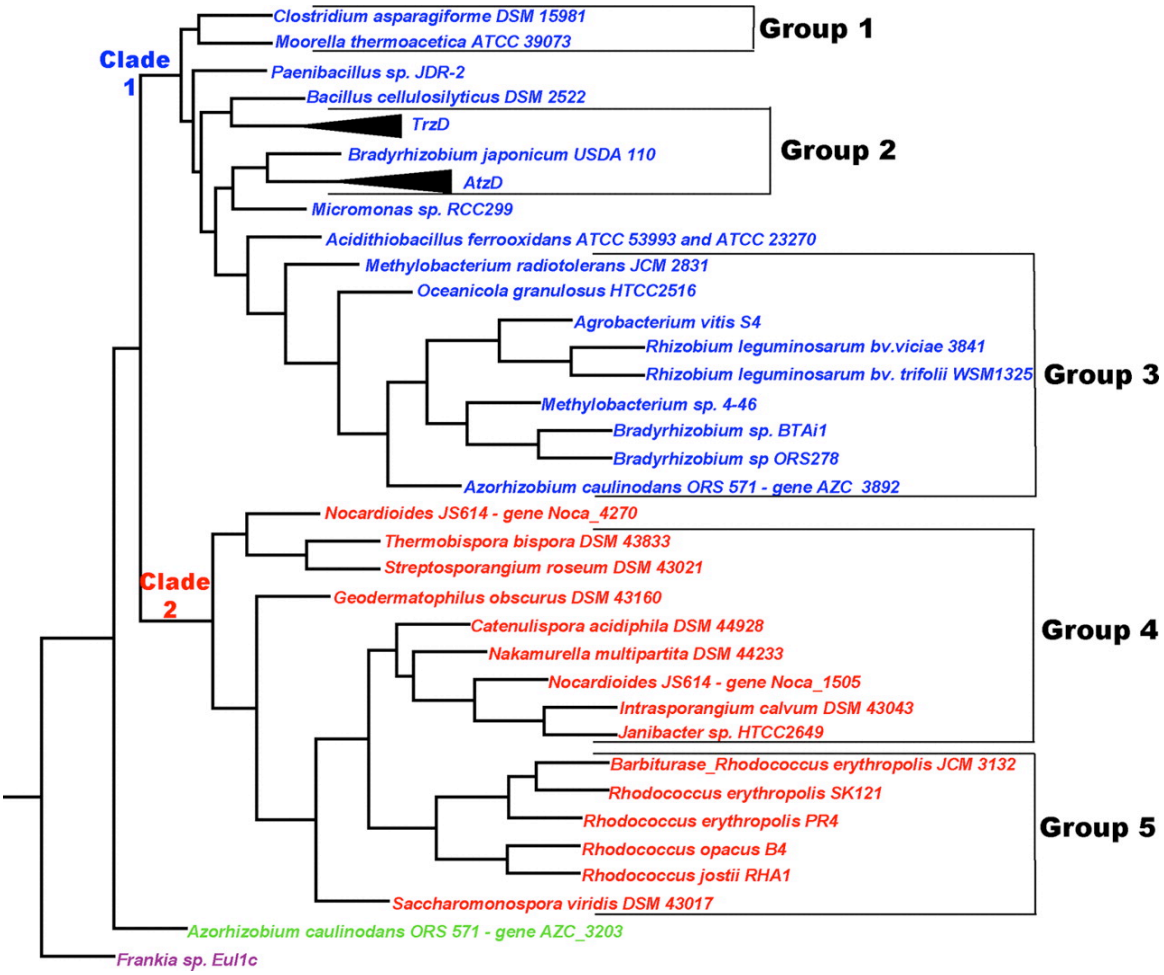


Figure 1-5. Maximum-likelihood phylogenetic tree for the AtzD/barbiturase family.
(Jennifer L. Seffernick et al. J. Bacteriol. 2012;194:4579-4588, Reproduced with Permission of ASM Press)

Overall different pairwise sequence alignments of members in the CAH/Bar family show 40-65 % sequence identity. However, it is difficult to functionally differentiate CAHs from barbiturases based on the sequence information alone. Considering their application in the real world and the relative stability of CAH enzymes compared to barbiturases, there has been more research on CAH enzymes. Hence, a couple of significant findings regarding CAH will be described in the following paragraphs.

Even though some effort has been invested in CAH research, the reaction mechanism of cyanuric acid transformation had not been clearly defined since CAH lacks obvious homology to other enzymes and there were no crystal structures.

Recently, X-ray crystal structures of CAH from *Pseudomonas* strain ADP (AtzD), *Azorhizobium caulinodans* ORS 571 (ACAH), and *Morella thermoacetica* (RMCAH) have been solved, the latter two being reported in this thesis (54, 55). The three CAHs for which structures have been solved are all categorized into clade 1. The three solved CAH have very similar 3-dimensional folds (Figure 1-6). The solution form of CAH is likely a tetramer based on gel filtration chromatography of ACAH and RMCAH (53).

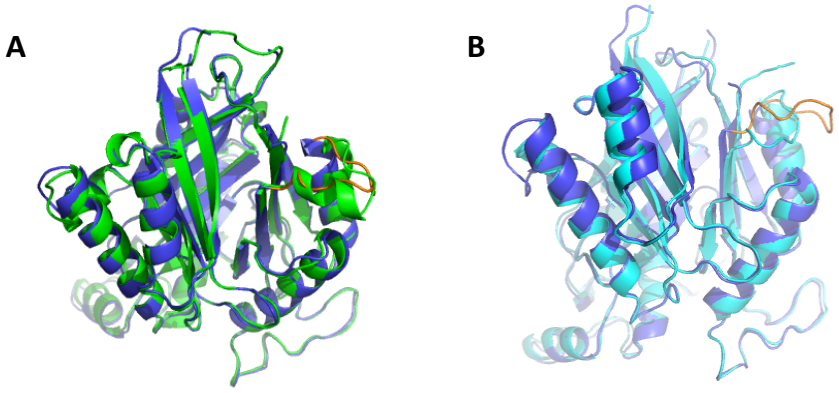


Figure 1-6. Structural alignments between RMCAH and ACAH or AtzD A. Structural alignment of RMCAH (Blue) and AtzD. (Green). B. Structural alignment of RMCAH (Blue) and ACAH (Cyan). Orange colored loops in RMCAH coordinate indicates the replaced loops from AtzD for engineering purpose.

The CAH monomer consists of 3 domains (domain A, B, and C) that have almost identical 3-dimensional folds to each other, although sequence identities within domains are less than 17% (Figure 1-7A, B, C) in case of ACAH. Each domain consists of 4 beta strands and 2 to 3 helices (Figure 1-7D).

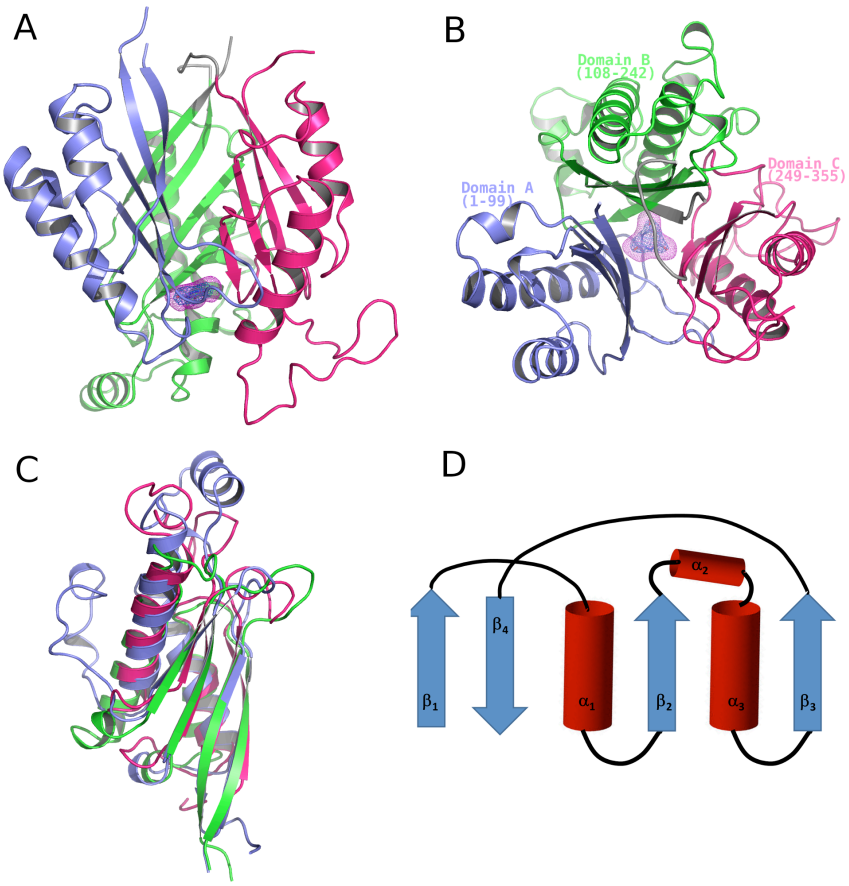


Figure 1-7. Monomer of CAH.
A. The side view of monomer structure. The selenium SAD-phased electron density for the barbituric acid inhibitor is shown in blue mesh, contoured at 2.0 σ. The simulated-annealing omit Fo-Fc map with the barbituric acid inhibitor omitted is shown in magenta mesh, contoured at 3.0 σ. Domains are depicted in different colors: Domain A, consisting of residues 1–99, in blue; Domain B, consisting of residues 108–242, in green; and Domain C, consisting of residues 249–355 in hot pink. B. The top view of monomer structure compared to A, looking through the center of the barrel along the pseudo three-fold axis. C. Overlay of the three domains of CAH. D. Depiction of the trimeric repeat domain fold in CAH. Red cylinders represent α-helices and blue arrows represent β-strands. α-helices 2 and 3 are present in the domain B, while domains A and C have a loop in place of α2 (Cho et al. PLoS One. 2014; Reproduced with Permission of PLoS One).

In the monomer, there are 3 pairs of serine-lysine dyads (S/K dyads) originating from each domain forming a trimeric conglomeration in CAH active sites (Figure 1-8).

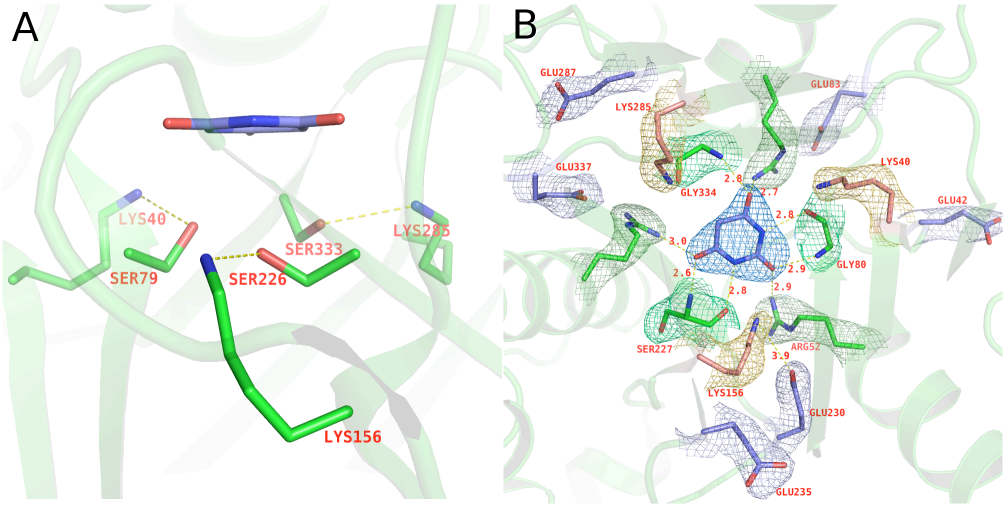


Figure 1-8. Trimeric conglomeration of CAH active site.
A. Three serine residues (Ser79, Ser226, Ser333) lie within hydrogen bond distances from corresponding lysine residues (Lys 40, Lys 156, and Lys 285, respectively). Each pair shows the typical arrangement for the Ser/Lys dyad hydrolase. B. Three arginine residues (Arg 52, Arg 188, Arg 314) and the backbone nitrogens from Gly80, Ser227 and Gly334 form hydrogen bonds with the carbonyl oxygens of the barbituric acid inhibitor. The simulated annealing composite-omit 2Fo-Fc map is shown in colored mesh, contoured at1.5 σ (Cho et al. PLoS One. 2014; Reproduced with Permission of PLoS One).

Figure 1-8 indicates the 3 Ser/Lys dyads are conserved across all of the groups of CAH/barbiturase family except for group H. Although no crystal structure of barbiturase has been solved, barburases in group 5 also show sequence conservation of the 3 Ser/Lys dyads as observed in CAHs. Together with its solution state as a likely tetramer like CAHs, it seems that barbiturase also uses Ser/Lys dyads for barbiturase hydrolysis in a similar active site.

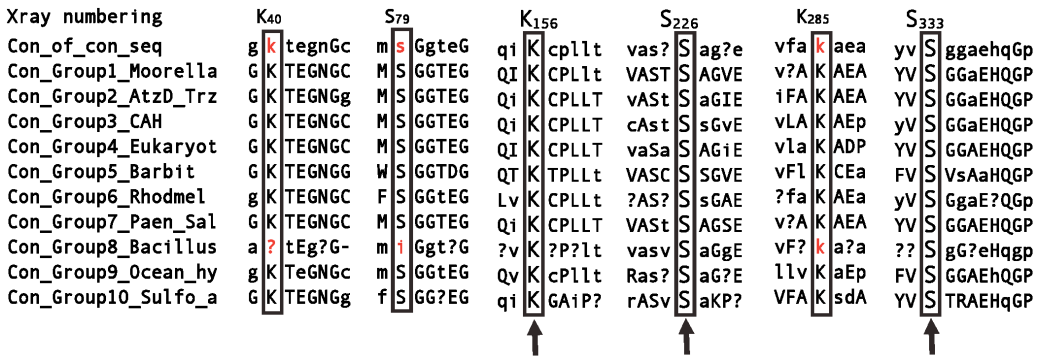


Figure 1-9. Sequence alignment showing conservation of three Ser/Lys dyads found within subgroups of the CAH/barbiturase family.
Subgroups of the updated family of 119 sequences were determined via phylogenetic clades. Consensus sequences were determined for each of the subgroups. Only sequences directly surrounding the dyads are displayed. Absolute, 100% conservation is represented by capital letters and greater than 50% conservation by lowercase letters. Columns in which no single residue was conserved at greater than 50% are depicted by a question mark. This usually occurred in regions of highly divergent sequence or where subgroups within a clade contained different residues. Residues equivalent to the Ser/Lys dyad residues are outlined with black boxes. Arrows point to Ser/Lys dyad residues with complete conservation across the family; residues not conserved in these positions are highlighted in red (Cho et al. PLoS One. 2014; Reproduced with Permission of PLoS One).

1.5. The application of CAH in degrading cyanuric acid in water treatment

As a massively produced chemical (56), cyanuric acid is used as a disinfectant in swimming pool treatments. Di- and tri-chloroisocyanuric acids maintain stable and high concentration of chlorine by slowly supplying chlorine (hypochlorite) to swimming pool water (57). Although the continuous supply of chlorine by *s*-triazine compounds is effective in disinfecting water, the accumulation of cyanuric acid (24), the product of this dechlorination impairs disinfection efficiency by shifting the chemical equilibrium. Therefore, swimming pool maintenance currently requires emptying and recharging pool water to remove cyanuric acid. Water conservation efforts have been a motivating factor to finding alternate ways, such as the use of CAH for remediating cyanuric acid due to the chemical stability of cyanuric acid. Recently, a paper by Yeom, et al (58) reported that heat-treated *E.coli* expressing CAH from the *Morella* thermophile (MCAH) efficiently removed cyanuric acid from swimming pool water.

Chapter 2

Crystallization and preliminary X-ray diffraction studies of cyanuric acid hydrolase from *Azorhizobium caulinodans*

2.1 Summary

Cyanuric acid is synthesized industrially and forms during the microbial metabolism of s-triazine herbicides. Cyanuric acid is metabolized by some microorganisms via cyanuric acid hydrolase (CAH) that opens the *s*-triazine ring as a prelude to further metabolism. CAH is a member of the rare cyanuric acid hydrolase (AtzD)/barbiturase family for which no three-dimensional structures are currently available. Here, we report crystallization and preliminary X-ray diffraction analysis of CAH from *Azorhizobium caulinodans*. CAH was co-crystallized with barbituric acid, a close analog of cyanuric acid that is a tight-binding, competitive inhibitor. Crystals suitable for X-ray diffraction experiments were grown in conditions containing PEG8K or magnesium sulfate as precipitants. An X-ray diffraction dataset from the CAH-barbituric acid crystals has been collected to 2.7Å resolution. The crystals were found to be in space group $I4_122$, with the unit cell parameters of a=b=237.9 Å, c=105.3 Å, $\alpha = \beta = \gamma = 90°$

2.2 Introduction

Cyanuric acid is found in nature but it occurs principally due to industrial synthetic processes. It is produced on the order of 350 million pounds annually, largely as a precursor in the synthesis of disinfectants, dyes, and pesticides (56). Cyanuric acid is problematic in swimming pools where it accumulates due to the decomposition of chlorinated isocyanuric acids that are added for water disinfection. Additionally, in recent pet food and milk poisonings, the presence of cyanuric acid in combination with melamine was shown to cause insoluble crystals to form in the kidneys, leading to renal failure (59). Cyanuric acid also forms inside cells during the microbial metabolism of melamine (60) and *s*-trazine herbicides (35). Certain bacteria further metabolize cyanuric acid by hydrolytically opening the *s*-trazine ring and subsequent amide-bond hydrolysis reactions (61). The ring opening reaction, the first step of the cyanuric acid metabolism, is catalyzed in microbes by a homologous set of enzymes known as cyanuric acid hydrolases (CAH) (47) (46). CAH catalyzes the addition of water to a C-N bond in cyanuric acid to produce carboxybiuret (Figure 2-1). Carboxybiuret subsequently undergoes spontaneous decarboxylation to yield biuret, the substrate for the next enzyme in the metabolic pathway (51) (62) CAH enzymes are related to the enzyme barbiturase that functions in purine catabolism (53), catalyzing a similar hydrolytic ring opening reaction. However, barbituric acid is not a substrate for CAH but was rather observed to be a tight-binding, competitive inhibitor (47).

The cyanuric acid hydrolase/barbiturase family was recently analyzed by bioinformatics, gene cloning, and protein characterization studies, which allowed identification of new genes that belong to this protein family (51). Still, despite the analysis of more than thousands of genomes and the GenBank database, only forty-one members of this family have been identified and they could not be linked by sequence to any known protein superfamily. In addition, to this date, there are no proteins in the cyanuric acid hydrolase/barbiturase protein family for which an X-ray structure has been elucidated. The absence of any homologs outside of the forty-one proteins identified by extensive bioinformatics analysis held open the possibility that

this protein family might represent a unique fold. In this context, we conducted X-ray crystallographic studies to obtain the three-dimensional structure of a cyanuric acid hydrolase and describe here the crystallization and preliminary X-ray diffraction studies of the CAH, AZC_3892, from *Azorhizobium caulinodans*.

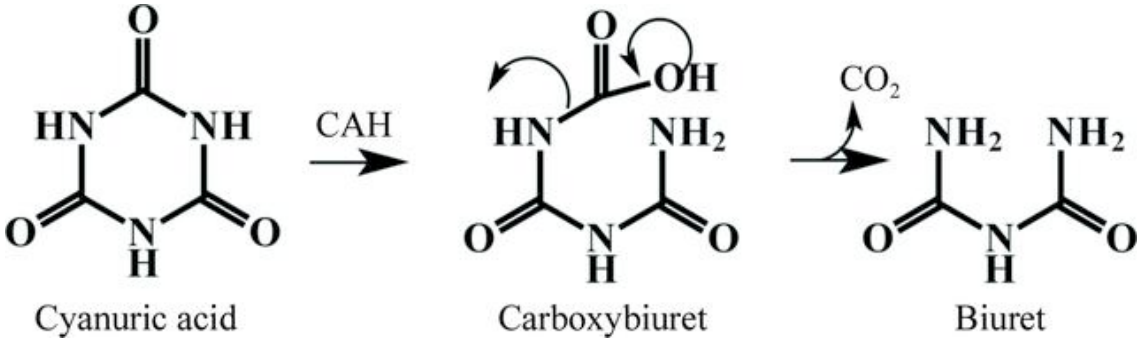


Figure 2-1. Ring opening reaction of cyanuric acid by cyanuric acid hydrolase.

2.3 Experimental design

2.2.1 Protein expression and purification

An NdeI-HindIII fragment encoding CAH from *Azorhizobium caulinodans* was subcloned into NdeI-HindIII sites of the pET28b+ vector (Novagen, Madison, WI). The resulting expression vector was introduced into the *Escherichia coli* strain BL21(DE3). The transformed strain was grown at 310K in LB medium (10 g Bacto tryptone, 5 g yeast extract, 10 g NaCl per litre) supplemented with 50 µg/ml kanamycin. Cells were grown until the optical density of the culture (OD600) reached 0.6. The expression of CAH was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM and the cell culture was continued for 30 hours at 298 K.

The bacterial cells were harvested by centrifugation (6400 x g, 10 min, 277 K) and resuspended in binding buffer (50 mM Tris-HCl pH7.0, 200 mM NaCl, 5 mM imidazole, 1mM β-mercaptoethanol) with protease-inhibitors (Roche). A microfludizer was used for cell disruption and the cell lysate was centrifuged at 63,988 x g for 40 min at 277 K. The supernatant was loaded onto a Ni-NTA resin (Qiagen)

that had been equilibrated with 3 column volumes of the binding buffer. Unbound proteins were washed away with 10x column volumes of washing buffer (50 mM Tris-HCl pH 7.0, 200 mM NaCl, 10 mM imidazole, 1mM β-mercaptoethanol). The target protein was eluted with two column volumes of elution buffer (50 mM Tris-HCl pH 7.0, 200 mM NaCl, 200 mM imidazole, 1mM β-mercaptoethanol). The eluted protein was concentrated by ultrafiltration and applied onto a Superdex-200 column (GE Healthcare) equilibrated with a running buffer containing 20 mM Tri-HCl pH 7.0, 200 mM NaCl, 5 mM dithiothreitol (DTT). The purity and size of the protein was confirmed using SDS-PAGE. To form a CAH-inhibitor complex, barbituric acid was added to the purified CAH at a molar ratio of 5:1 (inhibitor: protein). Either the unliganded CAH or the CAH-inhibitor complex was concentrated to 20-25 mg ml^{-1} for crystallization. Selenomethionine-substituted CAH was expressed using the methionine biosynthesis inhibition method as previously described (63) and purified as described above for the native CAH.

2.2.2 Crystallization

Crystallization screening of CAH or the CAH–inhibitor complex was performed by the sitting-drop vapor-diffusion method using a Rigaku CrystalMation system in the Nanolitre Crystallization Facility at the University of Minnesota. Conditions from the following screening kits were tested at 293 K: Crystal Screen HT, Index Screen, PEGRx HT, PEG/Ion and SaltRx (Hampton Research) and JCSG+ and PACT (Qiagen). Each drop was prepared by mixing 100 nl protein or protein–inhibitor complex solution with 100 nl reservoir solution. Within 3 d of incubation, small crystals were observed for the CAH–inhibitor complex, but not for unliganded CAH, in drops that contained magnesium sulfate and Tris–HCl pH 7.0 or PEG 8K and HEPES–NaOH pH 7.5. Optimizations for improving the sizes of these crystals were conducted by varying the precipitant concentration, buffer pH and protein concentration with the sitting-drop vapordiffusion method. After optimization, crystals suitable for X-ray diffraction experiments were obtained with the following two conditions: (i) 1.0–1.7 M magnesium sulfate, 0.1 M Tris–HCl pH 7.0– 7.5 and (ii) 5%

PEG 8K, 0.1 M HEPES–NaOH pH 7.5. Crystals obtained in the optimized conditions (Figure 2) grew to maximum size within 3 d in drops prepared by mixing 100 nl protein solution and 100 nl reservoir solution. The initial diffraction experiments were conducted at 100 K using a Rigaku R-AXIS IV X-ray diffractometer.

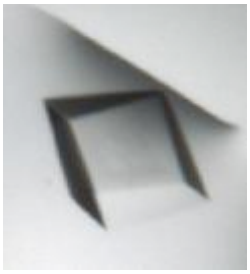


Figure 2-2. An optimized crystal of CAH-barbituric acid complex obtained from the crystallization condition containing 1.7 M magnesium sulfate and 0.1 M Tris-HCl pH 7.0.

2.2.3 Data collection and processing

The X-ray diffraction data were collected on beamline 4.2.2 of the Advanced Light Source, Berkeley, California, USA. The oscillation angle for each diffraction image (Figure 2-3) was 0.5°. The crystals were briefly soaked in the cryoprotectant (reservoir solution containing 25% glycerol) and flash-cooled in liquid nitrogen before X-ray irradiation. The diffraction data were indexed, integrated and scaled with HKL-2000 (64). We collected a complete data set from a selenomethionine-substituted CAH–barbituric acid complex crystal at the K absorption edge of selenium for singlewavelength anomalous dispersion (SAD) phasing (Table 2-1).

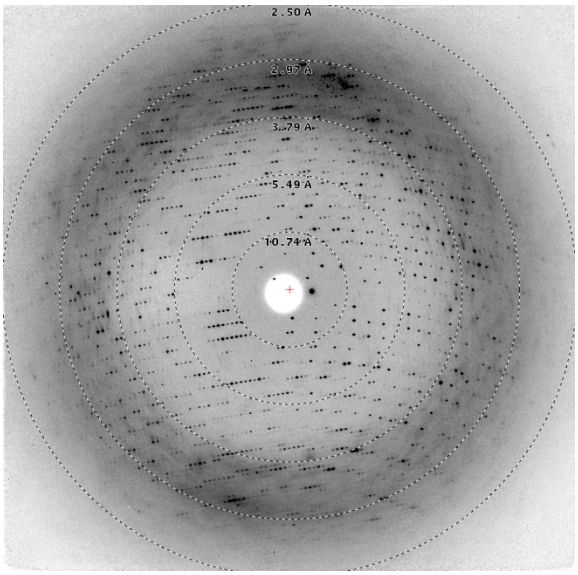


Figure 2-3. Diffraction image from the CAH-barbituric acid complex crystal.

2.4 Result and discussion

CAH was successfully expressed in *E. coli* strain BL21 (DE3) and purified to homogeneity. After affinity chromatography using a six-histidine tag attached to the N-terminus of the protein, CAH was further purified by size-exclusion chromatography. The specific activity of CAH purified using this protocol with cyanuric acid as the substrate was 7.9 mmol min^{-1} mg^{-1} as determined using the Berthelot reaction for ammonia and coupling with the enzyme biuret hydrolase (65, 66). We found that the molecular mass of CAH as estimated by size-exclusion chromatography was 160 kDa, suggesting that CAH exists as a tetramer. The size of the CAH monomer as determined by SDS– PAGE (15%) was consistent with the expected molecular weight of 36 040 Da for the native protein plus the N-terminal His-tag sequence (MGSSHHHHHHSSGLVPRGSH).

In the initial crystallization screening, the unliganded form of CAH did not yield any crystals. Hence, a competitive inhibitor, barbituric acid, was added to the purified protein to stabilize the enzyme and thereby facilitate crystallization. The bound competitive inhibitor is also helpful in discerning the active site within the CAH structure. Crystals of the inhibitor-bound CAH were observed within 3 d in conditions

containing magnesium sulfate, PEG 8K or PEG 10K as precipitants. On a rotating-anode home X-ray source, the CAH–barbituric acid complex crystals grown under PEG 8K and magnesium sulfate conditions showed good diffraction images (extending to ~3.0 Å resolution), while crystals grown with PEG 10K did not diffract well (diffraction extending to ~8.0 Å resolution). Because no homologous protein structure was available for CAH, we pursued structure determination by selenomethionine SAD phasing. Selenomethionine-derivative crystals could be produced in the magnesium sulfate condition, but not in the PEG-containing conditions. Therefore, a Se-SAD data set extending to 2.7 Å resolution was collected on a crystal obtained from the crystallization condition containing 1.7 M magnesium sulfate and 0.1 M Tris–HCl pH 7.0. The CAH–inhibitor complex crystals belonged to space group $I4_122$, with unit-cell parameters a = b = 237.9, c = 105.3 Å, $\alpha = \beta = \gamma = 90°$. The self-rotation function showed strong peaks on the $\kappa = 120°$ section, indicating that there might be three CAH–inhibitor complexes in the crystallographic asymmetric unit related by nocrystallographic threefold symmetry or that pseudo-threefold symmetry may exist within the CAH monomer structure (Figure 2-4). Assuming that two or three CAH monomers are in the asymmetric unit, the solvent content would be 73.5 or 60.2%, respectively.

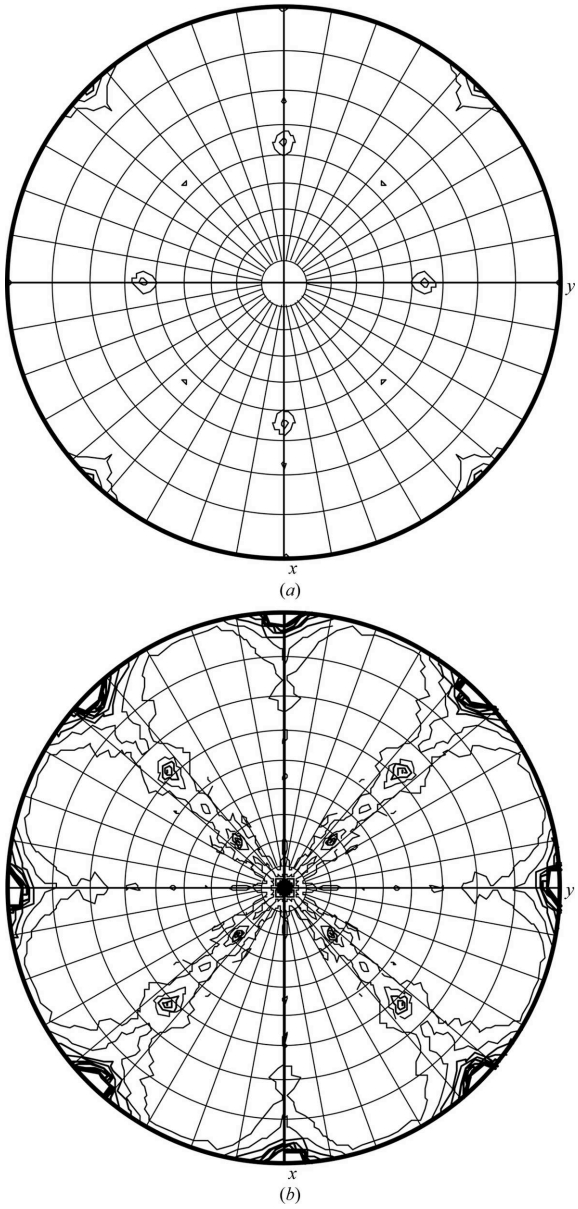


Figure 2-4. Self-rotation function showing the presence of three-fold rotational symmetry (κ = 120° section) in the CAH-barbituric acid complex crystal, calculated at 3.0 Å resolution using the program *MOLREP* (67) (a) κ = 120, (b) κ = 180 section. The integration radius was 29 A˚. The x and y axes are parallel to the a and b axes of the crystal, respectively.

The structure of the CAH–barbituric acid complex has been successfully determined by the selenomethionine SAD phasing method using PHENIX (Adams et al., 2010), giving figures of merit of 0.347 and 0.690 before and after density modification, respectively, based on 29 Se sites found out of a possible 30. Pseudo-threefold internal symmetry is evident within the CAH monomer, consistent with the self-rotation function results. Notably, while this manuscript was under review, the crystal structure was reported of AtzD, another cyanuric acid hydrolase/ barbiturase family member from *Pseudomonas* sp. strain ADP which shares 51% sequence identity with the *Azorhizobium* CAH crystallized in this study (55). Structural analyses of the CAH– barbituric acid complex, including detailed structural comparison between these two enzymes, will be published elsewhere.

	CAH
Data collection	
Space group	$I4_122$
Cell dimensions	
a, b, c (Å)	237.9, 237.9, 105.3
α, β, γ (°)	90.0, 90.0, 90.0
Resolution range (Å)	50.00 – 2.70 (2.75 – 2.70)
R_{merge}	0.088 (0.566)
Completeness (%)	98.4 (94.7)
Multiplicity	9.9 (8.2)
I/σ (I)	20.3 (2.20)
Total No. of reflections	403575 (15752)
No. of unique reflections	40849 (1921)

Statistics for the highest resolution shell are shown in parentheses.

Table 2-1. Data collection statistics for CAH

Acknowledgments

We thank Jay Nix at the beamline 4.2.2 of the Advanced Light Source for his support in data collection. We also thank Dr Jennifer Seffernick her valuable discussion. This work was partly supported by a grant from Lonza and from the Synthetic Ecology Initiative of the BioTechnology Institute at the University of Minnesota

Chapter 3

Cyanuric Acid Hydrolase from *Azorhizobium caulinodans* ORS 571: Crystal Structure and Insights into a New Class of Ser-Lys Dyad Proteins

3.1 Summary

Cyanuric acid hydrolase (CAH) catalyzes the hydrolytic ring-opening of cyanuric acid (2,4,6-trihydroxy-1,3,5-triazine), an intermediate in *s*-triazine bacterial degradation and a by-product from disinfection with trichloroisocyanuric acid. In the present study, an X-ray crystal structure of the CAH-barbituric acid inhibitor complex from *Azorhizobium caulinodans* ORS 571 has been determined at 2.7 Å resolution. The CAH protein fold consists of three structurally homologous domains forming a β-barrel-like structure with external α-helices that result in a three-fold symmetry, a dominant feature of the structure and active site that mirrors the three-fold symmetrical shape of the substrate cyanuric acid. The active site structure of CAH is similar to that of the recently determined AtzD with three pairs of active site Ser-Lys dyads. In order to determine the role of each Ser-Lys dyad in catalysis, a mutational study using a highly sensitive, enzyme-coupled assay was conducted. The 10^9-fold loss of activity by the S226A mutant was at least ten times lower than that of the S79A and S333A mutants. In addition, bioinformatics analysis revealed the Ser226/Lys156 dyad as the only absolutely conserved dyad in the CAH/barbiturase family. These data suggest that Lys156 activates the Ser226 nucleophile which can then attack the substrate carbonyl. Our combination of structural, mutational, and bioinformatics

analyses differentiates this study and provides experimental data for mechanistic insights into this unique protein family.

3.2 Introduction

Cyanuric acid, or 2,4,6-trihydroxy-1,3,5-triazine, is an industrially important compound used to make pesticides, dyes, and disinfectants. The latter consist largely of N-chlorinated cyanuric acid derivatives, used for cleaning and swimming pool disinfection. In 2005, 350 million pounds of cyanuric acid were produced industrially for those purposes (68). Although cyanuric acid alone is not very toxic, coingestion with melamine was the cause of recent poisonings due to adulterated pet food (59), leading to the largest pet food recall in North America and a $24 million settlement.

The use of cyanuric acid as a nitrogen source by bacteria generally prevents accumulation in the environment following disinfectant and pesticide degradation (60) (35). The first step of cyanuric acid metabolism is catalyzed by cyanuric acid hydrolase (CAH) and results in *s*-triazine ring opening (46, 47) to produce unstable carboxybiuret, which undergoes rapid, spontaneous decarboxylation to yield biuret (Fig. 1A) (51). Biuret is then metabolized by two more enzymes, biuret hydrolase and allophonate hydrolase, resulting in the complete mineralization of cyanuric acid and the release of three nitrogen atoms from the *s*-triazine ring as ammonia.

Recently, bioinformatics analyses determined that cyanuric acid hydrolases were members of an isolated family of enzymes and homologous to barbiturases (51). This cyanuric acid hydrolases/barbiturase family consisted of only 41 members when initially characterized. Since then, additional cyanuric acid hydrolases have been identified (69), but it still constitutes a unique protein family with no known linkages to other families. Only limited studies have been conducted with barbiturase, a rare enzyme involved in the catabolism of pyrimidines in a limited number of bacteria (53). Barbiturase is not reactive with cyanuric acid. Conversely, barbituric acid (Fig. 1B) is not reactive with cyanuric acid hydrolases, but it is bound tightly and competitively inhibits the enzyme's activity with cyanuric acid (46, 47). Since most of the characterized proteins in the cyanuric acid hydrolase/barbiturase protein family react

with cyanuric acid and cyanuric acid is not known to be a natural product of any bacterium, plant, or animal, it has been proposed that this might represent a relatively minor protein family, perhaps even a fold that was dying out and has undergone a resurgence with the recent input of commercial *s*-triazine ring compounds into the environment.

In light of the above, there is increasing interest to understand the structure and reaction mechanism of cyanuric acid hydrolases. We have conducted X-ray crystallographic studies of the cyanuric acid hydrolase from *Azorhizobium caulinodans* ORS 571 (NCBI reference sequence; YP_001526808.1), denoted here as CAH. Our preliminary account of the crystallization was previously reported (70). During the writing of this manuscript, the structure of AtzD, the cyanuric acid hydrolase in *Pseudomonas* sp. strain ADP, was reported (55). AtzD and CAH are 51% identical in a pairwise amino acid sequence alignment and share an overall unique fold consisting of three structurally analogous domains that form a barrel structure. Unique to the present study are details regarding structural features of subunit contacts and second-tier active site residues, proposed to contribute to catalysis. In addition, the present work developed a highly sensitive, coupled-enzyme assay using biuret hydrolase that allowed the detection of mutant activity diminished by eight orders of magnitude from that of wild-type CAH. This assay was used to differentiate the catalytic roles of the three active site serine residues in CAH which could not be done in the AtzD study. Mutant, bioinformatics, and structural studies within this paper have allowed us to propose a different catalytic serine nucleophile than was previously proposed for AtzD (55).

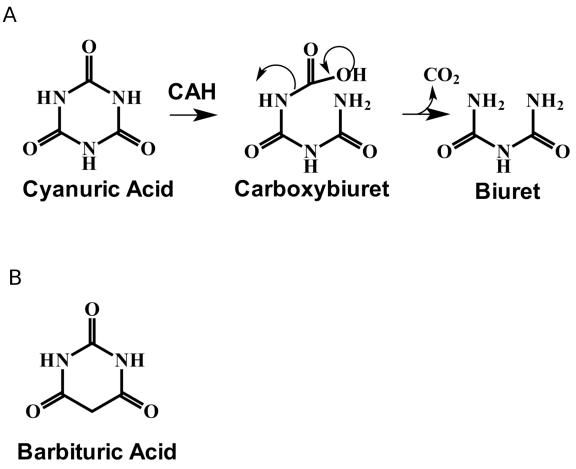


Figure 3-1. CAH catalytic reaction and inhibitor structure.
A. CAH conversion of cyanuric acid to carboxybiuret, which spontaneously decarboxylates to produce biuret. B. Structure of the CAH substrate analog and inhibitor, barbituric acid.

3.3 Materials and Methods

3.2.1 Cloning, Site Directed Mutagenesis, Protein Expression and Purification

Cloning the functional CAH gene from *A. caulinodans* ORS 571 (51), and heterologous expression and purification of the encoded native and selenomethionine-labeled proteins (70) were described previously. The protein can be identified by the following reference tags: A8IKD2 (Uniprot), AZC_3892 (locus name), YP_001526808 (accession), and GI158425516. Site directed mutagenesis was conducted with a QuikChange kit (Agilent Technologies, Santa Clara, CA), using the following primers: S79A used 5'-gcctcgtcatggccggcggcacc-3' and 5'-ggtgccgccggccatgacgaggc-3', S226A used 5'-gcgcgcgcgagctgtgccagcggt-3' and 5'-accgctggcacagctcgcgcgcgc-3', S333A used 5'-acggagatctatgtcgccggcggcggc-3' and 5'-gccgccgccggcgacatagatctccgt-3', K40A used 5'-cctcgccatctttggagcgaccgagggcaatggc-3' and 5'-gccattgccctcggtcgctccaaagatggcgagg-3', K156A used 5'-gcatttcgtgcaggtggcatgcccgcttctcacc-3' and 5'-ggtgagaagcgggcatgccacctgcacgaaatgc-3', K285A used 5'-catcgtgctcgccgcggcggagcccagc-3' and 5'-gctgggctccgccgcggcgagcacgatg-3', R188K used 5'-ctcaaatccatgggcctctcaaaggggggcgagcgc-3' and 5'-gcgctcgccccctttgagaggcccatggatttgag-

3', and R188Q used 5'-tgggcctctcacaggggcgagcgcg-3' and 5'-cgcgctcgccccctgtgagaggccca-3'. The protein yield of both wild type and mutant enzymes was 7-10 mg/L. CD spectroscopy experiments were conducted over the range of 200-250 nm on a JASCO J-815 CD spectrophotometer equipped with a Peltier temperature control (JASCO Inc.).

The wild type and mutant enzymes were analyzed by circular dichroism (CD) and indicated that the proteins had the correct secondary structure and were stable. The proteins were dissolved at a concentration of 14 μM in 0.1 M potassium phosphate buffer, pH 7.0. For far-UV CD data collection, the measurements were carried out in 1 mm path length cell at 25 °C. All measurements were an average of three scans. For thermal melting, the changes in CD spectra of the proteins were monitored at 220 nm during the course of heating the samples from 20 °C to 90 °C. No precipitation or signs of enzyme instability was observed.

3.2.2 Crystallization

Initial crystallization experiments were previously described (70). The crystals in this paper were obtained using the hanging drop vapor diffusion method at 20°C. Optimized crystallization conditions suitable for X-ray diffraction experiments contained 1.0-1.7 M magnesium sulfate, 0.1 M Tris-HCl pH 7.0-7.5. Suitable crystals for the X-ray diffraction experiments were observed within three days.

3.2.3 Structure Determination and Refinement

Crystal structure of the CAH-barbituric acid complex was determined by the single-wavelength anomalous dispersion (SAD) phasing method using a dataset collected at the Se peak wavelength (0.97864Å) on a selenomethionine-substituted CAH crystal. Details of the X-ray diffraction data collection were reported earlier (70). Twenty nine selenium atoms were located and phases calculated by PHENIX [12]. Autobuild by RESOLVE (71) built about 60% of the protein residues for two monomers in the asymmetric unit of the crystal. Iterative cycles of model refinement by PHENIX and model building using COOT (72) yielded a final model with R-work of 15.94% and

R-free of 19.13% at 2.7 Å resolution. The final model contains two protein chains with 351 and 347 amino acid residues, 3 barbituric acid molecules, 3 Mg^{2+} ions, 10 sulfate ions, and 172 water molecules.

Barbituric acid is not three-fold symmetrical and has a carbon instead of one of the nitrogens in the cyanuric acid substrate (Figure 3-1). The electron density readily identifies the position of the barbituric acid ring and the carbonyls within the ring, but the CAH structure resolution was not sufficient enough to immediately determine the position of a carbon versus a nitrogen in the ring. Therefore, the barbituric acid molecule at the active site was modelled in three different orientations, each with 120^{o} rotation around the axis perpendicular to the plane of the ring. Simulated annealing refinement was carried out for each barbituric acid orientation, and difference electron density maps were generated. The three different orientations yielded similar 2Fo-Fc maps but distinctive Fo-Fc maps. The present model yields a clean 2Fo-Fc map without any residual Fo-Fc map peaks, while the other two orientations of barbituric acid yield Fo-Fc maps with positive and negative peaks.

Quality of the model was assessed by PROCHECK (73), and no amino acid residue was found to be in the disallowed region of the Ramachandran plot. The summary of data collection, model refinement, and model quality is shown in Table 1. Atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank under the accession code 4NQ3.

	CAH
Space group	$I4_122$
Cell dimensions	
a, b, c (Å)	237.9, 237.9, 105.3
Wavelength (Å)	0.97864
Detector	NOIR-1
Resolution range (Å)	50.0–2.70 (2.75–2.70)
R_{sym}[a]	0.088 (0.566)
R_{pim} [30]	0.035 (0.242)
R_{meas} [31]	0.113 (0.731)
CC1/2 [32]	0.976 (0.870)
CC* [32]	0.994 (0.964)
Completeness (%)	98.4 (94.7)
Redundancy	9.9 (8.2)
I/σ (I)	20.3 (2.20)
Total No. of reflections	403575 (15752)
No. of unique reflections	40849 (1921)
FOM[b] after Density Mod.	0.69
R_{work}	0.1594 (0.2665)
R_{free}	0.1913 (0.3228)
Average B-factor	53.40
Macromolecules	53.30
Barbituric acid	
active site (A chain)	26.79
active site (B chain)	42.74
surface-bound	73.62
SO_4^{2-} and MG^{2+}	99.46
Water	43.40
R.m.s.d from ideality	
Bond lengths[c] (Å) (r.m.s.d[b])	0.009
Bond angles[c] (°)	1.20
Ramachandran plot (%)	
Most favored	92
Additional allowed	8
Generously allowed	0
Disallowed	0

a = $\Sigma|I_{obs}—I_{avg}|/\Sigma I_{obs}$, ($I_{obs}$ stands for the observed intensity of individual reflection, I_{avg} is average over symmetry equivalents).
Statistics for the highest resolution shell are shown in parentheses.
b = Figure of merit in phasing.
c = root mean square deviation from idealized geometry.
doi:10.1371/journal.pone.0099349.t001

Table 3-1. Data collection and refinement statistics

3.2.4 Biuret Hydrolase-Coupled Ammonia Detection Assay

A coupled-protein assay was developed as a highly sensitive method to measure CAH activity. CAH and mutant enzymes (25-28 μM) were incubated with 0.5 ml of 10 mM cyanuric acid in 0.1 M potassium phosphate buffer (pH 7) for time periods ranging from 0.5-65.0 h. The CAH reactions were stopped at four discrete time points by boiling the reaction tubes for ten minutes. Next, the reaction tubes were cooled to room temperature, a 5 μg aliquot of purified biuret hydrolase was added to each tube, and then the tubes were incubated at room temperature for 1 h. After incubation, ammonia was quantitated colorimetrically via the Berthelot reaction (66). Specific activity of the mutant CAHs was calculated at each time point based on 1 mole ammonia/1 mole biuret, and 1 mole biuret/1 mole cyanuric acid cleaved. Control samples without enzyme(s) were incubated in parallel to determine background levels of cyanuric acid hydrolysis or ammonia release. All samples were analyzed in triplicate.

3.2.5 Sequence Family Analysis

The CAH/barbiturase famly was updated, using methods described previously (51). As of October 5, 2013, 119 sequences were collected. These sequences were subdivided into ten clades based upon phylogenetic neighbor-joining trees produced by PHYLIP. Consensus sequences were determined for each of the clades, making distinctions between 100% and 50% conservation. Conserved residues were cross-referenced with the structure using the program Chimera (74).

3.2.6 Structural Analysis

The program LIGPLOT was used to analyze the tetramer interface (75). The programs Areaimol (76), within the program suite CCP4 (77), and CNS (78) were used to calculate accessible surface areas of amino acid residues and buried surface area.

3.2.7 Figure Preparation

Pymol was used for figure preparation in the manuscript. (http://pymol.sourceforge.net/)

3.4 Results

3.3.1 Overall Architecture of CAH

The crystal structure determination of a selenomethionine substituted CAH-inhibitor complex was accomplished using the single wavelength anomalous diffraction (SAD) phasing. The final 2.7 Å resolution model revealed a protein fold consisting of a β-barrel-like core with α-helices organized on the outside, forming a central cavity ~10 Å wide at its opening. (Figure 3-2A-B). At the bottom of the central cavity, the barbituric acid inhibitor is bound.

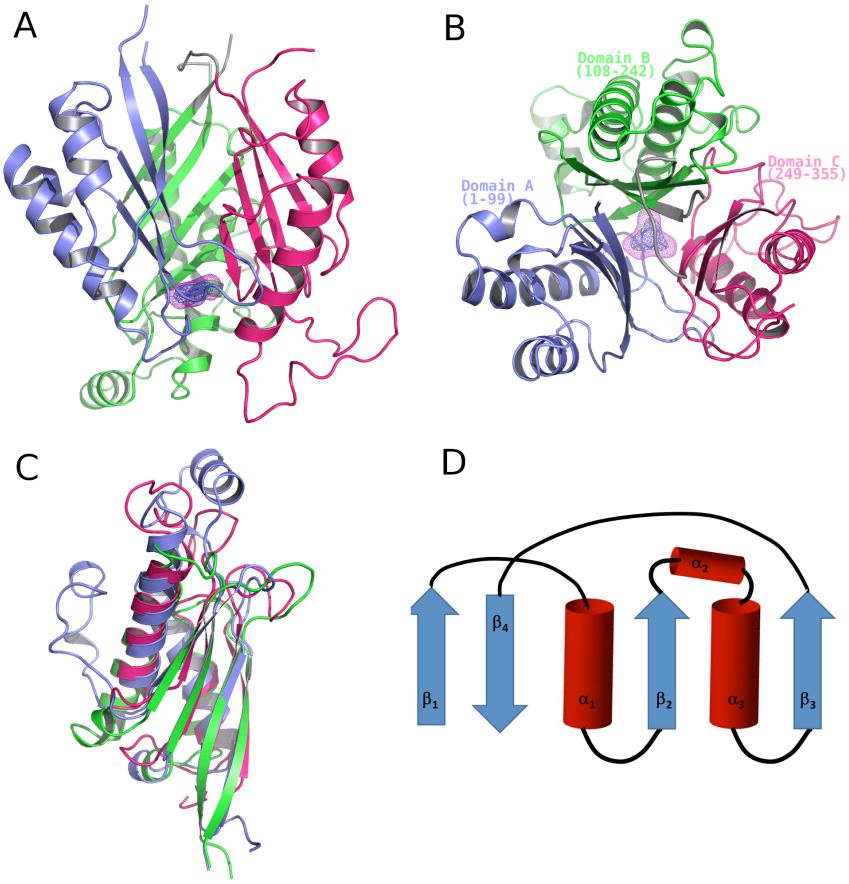


Figure 3-2. Monomer of CAH.

A. The side view of monomer structure. The selenium SAD-phased electron density for the barbituric acid inhibitor is shown in blue mesh, contoured at 2.0 σ. The simulated-annealing omit Fo-Fc map with the barbituric acid inhibitor omitted is shown in magenta mesh, contoured at 3.0 σ. Domains are depicted in different colors: Domain A, consisting of residues 1–99, in blue; Domain B, consisting of residues 108–242, in green; and Domain C, consisting of residues 249–355 in hot pink. B. The top view of monomer structure compared to A, looking through the center of the barrel along the pseudo three-fold axis. C. Overlay of the three domains of CAH. D. Depiction of the trimeric repeat domain fold in CAH. Red cylinders represent α-helices and blue arrows represent β-strands. α-helices 2 and 3 are present in the domain B, while domains A and C have a loop in place of α2.

In the CAH-barbituric acid complex crystal, the asymmetric unit contains a dimer related by a non-crystallographic two-fold axis (Figure 3-S1A) that further dimerizes via a crystallographic two-fold axis to form a tetramer with 222 point symmetry (Figure S1B).

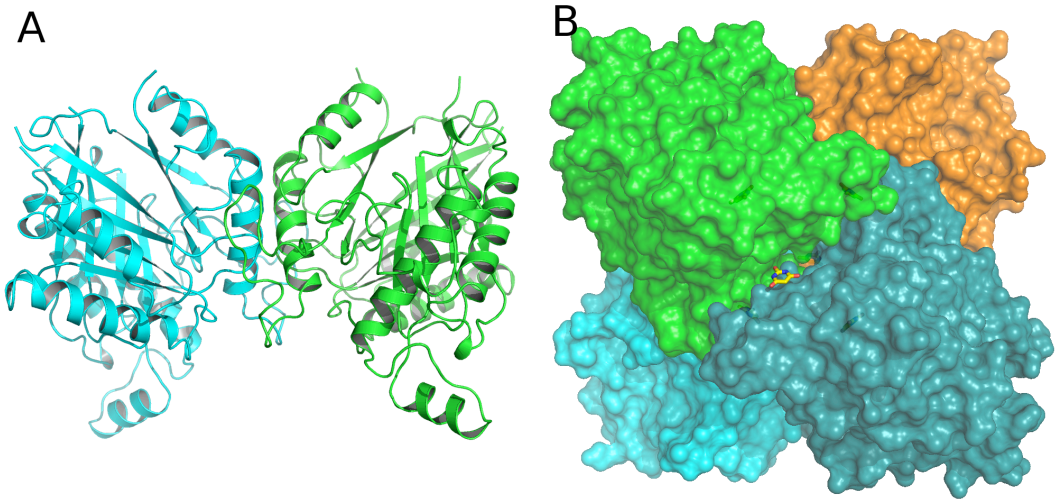


Figure 3-S1. The oligomerization of CAH. A. Dimer of CAH found in the asymmetric unit of the crystal. The two molecules are related by a non-crystallographic two-fold axis. The two protein chains are colored green and cyan. B. Tetramer of CAH. The CAH dimer shown in (A) further dimerizes to form a tetramer through a crystallographic two-fold axis.

The CAH tetramer is stabilized by extensive interactions between the monomers, burying a total of 3840 $Å^2$ of accessible surface area which is 12.41% of the total surface area. The molecular interface consists of a large number of polar interactions,

particularly in the core of the tetramer, involving a total of 88 residues forming 58 hydrogen bonds (counted with a distance cutoff of 2.5 ~ 3.2 Å) as well as several aromatic residues making hydrophobic interactions (Figure 3-S2). The formation of the CAH tetramer in the crystal is consistent with our earlier observation by size-exclusion chromatography (70).

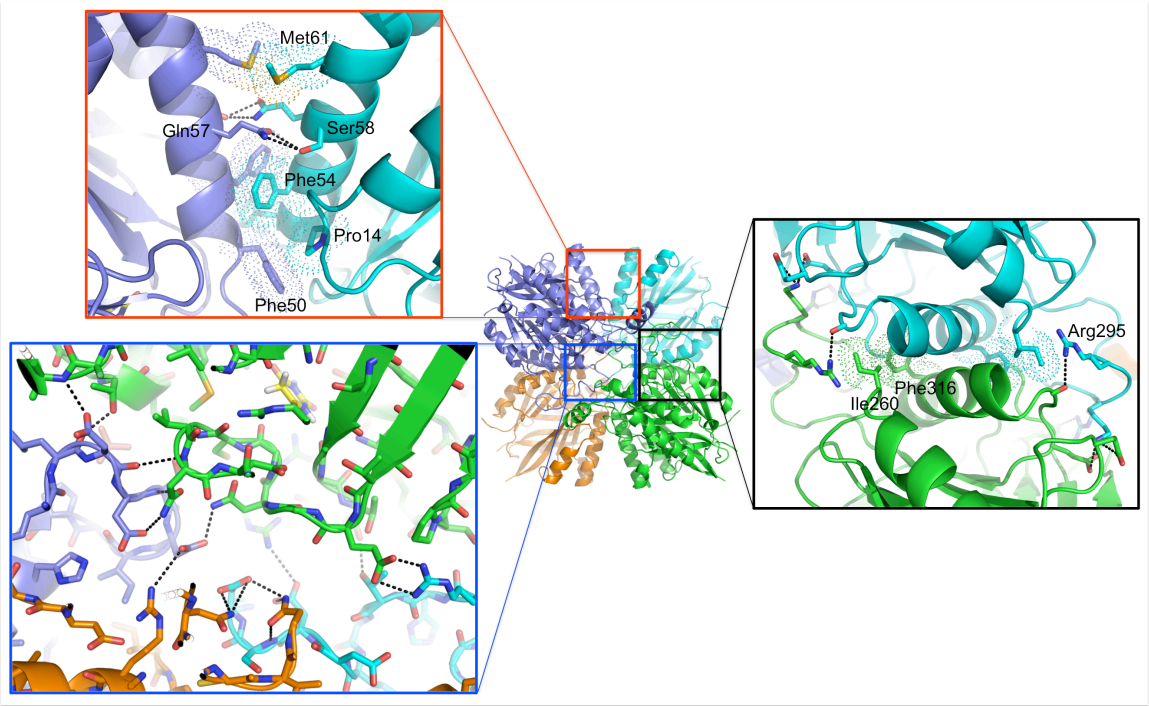


Figure 3-S2. Interactions in tetramer interface. The panel framed in blue is showing the abundance of polar interactions (hydrogen bonds and salt bridges) in the core of the tetramer. The other two panels (red and black) are showing the presence of both polar and hydrophobic contacts stabilizing different dimeric interfaces.

The structure of the CAH monomer exhibits an internal pseudo-three fold symmetry, consistent with our preliminary analyses of the X-ray diffraction data (Figure 3-2B) (70). Structural similarity is evident between three separate sections of the CAH monomer: domain A, 1-99 (Figure 3-2, blue); domain B, 108-242 (Figure 3-2, magenta); and domain C, 249-355 (Figure 3-2, cyan). The three domains are superimposable with root mean square deviations (RMSD) of alpha carbons of 3.2 Å, 2.0 Å, and 2.2 Å for domains A and B, domains A and C, and domains B and C,

respectively (Figure 3-2C), though protein sequence identity among these three domains is limited to 13-17%. Each domain of the protein has a conserved fold consisting of a 4-stranded β-sheet, and either 2 or 3 α-helices (Figure 3-2D).

3.3.2 Comparison of CAH to AtzD and Evolutionary Links to Other Protein Families

The pseudo three-fold symmetrical protein fold of *Azorhizobium* CAH has been recently described as the Toblerone or tricorne fold for AtzD from *Pseudomonas* sp. stain ADP, another member of the CAH/barbiturase protein family (55). These proteins share 51% sequence identity and differ in their enzyme kinetic values. Though k_{cat} values of both enzymes are close (AtzD: (73 ± 6) s^{-1} and *Azorhizobium* CAH (50 ± 9) s^{-1}), there is an order of magnitude difference for k_{cat}/K_m values, (AtzD: (3.2 ± 1.2) × 10^6 s^{-1} M^{-1} and CAH: (1.3 ± 0.6) × 10^5 s^{-1} M^{-1}), due to a ~10-fold difference in their K_m values (AtzD: (23 ± 7) μM and CAH: (370 ± 90) μM (51). The difference in K_m values implies that there may be differences affecting substrate-protein interactions, despite similar structures (RMSD for 334 Cα atoms, 0.86 Å) (Figure 3-3A) with highly similar active site residues (Figure 3-3B), which suggests a common catalytic mechanism.

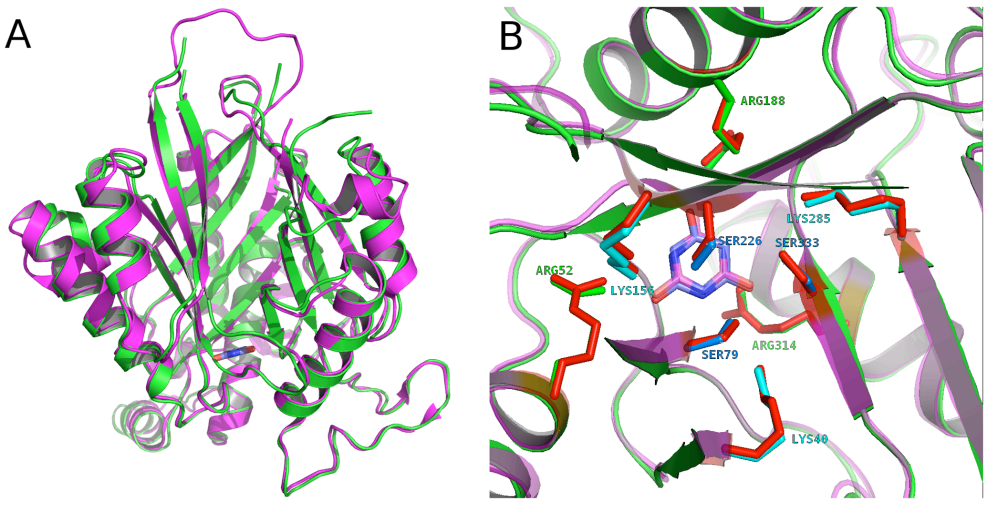


Figure 3-3. Structural comparison between CAH and AtzD.

A. Comparison of overall structures between CAH and AtzD monomers. CAH is colored in green while AtzD is colored in magenta. B. Comparison of the active sites between CAH and AtzD showing highly similar arrangement of the residues involved in substrate binding and catalysis. Residue labels are for CAH.

The residues positioned within 5 Å of the barbituric acid inhibitor are, in fact, identical between the two proteins except for Ser227 in CAH corresponding to Ala233 in AtzD (4BVR). The backbone nitrogens of these differing residues are 2.64 Å and 3.01 Å from a substrate carbonyl, respectively, and the backbone carbonyls are 2.82 Å and 2.71 Å from a substrate ring nitrogen. Ser227(CAH) and Ala233(AtzD) may, therefore, assist in substrate binding. This single residue difference is therefore not expected to alter substrate binding to the degree observed. Remote mutations are also known to affect substrate binding (79, 80), making more study into K_m differences between these two proteins necessary.

The DALI server (81) was used to identify proteins with folds similar to that of *Azorhizobium* CAH. Besides AtzD, which is closely related to *Azorhizobium* CAH structurally and functionally, the best scoring protein was the AroH-type chorismate mutase (CM) (PDB ID: 1XHO) (Figure 3-4A) with a Z-value of 7.8. This protein belongs to the YjgF-like SCOP structural superfamily (82, 83), which contains homotrimeric proteins with active sites located at each of the subunit interfaces. These active sites are on the external face of the barrel (Figure 3-4A), and the internal cavity or the head of the barrel is not known to have a catalytic function. This contrasts with CAH, which is a single polypeptide with a single active site on the inside of the barrel structure. Overlays of the individual domains of CAH on AroH-type CM (Figure 3-4B) yield an average RMSD of 2.88 Å, despite only ~9% sequence identity. The difference in the active site arrangements and low protein sequence identities between CAH and the YjgF-like superfamily proteins make it unclear whether this similarity is due to convergent or divergent evolution.

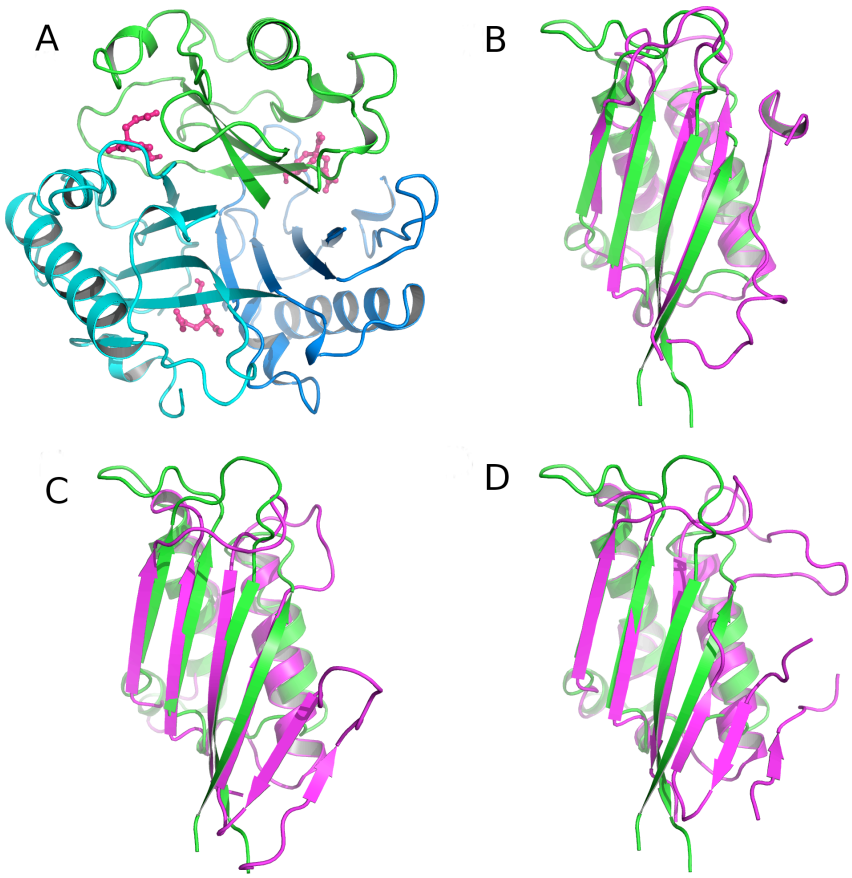


Figure 3-4. Comparison between CAH fold and the homotrimeric proteins in the YjgF-like SCOP structural superfamily.
A. The trimer of AroH-type chorismate mutase (CM) with active sites in each subunit interface that is accessed from the exterior of the barrel. B, C and D show the overlays of CAH domain A (residues 1–99, green) and a monomer of various structurally similar homotrimeric proteins (magenta): B. AroH-type CM (1COM). C. Proposed translation initiation inhibitor (PDB ID: 3GTZ). D. Putative aminoacrylate peracid reductase, RutC (PDB ID: 3V4D).

3.3.3 Active Site Architecture of CAH

In both subunits in the asymmetric unit of the CAH-barbituric acid complex crystal, a molecule of the highly analogous substrate analog, barbituric acid was found at one end of the barrel structure, unambiguously identifying the location of the active site and the mode of substrate-binding (Figure 3-2B). In addition, another barbituric acid molecule, albeit at lower occupancy, was observed on the surface of one of the CAH subunits near the opening of the potential substrate channel (Figure 3-9 A and D), which is discussed later in more detail.

The three-fold structural symmetry of CAH extends to the arrangement of residues that compose the active site. A highly organized, trimeric conglomeration of serines, lysines, and arginines surrounds the bound barbituric acid (Figure 3-3B), mirroring the three-fold symmetrical shape of the substrate. Three serine residues (Ser79, Ser226, Ser333) lie in close proximity to the carbonyl carbons of barbituric acid, suggesting that one could act as the nucleophile in a nucleophilic attack on the substrate. These residues are analogous to the three serines identified in AtzD (55). Because the serines are positioned below the plane of the barbituric acid ring and the bound inhibitor is not exactly centered in the active site, each of the three serines are at various distances from the carbonyl carbons within the bound barbituric acid (Figure 3-5B). Ser79 γO has distances of 3.65 Å, 4.66 Å, and 3.21 Å from the carbonyl carbons facing domains A, B, and C, respectively. Likewise, Ser226 γO is 3.65 Å, 3.57 Å, and 4.16 Å; and Ser333 γO is 5.59 Å, 4.71 Å, and 3.73 Å, to those same domain-facing carbonyl carbons, respectively.

Measurement of Bürgi–Dunitz (BD) angles was used to obtain greater insights into the residues potentially involved in nucleophilic attack on the substrate. The BD angle is the angle that defines the geometry between a nucleophile and a trigonal unsaturated center like a carbonyl carbon, with acceptable BD angles of 105 ± 5° (84). The BD angles for Ser226 attacking either of two carbonyl carbons in the bound inhibitor were within reasonable values (105.8° for the carbonyl facing domain A and 102.6° for the carbonyl facing domain B). In contrast, only one BD angle was reasonable for Ser79 (103.4° for the carbonyl facing domain A) or Ser333 (110.0° for the carbonyl facing domain B). Combining distances and BD angle data, it appears that Ser79 has an unreasonable BD angle for the carbonyl carbon facing domain C, despite the close distance. This was proposed to be the catalytic serine and carbonyl that is attacked for AtzD (55). The carbonyl carbon with the best BD angle for Ser79 has a distance (3.65 Å) which is a bit far unless there is some conformational change. Ser226 has the domain A and B-facing carbonyl carbons with reasonable BD angles and distances. Ser333, however, has a very long distance (4.71 Å) for the carbonyl with the only BD angle within reasonable parameters. The combination of distance and BD angles data

indicate that Ser226 is the most likely nucleophile of the three serines in the CAH active site.

As seen with AtzD, each serine is accompanied by a lysine residue (Lys40, Lys156, and Lys285, respectively for Ser79, Ser226, and Ser333), positioned within hydrogen bonding distances (Figure 3-5A). These lysines are positioned optimally to activate the serine residues. The guanidinium group of three arginine residues (Arg52, Arg188, Arg314), one from each domain, and the backbone amide nitrogens of the residues following the active site serines (Gly80, Ser227 and Gly334) are within 3.0 Å from the three carbonyl oxygens of the bound barbituric acid (Figure 3-5B), demonstrating potential protein-substrate interactions which could be involved in substrate activation and be the oxyanion hole for the carbonyl that undergoes nucleophilic attack.

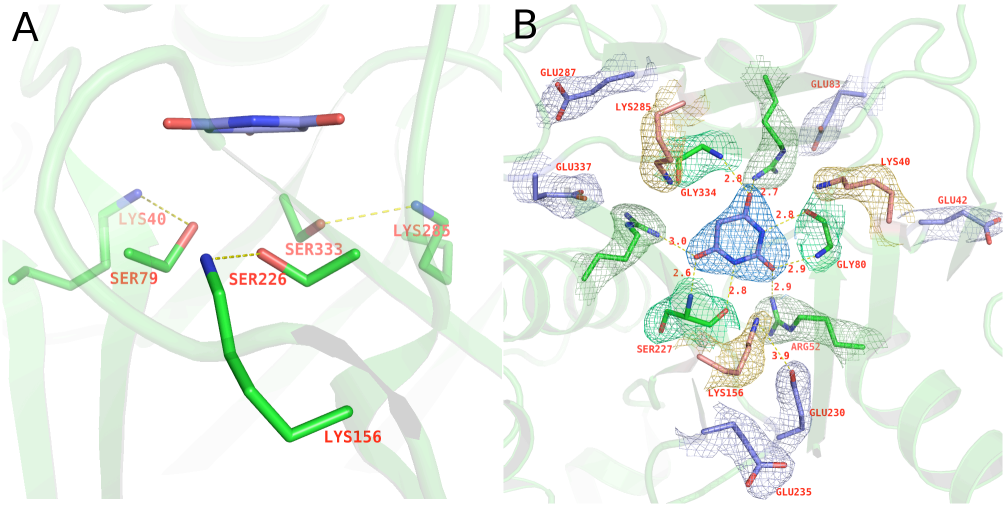


Figure 3-5. Trimeric conglomeration of CAH active site.
A. Three serine residues (Ser79, Ser226, Ser333) lie within hydrogen bond distances from corresponding lysine residues (Lys40, Lys156, and Lys285, respectively). Each pair shows the typical arrangement for the Ser/Lys dyad hydrolase. B. Three arginine residues (Arg52, Arg188, Arg314) and the backbone nitrogens from Gly80, Ser227 and Gly334 form hydrogen bonds with the carbonyl oxygens of the barbituric acid inhibitor. The simulated annealing composite-omit 2Fo-Fc map is shown in colored mesh, contoured at 1.5 σ.

3.3.3 Catalytic Properties of Mutants

To obtain further insights into catalytic roles of the active site residues, site-directed mutagenesis was performed. Development of a sensitive, coupled enzyme assay in this study specifically allowed for determination of activities that were not possible in

previous studies. All of the Ser to Ala mutants (S79A, S226A, S333A) exhibited reduced activities with cyanuric acid compared to the wild-type enzyme (Table 3-2).

	Specific activity
	(μmol/min per mg)
Wild type	11.5±3.5
S79A	$(6.0\pm1.3)\times10^{-7}$
S226A	$(2.7\pm0.4)\times10^{-8}$
S333A	$(1.7\pm0.4)\times10^{-7}$
K40A	$(8.0\pm2.8)\times10^{-4}$
K156A	$(1.0\pm0.2)\times10^{-3}$
K285A	$(1.4\pm0.1)\times10^{-3}$
R188K	$(1.5\pm0.1)\times10^{-3}$
R188Q	$(3.2\pm0.5)\times10^{-4}$

*Results were performed in triplicate.
doi:10.1371/journal.pone.0099349.t002

Table 3-2. Specific activity of wild type and mutant CAH on cyanuric acid as determined by the biuret hydrolase-coupled ammonia detection assay

The S226A mutant showed the greatest decrease, approximately 10^9-fold. The activities of S79A and S333A were 10^8-fold lower than that of wild type. These results indicate that all three serines play roles in CAH catalysis. The 20-40 fold greater effect of the S226A mutant is consistent with the possibility of this serine serving as a nucleophile. Moreover, when analyzing the primary sequences of all homologs within the CAH/barbiturase family, only one Ser/Lys dyad is absolutely conserved and corresponds to the Ser226/Lys156 dyad in CAH (Figure 3-6). In total, the results suggest that Ser226 acts as a nucleophile and that the other two serines are involved in binding and activation of the substrate.

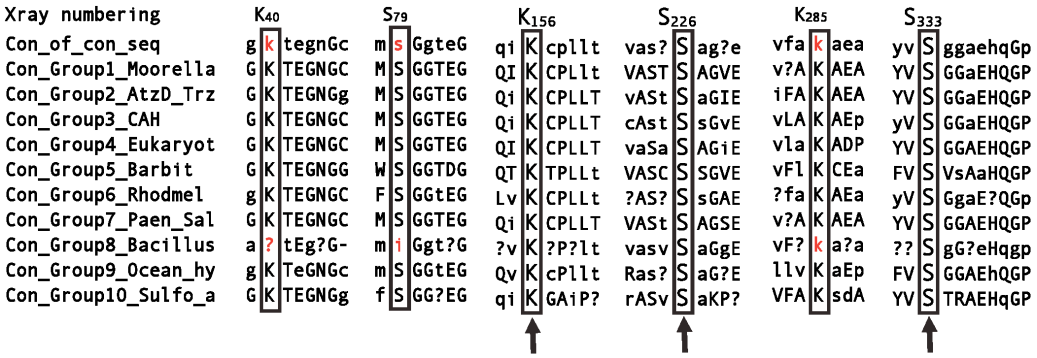


Figure 3-6. Sequence alignment showing conservation of three Ser/Lys dyads found within subgroups of the CAH/barbiturase family.
Subgroups of the updated family of 119 sequences were determined via phylogenetic clades. Consensus sequences were determined for each of the subgroups. Only sequences directly surrounding the dyads are displayed. Absolute, 100% conservation is represented by capital letters and greater than 50% conservation by lowercase letters. Columns in which no single residue was conserved at greater than 50% are depicted by a question mark. This usually occurred in regions of highly divergent sequence or where subgroups within a clade contained different residues. Residues equivalent to the Ser/Lys dyad residues are outlined with black boxes. Arrows point to Ser/Lys dyad residues with complete conservation across the family; residues not conserved in these positions are highlighted in red.

Mutation of any of the three active site lysine residues within hydrogen bonding distance to the active site serines (Lys40, Lys156, and Lys285) causes ~10^4-fold reduction of specific activity (Table 3-2). One of these lysine residues could act as a base to activate the serine nucleophile (Figure 3-3B). The similarity between CAH and Ser-Lys serine proteases (85) is currently unclear, but it is interesting that CAH is not inhibited by classical serine protease inhibitors like PMSF (data not shown) similar to many of the Ser-Lys dyad proteins. This differs from the AtzD enzyme in which inhibition was observed upon treatment with PMSF (55). Aside from the residues mentioned above, there are no other ionizable groups within 8 Å of the three serines except for Asp49 and Tyr331. The side chain oxygens (δO1 and δO2) of Asp49 are 7.5 -12.5 Å remote from the γO of the active site serines and Tyr331 side chain oxygen at 4.4 Å is 1.5 Å further away from the Ser333 γO than the nitrogen side chain of the corresponding Lys285 making either residue a less likely candidate for a general base catalyst.

To further assess the potential catalytic roles of these Lys residues, we examined their environments. All three lysines have extensive hydrogen bonding to backbone carbonyl groups in addition to respective serine residues. Lys40 is hydrogen-bonded to

Gly80 and Gly81; Lys156 with Met78, and Ser227; and Lys285 with Gly334 and Gly335 (Figure 3-7). Furthermore, these lysine residues are completely buried in the interior of the protein, with the accessible surface area (ASA) (73) for the ε-amino group of 0 $Å^2$ compared with the average ASA of other lysine residues in the CAH structure of 30.3 $Å^2$. This likely is responsible for adjusting the p*K*a of the lysine residues for catalysis.

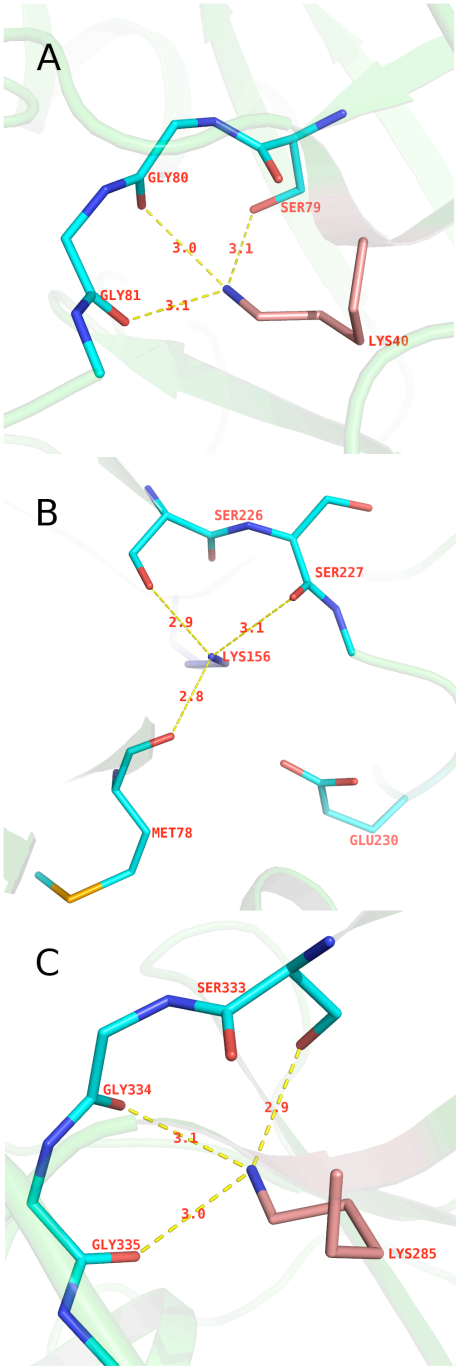


Figure 3-7. Hydrogen bonding and electrostatic interactions network around three lysines from each domain in CAH.
A. Lys40 forms hydrogen bonds with the backbone carbonyls of Gly80 and Gly81, and the Ser79 side-chain. B. Lys156 is hydrogen-bonded to backbone carbonyls of Met78 and Ser227, and Ser226 side-chain. In addition, Glu230 forms electrostatic interaction with Lys156. C. Lys285 is hydrogen bonded to backbone carbonyls of Gly334 and Gly335, and Ser333 side-chain.

Three arginine residues (Arg52, Arg188, Arg314) were within 3.0 Å of the substrate analog's carbonyl oxygens. To investigate the potential role of these arginines in catalysis, two mutants were created (R188K, R188Q). We measured the specific activities of two mutants with the biuret hydrolase-coupled ammonia detection assay and found that R188K and R188Q showed 5,000 and 23,000 fold reductions of specific activities respectively compared to the wild type (Table 3-2). The greater loss of activity in the R188Q mutant could indicate that the positive charge on the arginine is important for catalysis. The arginines could be involved in substrate binding by interactions with the carbonyl oxygens of the substrate and act as part of the oxyanion hole that stabilizes that charge on tetrahedral intermediates.

3.3.4 Proposed Catalytic Mechanism of CAH

The active site of CAH with the pseudo 3-fold rotational symmetry constructed from 3 structurally homologous domains seems optimal for catalyzing hydrolysis of the 3-fold symmetrical cyanuric acid. The classical serine hydrolases with the Ser-His-Asp catalytic triad, such as trypsin, initiate catalysis using a negatively charged oxygen atom of serine as a nucleophile. This negatively charged serine oxygen is formed by deprotonation of the gamma hydroxyl group by the catalytic histidine. By stabilizing a positively charged imidazole-ring of the histidine, the aspartate improves the base property of the histidine. Due to the absence of histidine and aspartate residues in the CAH active site and the different chemical environment of the Ser-Lys pair, the same enzymatic mechanism is not possible in the CAH.

Based on the results described above suggesting that Ser226 is the nucleophile, the following reaction steps could be envisioned (Figure 3-8). Lys156 deprotonates the gamma oxygen of Ser226 to facilitate nucleophilic attack on one of the substrate's carbonyl carbons. The first tetrahedral intermediate would ensue and the negative charge that results is stabilized by an oxyanion hole composed of an arginine and backbone amide as the carbonyl reforms, the ring C-N bond is cleaved. A water molecule then attacks the acyl-enzyme intermediate to regenerate the resting enzyme after product release.

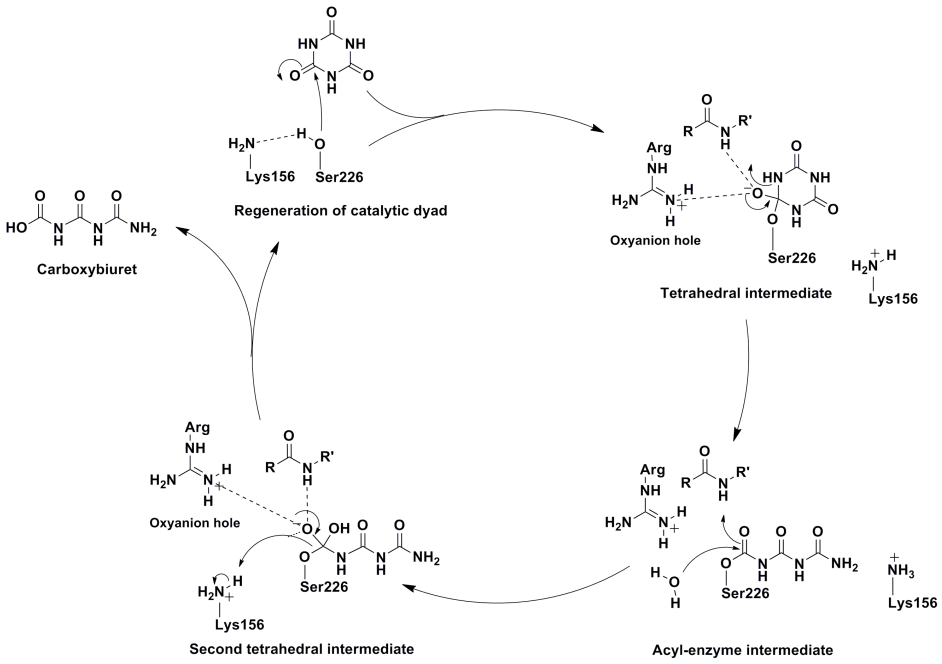


Figure 3-8. Proposed mechanism for CAH.
After substrate binding, Ser226 attacks the substrate carbonyl, producing a tetrahedral intermediate. Arginine and a backbone amide group form an oxyanion hole to stabilize the negative charge on intermediates. The substrate carbonyl reforms with the cleavage of the C-N ring bond. Water then attacks the acyl intermediate to release the product.

3.3.5 Surface-Bound Barbituric Acid and Channels to the Active Site

There are two entries ("front" and "back" doors) to the active site of CAH (Figure 3-9A-B). The central cavity (front door, highlighted by a red arrow) has a funnel-like shape and runs down to the active site along the pseudo 3-fold axis. This channel likely serves as the path for the substrate/product to diffuse in and out. The channel has a diameter of ~10 Å at its mouth and ~4 Å at the narrowest point, the latter of which is smaller than the size of the substrate, cyanuric acid. Therefore, some conformational changes would be required for the residues lining the wall of the cavity to allow passage of cyanuric acid. In addition, the opening of this channel to the solvent is partially blocked due to protein tetramerization (Figure 3-9D), which limits

the access of the entering substrate. In this regard, it is notable that, for one of the molecules in the asymmetric unit, a barbituric acid molecule is found bound on the protein surface near the opening of the central channel (Figure 3-9A, 9D). This non-active site barbituric acid molecule bound near the protein-protein interface might provide insights into how the substrate is attracted to the long and narrow tunnel leading to the active site.

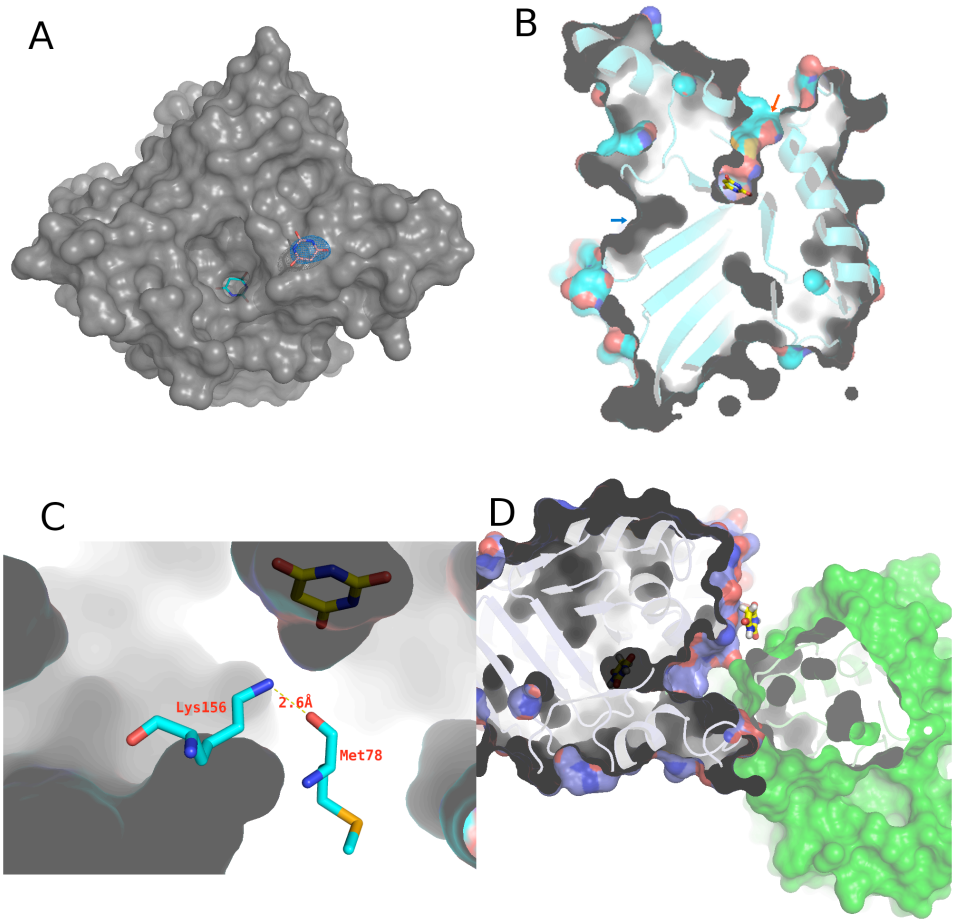


Figure 3-9. Front and back channels around the active site of CAH.

A. A view looking along the front channel down to the active site-bound barbituric acid. A second barbituric acid molecule, found on the surface is also shown in sticks with a simulated annealing composite omit 2Fo-Fc map contoured at 2.0 σ. B. Cut-away view of the CAH monomer. Barbituric acid is seen at the base of the front-side channel (red arrow). The back channel (blue arrow) is blocked by two hydrogen-bonds formed by Lys156 and Met78 (shown in sticks), and by Lys156 and Ser226, and therefore does not reach the substrate-binding site. C. Close-up view of the hydrogen bonds

between Lys and Met78. D. Barbituric acid molecule bound near the entry of the front side channel, on the tetramer interface cleft.

The cut-away view of CAH (Figure 3-9B) shows that it also has the "back door" between domains A and B, similar to that seen in AtzD (55). This back-tunnel (highlighted by a blue arrow in Figure 3-9B) is blocked by Lys156 that is hydrogen-bonded to the main-chain carbonyl oxygen of Met78 and the Ser226 side-chain, therefore is not reaching the active site (Figure 3-9B, 9C). The AtzD structure has an equivalent residue Lys162, hydrogen bonded to Met84 (55). It is unclear whether this tunnel in the back has any role in substrate/product passage.

3.5 Discussion

The reported crystal structure of CAH has internal three-fold symmetry, originating from the concatenation of three homologous domains. Each domain contributes equivalent active site residues, producing a three-fold symmetry within the active site. The symmetry of the cyanuric acid substrate extends this theme to the active site's substrate binding. Geometry and substrate analog positioning within the active site indicates that CAH likely utilizes a Ser-Lys dyad for catalysis. However, as there are three Ser-Lys dyads in the active site, additional analyses were required to address their specific roles.

With the highly sensitive biuret hydrolase-coupled ammonia detection assay, the catalyticcontribution of each serine to enzyme catalysis could be determined, in contrast to the study with AtzD in which no activity could be detected for any of the serine mutants. The greater loss of activity from the Ser226 mutant in the present study suggested that this could be a catalytic nucleophile. This assignment was further supported by sequence alignments between the cyanuric acid hydrolase/barbiturase family members which revealed that only the Ser226-Lys156 dyad was absolutely conserved across the entire family.

We have proposed that Lys156 serves as the general base to deprotonate Ser226 for a nucleophilic attack on cyanuric acid. The side-chain of lysine is typically protonated at physiological pH (around 7) due to its relatively high p*K*a (~10). This would not be

compatible with the proposed role for Lys156 as a base. The crystal structure, however, revealed that Lys156 is in a unique microenvironment compared to other lysines in CAH.

First, as mentioned above, the ε-amino groups of the three active site lysines (Lys40, Lys156, Lys285) are buried in the interior of the protein. When substrate binds to the active site, the ε-amino groups become completely inaccessible to solvent. Second, the hydrogen-bonding and electrostatic interactions involving Lys156 specifically provide it with a unique environment amongst the three lysines residues in the active site. The ε-amino group of Lys156 is located within ~3 Å from three oxygen atoms (side-chain hydroxyl group of Ser226 and main-chain carbonyl groups of Ser227 and Met78) as well as ~3.9 Å from the side chain carboxyl group of Glu230 (Figure 3-7B).

There are two glutamic acids around each of the three lysines from each domain (Glu42 and Glu83 around Lys40, Glu230 and Glu235 around Lys156, Glu287 and Glu337 around Lys285) within a distance of 6 Å, expanding the three-fold symmetrical arrangement of protein residues surrounding the active site (Figure 3-5B). Only Glu230 is absolutely conserved amongst the AtzD/barbiturase family members. This residue is within a distance to allow direct ionic interaction with Lys156 (~3.9 Å). This would indicate that the Ser-Lys dyad may in fact be a Ser-Lys-Glu active site with Glu230 stabilizing the positive charge on Lys156 after it takes the proton and activates the gamma oxygen of Ser226. This configuration of active site residues has not, to our knowledge, been identified in the literature, but mirrors the other assortment of residues used to activate a serine nucleophile within serine hydrolytic enzymes (85).

The active site of AtzD from *Pseudomonas* sp. stain ADP (55) also consists of three pairs of Lys/Ser dyads (Lys42/Ser85, Lys162/Ser233, and Lys296/Ser344) with an arrangement that mirrors that of CAH (Figure 3-3). Peat et al. proposed that Ser85 of AtzD is acting as a nucleophile in the catalytic cycle. Ser85 of AtzD corresponds to Ser79 from the domain A of CAH. However, Peat, et al. were unable to quantitate activity for any of the three serine mutants. With the additional data from the mutational studies provided by the current study, the evidence favors Ser226 of the

Azorhizobium CAH as the nucleophile. This residue corresponds to Ser233 of AtzD (Figure 3-3B).

Mutational studies also indicated Arg188 to have a role in catalysis. Here, it is proposed that the triplicate arginines (Arg52, Arg188, Arg314) are involved with substrate binding, providing a positive charge to interact with the three carbonyl oxygens.

The arginine associated with the carbonyl that is attacked by the nucleophile would also be part of the oxyanion hole that neutralizes the negative charge that forms on the oxygen of tetrahedral intermediates. Family analysis shows that Arg188 is conserved in all except two subgroups. The *Bacillus* subgroup has a lysine, and the barbiturase subgroup has a glutamine. Barbiturase uses barbituric acid as a substrate. Though this compound is analogous to cyanuric acid, it has a carbon in the place of one of the ring nitrogens. This difference results in only one of the carbonyl groups being surrounded on both sides by ring nitrogens, making it asymmetric. This difference could alter the substrate binding requirements for the enzyme and allow a glutamine in this position. Because the glutamine is less likely to stabilize the negative charge of the tetrahedral intermediate than an arginine, it is unlikely that the carbonyl near this residue (facing domain B) is the site of nucleophilic attack by the catalytic serine, at least in the case of barbiturase. The only absolutely conserved arginine is Arg52, making the carbonyl by this residue (in domain A) the most likely site for nucleophilic attack for the whole family.

In one of the CAH molecules in the asymmetric unit, a water molecule (A #509) was found within hydrogen bonding distance (3.0 Å) of the terminal side chain nitrogens of Arg188 This water is positioned 5.2 Å from one of the carbonyl carbons of the bound inhibitor and on the opposite face of the ring compared to the serine nucleophile. Another water (A #609; occupancy of 1 and a B-factor of 45.5 $Å^2$, a value comparable to nearby water molecules) is located directly above the barbituric acid ring in a pocket within 2.4-3.1 Å from the OG1 of Thr310, the NH1 of Arg314, and water (A #509). It is 4.2-5.0 Å from the carbonyl carbons of the substrate ring. With either of these waters, some conformational change that reduces the distance to

one of the carbonyl carbons would be required for involvement in catalysis, but more studies are required to confirm the function of specific waters in the crystal structure or the exact role of the arginine residues.

The structure and proposed mechanism presented here will serve as a framework for better characterizing the catalytic mechanism of CAH and engineering CAH for improved catalytic properties. Since cyanuric acid is a product of pool water chlorination, a CAH mutant with enhanced activity would be industrially useful for bioremediation and water conservation efforts. Developing an *in situ* swimming pool water treatment system would prevent the need for complete water exchange to remove excess cyanuric acid in pool water. Based on the current CAH structural data and gel-filtration analysis (70), the biological unit of CAH is a tetramer (Figure 3-S1B). AtzD was also reported to form a stable homotetramer, based on the X-ray crystallographic and SAXS analyses (55). Although tetramerization could be contributing to the stability of CAH, it does not seem to provide any benefit in enzyme catalysis and rather appears to partially block the substrate entry pathway. Therefore, dissociation of CAH tetramer into monomers or dimers by modifying the molecular interfaces could potentially allow more efficient substrate diffusion to the active site and increase the catalytic turnover of the enzyme.

Acknowledgments

We thank J. Nix of the Molecular Biology Consortium at the Advanced Light Source (ALS) for help in data collection. Computer resources were provided by the Basic Sciences Computing Laboratory of the University of Minnesota Supercomputing Institute. We also thank Kwon-Ho Hong for valuable discussion.

Chapter 4

Redesign and X-ray structure of a commercial, thermophilic cyanuric acid hydrolase from *Moorella thermoacetica* : Enzyme-substrate intermediates trapped in action

4.1 Summary

CAH from *Morella thermoacetica* (MCAH) is currently being tested for commercial use in removing cyanuric acid from swimming pool water and its X-ray structure is of considerable interest. MCAH was engineered using the Surface Entropy Reduction (SER) method and loop redesign to render it more likely to crystallize. X-ray structures were obtained from crystals grown with barbituric acid to 1.5 Å resolution and with cyanuric acid to 2.9-3.2 Å resolution. The barbituric acid grown crystals contained 4 subunits in the asymmetric unit and all showed electron density consistent with bound barbituric acid. The cyanuric acid grown crystals contained 24 subunits in the asymmetric crystallographic unit and different stages of the reaction with cyanuric acid were observed in different subunits. The electron density for an acyl-enzyme intermediate formed between γO of the Ser231 and the carbonyl carbon of the cyanuric acid substrate was identified and support the hypothesis that Serine/Lysine dyad from domain B is involved in nucleophilic attack. Also, carboxybiuret and biuret were found bound in other subunits, providing "snapshots" of different stages of the reaction.

4.2 Introduction

Pioneering work on the serine hydrolase protein family (86-89) has established the idea of a catalytic triad consisting of serine, histidine, and aspartate in numerous enzyme active sites, making this a paradigm for amide bond hydrolysis. In proteases following this mechanism, the serine is deprotonated by histidine, where the histidine is a general base, and the aspartate serves in charge stabilization of the positive charge

of histidine during catalysis. However, another class of the serine hydrolases that use a Ser/Lys dyad was recently identified (90-93). Two Cyanuric Acid Hydrolase (CAH) structures (54, 55) revealed three pairs of Ser-Lys dyads forming the catalytic core of the enzyme. Because the three Ser/Lys dyads are arranged symmetrically, it is essential to determine the Ser/Lys pair that plays a role as the catalytic dyad to understand the exact catalytic mechanism of the cyanuric acid hydrolase/barbiturase family members. When the AtzD cyanuric acid hydrolase structure was determined in the presence of cyanuric acid by Colin Scott's group (55), they proposed that Ser85 from the first domain is the nucleophile that attacks a substrate carbonyl and opens the ring, although this hypothesis was not supported by strong empirical data. In our previous paper on the cyanuric acid hydrolase from *Azorhizobium calinodans* (54), we proposed Ser226 from the second domain is the nucleophile based on indirect evidence from several experimental approaches.

The data supporting our proposal consisted of a mutational study showing the most drastic loss of activity was from replacing Ser226 with an alanine, observing an allowed Burgi-Dunitz angle between the serine oxygen and a carbonyl carbon of the substrate, and bioinformatics arguments (84). Although these constituted important indirect evidence, the catalytic mechanism of CAH remained an unanswered question because there was no direct demonstration of the role of any of the serines in the reaction.

A thermophilic enzyme from *M. thermoacetica* (94) is at present being evaluated for commercial use to remove cyanuric acid from swimming pools and is being mutagenized to increase its resistance to hypochlorite and other oxidants. We were interested in the crystal structure of this enzyme so as to gain insights into the structural basis of its enhanced thermostability relative to other cyanuric acid hydrolases (94). However, attempts to crystallize the native protein failed. By applying a surface entropy reduction method and modifying a loop domain (the active site remained intact) in accordance with the published AtzD structure, we were able to crystallize the *Moorella* cyanuric acid hydrolase in the presence of cyanuric acid.

The present data sets obtained from the crystals with cyanuric acid have 24 subunits comprising the asymmetric unit (ASU). Different subunits in the asymmetric unit contained cyanuric acid, a modified serine representing an acyl-enzyme intermediate, and the known reaction products carboxybiuret and biuret. The observed acyl-enzyme intermediate structure supports our previous hypothesis that Ser226 of CAH is the nucleophile since we could observe extending electron density from Ser231 (the Ser that corresponding to Ser226 in our modified CAH) to a carbonyl carbon in the substrate. Hence, our results provided the first direct evidence to identify the cyanuric acid hydrolase active site nucleophile and thus give the most concrete insight yet into the mechanism of these enzymes.

4.3 Materials and methods

4.2.1 Designing CAH_Morella for crystallization

Surface entropy reduction methods were applied to crystallize *Morella* CAH. Both XtalPred (95) server and SERp (96) server predict residues 130 to 134, a part of one of the loops in domain A, as candidates for alanine replacement to reduce entropy. The loop region "LKFDCCPPAEELAK" in MCAH that was predicted to be flexible was replaced with the loop region "IRSDDEMDR" from AtzD which is ordered in the AtzD crystal structure (Figure 4-1). The DNA sequence corresponding to the designed amino acid sequence was synthesized and cloned in pET28B+ by GenScript (Piscataway, NJ).

4.2.2 Protein expression and purification

The engineered full length *Morella* CAH (RMCAH) gene was expressed in *Escherichia coli* BL21 (DE3) cells (Invitrogen) from a construct based on pET28 b+ vector in LB medium (10 g Bacto tryptone, 5 g yeast extract, 10 g NaCl per liter) supplemented with 50 mg ml_1 kanamycin. Cells were grown in 37 °C in LB medium and induced with 0.5 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration for 30 hours at 20 °C.

After collection by centrifugation (6400 x g, 10 min, 277 °K), the cells were resuspended in binding buffer (50 mM Tris–HCl pH 7.0, 200 mM NaCl, 5 mM imidazole). Sonication was used for cells disruption and the cell lysate was centrifuged at 63,988 x g for 1 h and 30 min at 277 °K. After centrifugation, the cell lysate was applied onto a column packed with Ni–NTA resin (Thermo scientific) that had been equilibrated with the binding buffer. After washing the unbound protein off of the column with washing buffer (50 mM Tris–HCl pH 7.0, 200 mM NaCl, 10 mM imidazole), the target enzyme was eluted with four column volumes of elution buffer (50 mM Tris–HCl pH 7.0, 200 mM NaCl, 200 mM imidazole). For further purification, gel filtration fast protein liquid chromatography (GE Healthcare AKTA FPLC) on a Superdex 200 10/300 GL gel-filtration column (GE Healthcare) equilibrated with a running buffer consisting of 20 mM Tris–HCl pH 7.0, 200 mM NaCl was applied. The molecular weight of the enzyme was determined by NanoDrop 8000 UV–Vis spectrophotometer (Thermo scientific) and concentrated with ultrafiltration (EMD, Amicon®).

4.2.3 Crystallization

The enzyme was concentrated to 12mg ml^{-1}. Co-crystallization of RMCAH with barbituric acid or cyanuric acid was carried out in hanging drops at 20 °C. The initial crystallization condition consists of 3.0 % tacsimate pH 5.0 and 18% of PEG 3350 (Hampton research) and was further optimized. Needle shape crystals appear in 2 weeks for harvest. For data collection, crystals were transferred into the cryoprotectant solution composed of the reservoir solution and 20 % glycerol (v/v) and flash-cooled.

4.2.4 Data collection and processing

The diffraction of flash-cooled crystals was measured on beamline 24-ID-E at the Advanced Photon Source (APS). Data were processed with program HKL2000 (64). Crystals belonged to monoclinic $P2_1$ space group.

4.2.5 Structure determination and refinement

The diffraction images were integrated and scaled to 2.9 Å respectively with HKL2000. PHENIX (97) was used for further crystallographic calculations. Molecular replacement was performed with Phaser (98) to get initial atomic models using the previously determined structure CAH (4NQ3). To get the final model, iterative cycles of model building and refinement were conducted utilizing COOT (72) and PHENIX refinement module. PROCHECK (73) was used to check the model validity.

4.2.6 Structural analysis and figure preparation

Ligplot (75) and Areaimol (76) in CCP4 (24) were used to analyze the interfaces between chains and accessible surface areas of amino acids respectively. All of the figures in the chapter were prepared with pymol (http://www.pymol.org/).

4.4 Result

4.3.1 Redesign of MCAH

We have cloned, expressed, and purified the redesigned CAH from *Morella* (RMCAH). Because wild type MCAH was resistant to crystallization, we applied surface entropy reduction strategy to enhance crystallizability. The SERp server predicted Q103, E104, and K107 as candidates for alanine mutations for reducing entropy. Additionally, we replaced one of the potential loop sequences of the *Morella* CAH (MCAH) (280- 292) with a corresponding loop sequence from AtzD (46) to make RMCAH that has a sequence shorter by 5 amino acids than the wild-type MCAH (Figure 4-1).

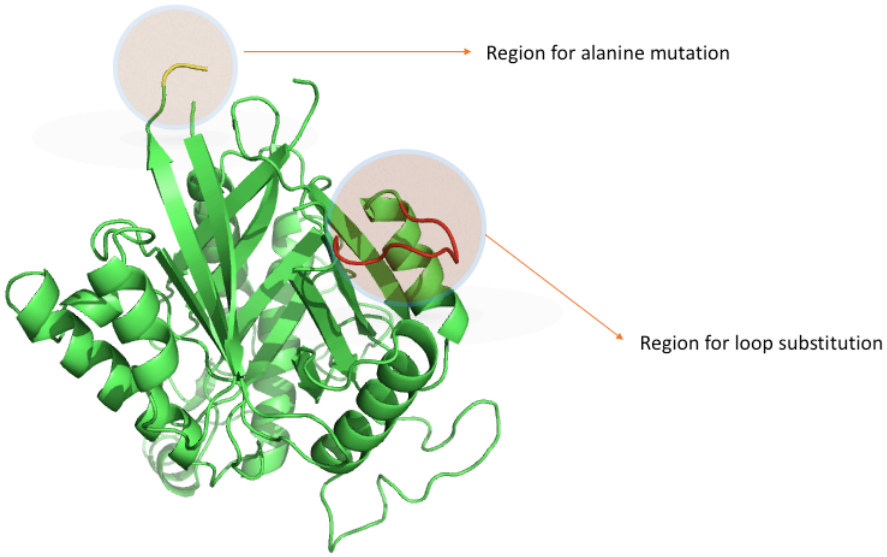


Figure 4-1. The scheme for the mutations of loops for enhancing crystallization of MCAH.
The colored loops in shaded circles indicate target amino acid sequences for mutations. The yellow colored loop was subjected to alanine mutation while red colored loop was replaced with the corresponding loop in AtzD.

The MCAH sequence has a higher sequence identity with the AtzD sequence (58%) than the ACAH sequence (56%). Also in AtzD, there is a visible helix in the region that the ACAH structure could not be built due to likely disorder. Hence we used the region of the AtzD gene to replace the analogous MCAH region and ensure that the two protein sequences on either side connect correctly. We ran this new sequence through SERp and got no additional sites than the ones in the original protein sequence.

4.3.2 Overall RMCAH_CYA and RMCAH_BAR architecture

We determined the crystal structure of RMCAH in both cyanuric acid (CYA) and barbituric acid bound forms by molecular replacement. The monomer of the BAR bound form of ACAH was used as a search model to find the phase of RMCAH_Bar and RMCAH_CYA. The R_{work} and R_{free} values of the 1.5 Å resolution RMCAH_Bar structure that consists of 4 monomers in the ASU are 0.1513 and 0.1780 (Table 4-1) respectively and 97.39% of residues are located in favored regions of the

Ramanchandran plot (99). RMCAH_CYA structures which contains 24 monomers in the ASU was refined to 3.1 Å with an R_{work} and R_{free} of 0.2213 and 0.2614 respectively (Table 4-1) and 95.99% of residues are located in the allowed regions of the Ramanchandran plot (99)

	RMCAH-CYA	RMCAH_BAR_1	RMCAH-BAR_2
Space group	$P2_1$	$P2_12_12_1$	$P2_1$
Cell parameters (*a, b, c)* (Å) (α, β, γ) (°)	134.50,171.02,194.70 90.00, 90.10, 90.00	81.67,89.18, 205.46 90.00, 90.00, 90.00	134.50,171.02,194.70 90.00, 90.10, 90.00
Resolution (highest shell), (Å)	100.00-3.16 (3.23-3.16)	100.00-1.54 (1.57-1.54)	100.00-3.19 (3.31-3.23)
No. of measured intensities	998839	1421024	1045031
No. of unique reflections	147537	211507	141576
Redundancy	6.8	6.7	7.4
Completeness (%) (last shell)	99.6 (92.7)	96.7 (65.1)	99.9 (99.8)
$I/\sigma(I)$ (last shell)	9.1 (1.3)	17.0 (0.9)	7.3 (1.1)
R_{merge} % (last shell)[a]	0.160 (1.00)	0.089 (0.94)	0.244 (1.00)
Refinement Statistics			
Resolution range (Å)	50.00-3.16	81.81-1.54	97.33-3.19
No. of reflections used	149896	211343	146731
No. of protein atoms	64560	12320	64740
No. of waters	72	1222	83
R_{work}[b]	0.224	0.133	0.215
R_{free}[c]	0.265	0.160	0.233
rmsd bond lengths (Å)	0.002	0.009	0.004
rmsd bond angles (°)	0.461	1.20	0.708
Ramachandran Plot Favored/ allowed Outlier (%)>	99.66 0.34	100.00 0.00	99.72 0.28
Average B-factors protein/other ($Å^2$)	94.43/63.56	19.68/32.46	85.49/89.04

Table 4-1. Data collection and Refinement Statistics

[a] $R_{merge} = \Sigma |I - (I)| / \Sigma I$, where I is the intensity of an observed reflection and (I) is the average intensity of multiple observations.
[b] $R = \Sigma||F_{obs}| - |F_{cal}||/\Sigma|F_{obs}|$.
[c] $R_{free} = \Sigma||F_{obs}| - |F_{cal}||/\Sigma|F_{obs}|$, where F_{obs} is from a test set of reflections (5 % of the total) that are not used in structural refinement.

In both monoclinic structures, the twenty-four monomers had the same overall fold within the asymmetric unit (RMSD deviation <0.1 Å for all common C alpha atoms).

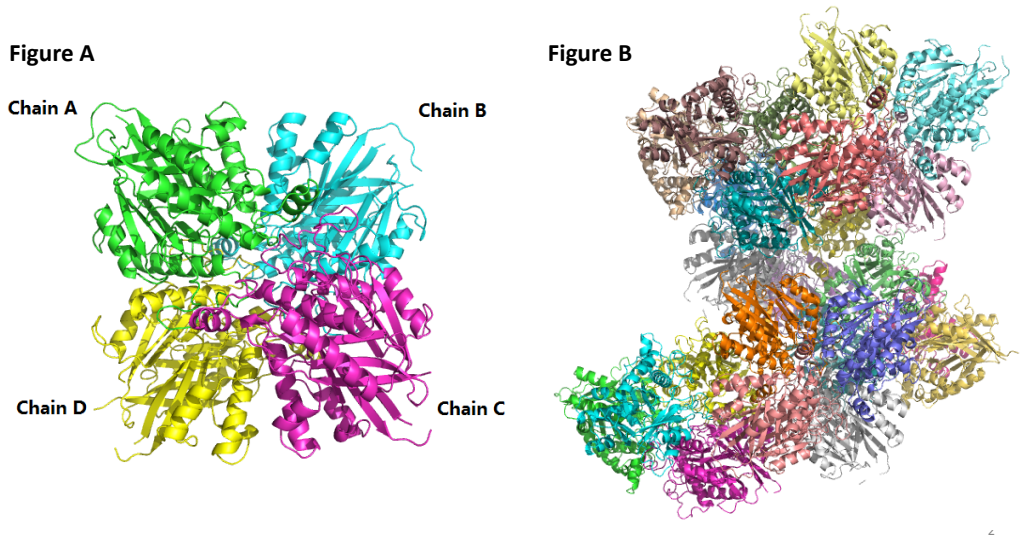


Figure 4-2. Oligomerization states of RMCAH. A. The figure shows the tetramer state in solution. B. 24 monomers exist in asymmetric unit.

Both RMCAH_CYA and RMCAH_BAR structure consists of a β-barrel-like core with α-helices analogous to the ACAH and AtzD structures in the same barbiturase/cyanuric acid hydrolase family. For crystals grown with cyanuric acid, different shapes of electron densities in the active sites indicate different ligands were bound. The different ligands in different RMCAH_CYA monomers were identified as: cyanuric acid, the acyl-enzyme intermediate, carboxybiuret, and biuret. By contrast, for crystals grown with barbituric acid, essentially identical electron densities for barbituric acid were found in the 24 active sites of RMCAH_BAR.

Because the biological oligomerization states of ACAH and RMCAH are assumed to be a tetramer based on size exclusion chromatography, the interfaces of a tetramer (Chain A, B, C, and D) were analyzed with respect to the tetramer itself and another tetramer interfaced in twenty-four molecules of the RMCAH oligomers (Fig. 4-2). Like the ACAH tetramer, RMCAH tetramer is also stabilized by extensive hydrogen bonds and hydrophobic interactions forming interfaces. The tetramer interface

contains a total of 45 hydrogen bonds (a distance cut off of 3.2 Å) with extensive hydrophobic interactions involving 106 amino acids. The dimer interface between chain A and B is formed only via hydrophobic residues, while other dimer interfaces (AC, AD, BC, BD) form via both hydrogen bonds and hydrophobic interactions. The ABCD tetramer further contacts with another tetramer through three hydrogen bonds and seventeen residues involved hydrophobic interactions in the 24 subunit oligomer.

The structure of the RMCAH monomer has an internal pseudo-three fold symmetry whose characteristics also are evident in monomers of ACAH and AtzD (54, 55).

4.3.3 Structure Comparison between RMCAH and other structurally determined CAHs

Structural alignments were done between the two previously determined CAH structures, ACAH and AtzD, and chain R of RMCAH_CYA. Structural alignments indicate that RMCAHs shares a high degree of structural similarity with ACAH and AtzD at RMSD values of 0.562 and 0.531 Å for all common C alpha atoms respectively (Fig. 4-3).

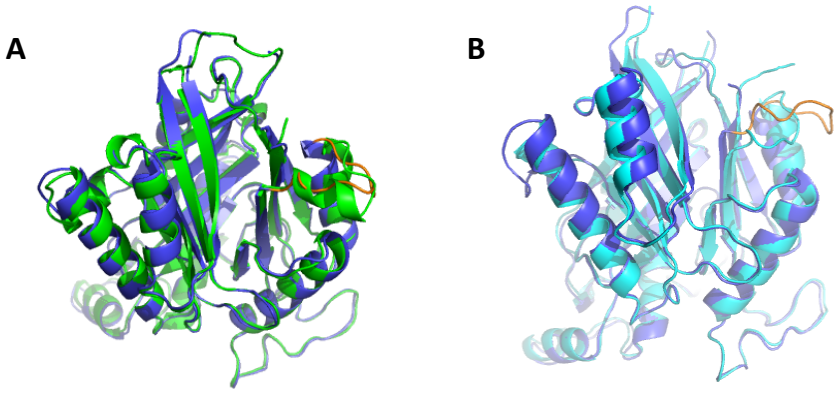


Figure 4-3. Structural alignments between RMCAH and ACAH or AtzD**.** A. Structural alignment of RMCAH (Blue) and AtzD. (Green). B. Structural alignment of RMCAH (Blue) and ACAH (Cyan). Orange colored loops in RMCAH coordinate indicates the replaced loops from AtzD for engineering purpose.

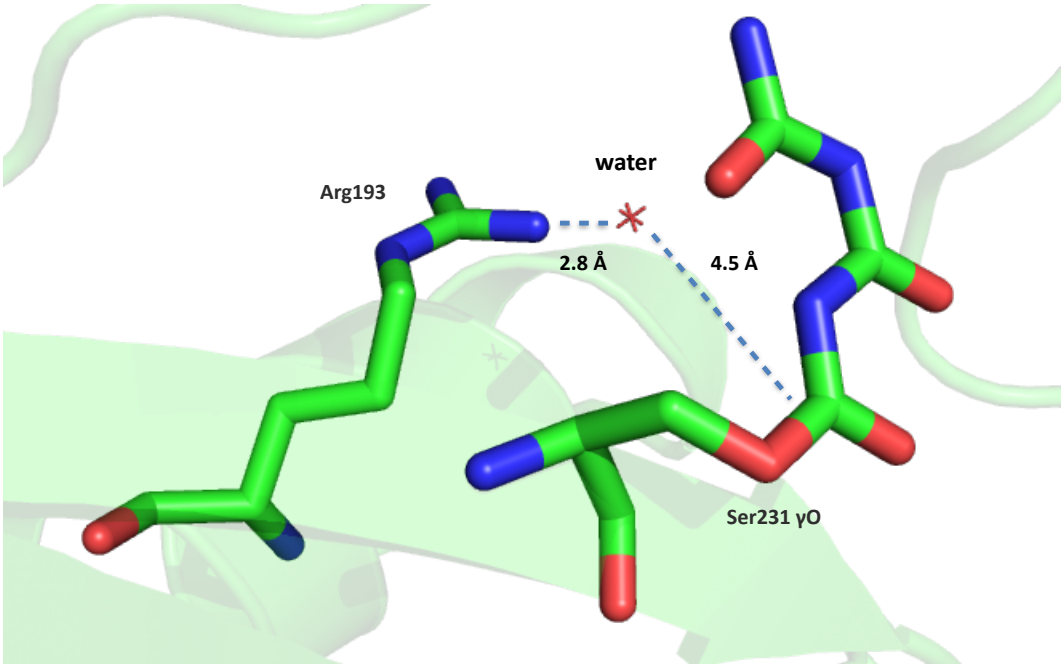
4.3.4 Identification of an Acyl Enzyme intermediate

Although the deacylation step of the catalytic triad serine hydrolase mechanism has been shown to be rate limiting, only a few crystal structures of acyl-enzyme intermediate for serine hydrolases have been reported and they used different strategies to trap the intermediate (100-105). To determine the crystal structure of the acyl-enzyme intermediate of RMCAH_CYA, RMCAH was purified and crystallized with cyanuric acid at pH 5.0. The RMCAH_CYA structure consisted of twenty-four molecules per asymmetric unit harboring different stages of the reaction. The chains C and I are of particular interest because there is clear electron density with trigonal planer geometry extending from the γO of the Ser231 and fits the previously suggested acyl-enzyme intermediate (54). The planarity in the ester bond of the acyl-enzyme intermediate in RMCAH_CYA structure is also observed in other paper that reports an acyl-enzyme intermediate of a serine hydrolase. This suggested that the nucleophilic attack by water prior to formation of the 2nd tetrahedral intermediate formation is not restricted by geometric constraint of the acyl-enzyme intermediate. The distance between γO of the Ser231 and the carbonyl carbon in the ester intermediate is 1.4 Å, only consistent with a covalent bond. This covalent bond distance indicates that the electron density represents the acyl-enzyme intermediate. If the electron density implies another reaction intermediate bound in the active site, we would expect a significantly greater distance between γO of the Ser231and the carbonyl carbon of the ester intermediate.

A water molecule is also located between the nitrogen of the terminal side chain in Arg 193 and the carbonyl carbon in the ester intermediate at distances of 3.1 Å and 4.1 Å respectively, potentially in position for nucleophilic attack on the ester. If this is the second nucleophile that is activated to a hydroxide ion, it is currently unclear what residue acts to activate the water. In the previously solved ACAH structure (54), a water molecule was found within H-bond distance of the nitrogen of the guanidinium group in the side chain of Arg188 that corresponds to Arg193 in RMCAH. We suggested the possibility of the water's role in the deacylation step in the catalysis. Hence, enzyme assays were conducted with R188K and R188Q mutants and showed

the activity decreases by 10^{-3} to 10^{-4} fold (54). The position of the water molecule in the RMCAH_CYA structure reinforces our previous observation that the Arg helps to position the water molecule that participates in the deacylation step (Fig. 4-4).

In the other facets of the mechanism, the side chain -NH_2 of Lys 161 is located within hydrogen bond distance (1.9 Å) of the carbonyl carbon in the ester intermediate. Presumably, Lys161 plays a role as part of an oxyanion hole that neutralizes the developing negative charge on the ester carbon. In chain W, any electron density that corresponds to a reaction intermediate was not found and indicating the complete departure of the product biuret from the active site.



4

Figure 4-4. Active site environment around the water molecule proposed to be the nucleophile in the deacylation step. The potential water for nucleophilic attack is located in between side chain of Arg193 and the carbonyl carbon of ester enzyme intermediate retaining distance 2.8 Å and 4.5 Å respectively.

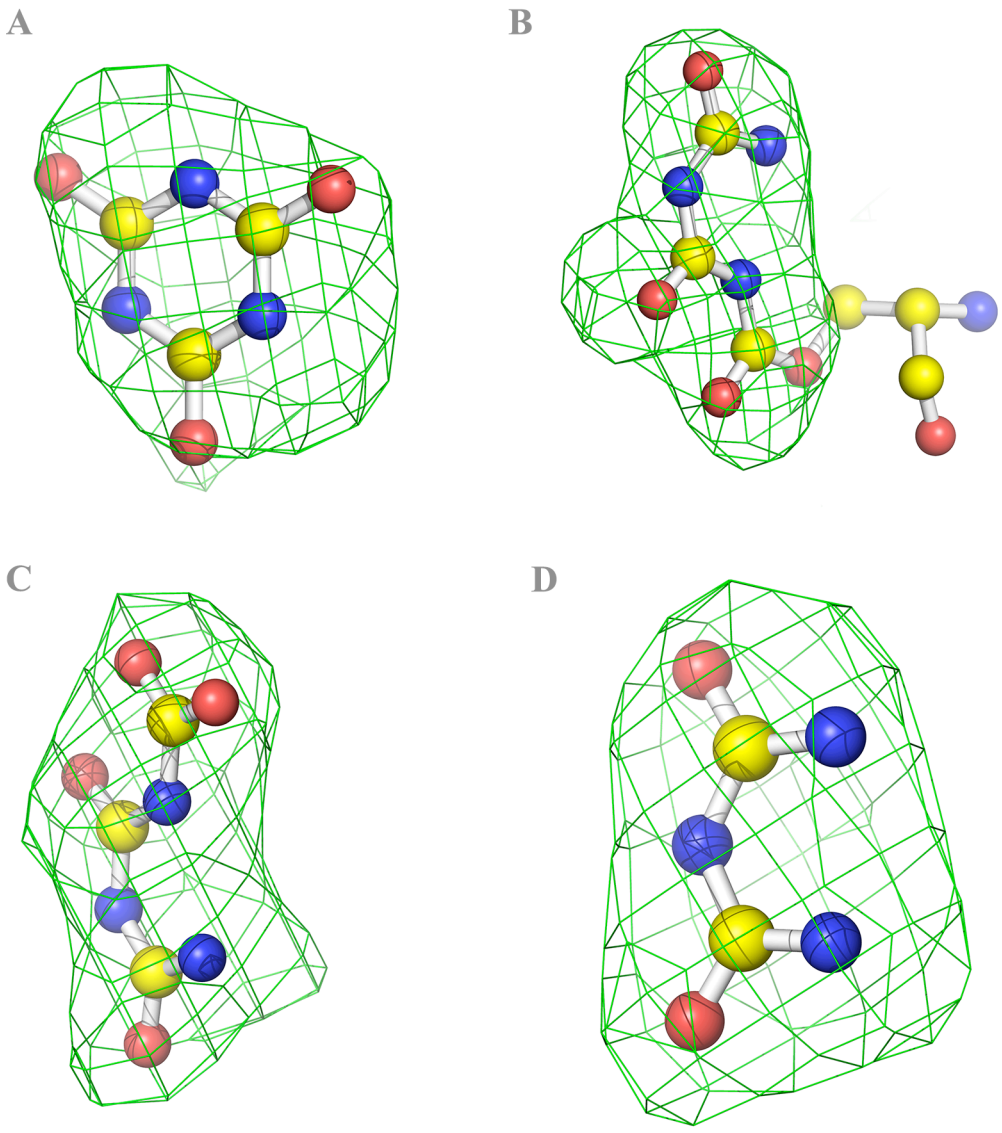


Figure 4-5. *Fo-Fc* omit electron density map contoured at 3.0 σ level, depicted in green mesh and calculated the A) Cyanuric Acid, B) Covalent intermediate, C) first product Carboxybiuret and D) final product Biuret omitted respectively and shown in ball-and-stick representation.

The electron densities for carboxybiuret, a direct hydrolysis product of the CAH catalysis, and biuret, the final product after spontaneous decarboxylation of carboxybiuret were also identified in chain L, N, F and chain X respectively in the 24 molecules of RMCAH_CYA crystal (Fig. 4-5).

4.3.5 Active site analysis of RMCAH_CYA

The electron density of the active sites of chain P and chain R were consistent with the substrate, cyanuric acid. For simplicity, the active site of chain R is analyzed. There are 3 pairs of Ser-Lys dyads that are contributed by each of the three domains in the active site. As observed in the ACAH and AtzD structures, each of the three serines (S83, S231, S342) lie within hydrogen bond distances (<3.2 Å) from the accompanying three lysines (K40, K161, K294), extending 3-fold symmetry of the enzyme active site on the 3 fold symmetry of the substrate's structure.

As seen with the other two CAHs, the three lysines (K40, K161, K294) in the active site were located within hydrogen bond distances for accepting the protons from the hydroxyl groups of three serines (S83, S231, S342) like other Ser/Lys dyad containing hydrolases (90, 105). We also analyzed the Burgi-Dunitz angles and bond lengths between γ oxygen from the potential catalytic serines and carbonyl carbons in CYA to obtain insight about the environment around three serines that potentially are involved in the nucleophilic addition within the reaction (Table 4-2). S231 has at least one optimal combination of angle and bond distance while the other two serines do not have any optimal combinations of angles and bond length. This matches the observation that S231 is the serine that supports the acyl-enzyme intermediate electron density found in chain C.

	Bond distances (corresponding Burgi-Dunitz angles)
S83	3.6 Å (101°), 2.9 Å (70°), 4.7 Å (155°)
S231	2.3 Å (106°), 3.0 Å (131°), 2.7 Å (121°)
S342	3.2 Å (44°), 4.5 Å (110°), 5.4 Å (151°)

Table 4-2. Burgi-Dunitz angles and their corresponding bond distances. The values in the table indicate Burgi-Dunitz angles and bond distances between γ oxygens of serines and carbonyl carbons of cyanuric acid in cyanuric acid bound monomer.

4.5 Discussion

In a previous paper describing the crystal structure of ACAH, we proposed that the serine in domain B (S231 in RMCAH and S226 in ACAH) is the likely nucleophile, while the other two serines (S83, S342 and S79, S333 in RMCAH and ACAH respectively) are involved in substrate activation and positioning. First, the mutational study with S79A, S226A, S333A of ACAH using the sensitive biuret hydrolase coupled ammonia detection assay showed that the activity decrease of the S226A mutant was the greatest by 20-40 fold. Second, the sequence alignment within subgroups of the CAH/barbiturase family indicated that only the Ser/Lys dyad pairs in domain B from different sub-groups are conserved throughout the cyanuric acid hydrolases. Third, measurement of Bürgi–Dunitz (BD) angles showed that S226 was the only serine that has an optimal angle and distances for nucleophilic attack on one of the carbonyl carbons on the bound substrate molecule.

Despite this indirect evidence, there has been a disagreement in the literatures as to which serine among the three serines in the active site plays the role of the nucleophile. Scott and co-workers reported the crystal structure of AtzD and proposed that the serine in domain A may be the nucleophile based on the polarized electron density of the cyanuric acid molecule that they observed positioned close to the serine in domain A. However, their interpretation was not consistent with our observations based on the above three experimental points. Our observations also have limitations. Although the serine 226 mutant had the lowest activity, mutating any of the active site serines resulted in significant activity loss. Moreover, measurement of Bürgi–Dunitz (BD) angles could not conclusively determine the nucleophilic serine, since the measurement was calculated from a single structure and there may be movement of the protein not determined from the single structure at medium resolution.

Hence, we thought that trapping an acyl-enzyme intermediate formed between CAH and CYA would provide the most conclusive evidence as to which serine participates as a nucleophile in the catalytic mechanism. For this purpose, we crystallized and solved the crystal structure of CAY treated or Bar bound RMCAH after we redesigned MCAH for better crystallization using the surface entropy reduction method.

There are 24 monomers in RMCAH_CYA and RMCAH_Bar per ASU. Importantly, we could observe different states within the catalytic cycle including substrate, acyl-enzyme intermediate, and product bound states all in the same crystal (Fig. 4-7). Of particular interest is the observation of the acyl-enzyme intermediate that has the planar ester group. The acyl-enzyme intermediate data clearly support our hypothesis in the previously solved ACAH that the serine in domain B participates in the nucleophilic attack of the ring opening reaction.

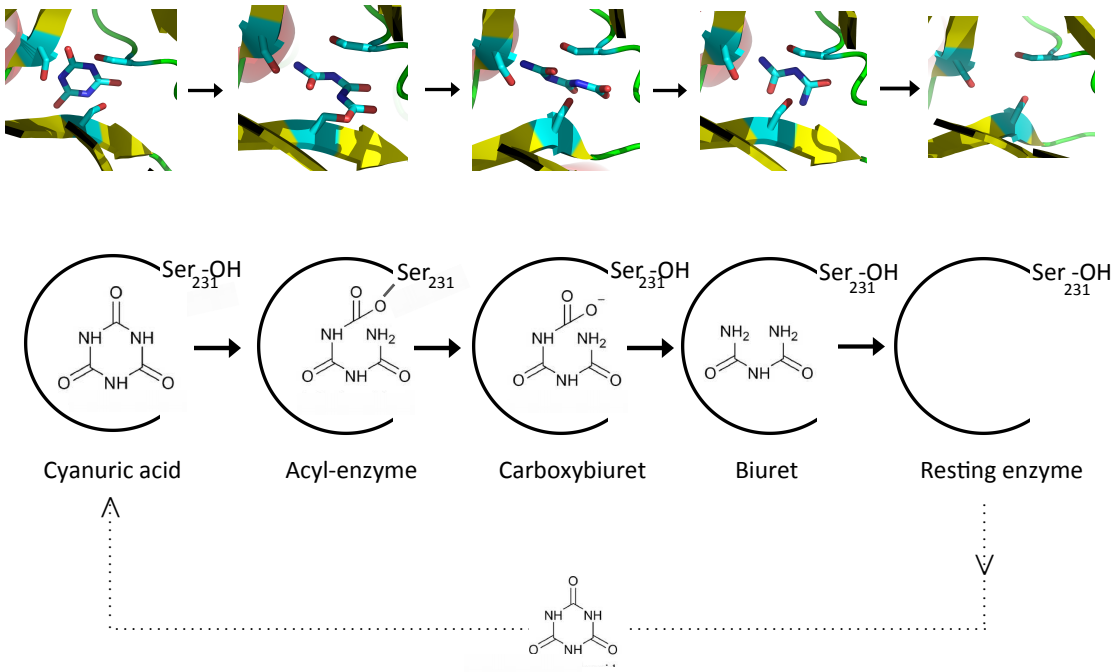


Figure 4-6. The entire catalytic cycle of CYA hydrolysis. The hydrolysis steps by CAH were drawn based on the electron densities for the different states of intermediates trapped in single crystal.

MCAH is currently used commercially for degrading cyanuric acid because it is from a thermophilic bacterium and so the protein is highly stable in normal use and storage. The melting point (Mp) for MCAH is 11 °C higher than RMCAH based on circular dichronism experiments (Mp of MCAH : 84 °C, Mp of RMCAH : 73 °C). The loop substitution and/or alanine mutations introduced during the construct redesign

appeared to negatively affect the thermostability of the RMCAH. However the Mp of RMCAH is still higher than the Mp of ACAH (69 °C) and it has activity comparable to the MCAH.

There are a small number of papers that describe the crystal structures of acyl-enzyme intermediates from different serine hydrolases in a low pH medium such as 5 (103, 105). The low pH crystallization condition may stabilize the protonated state of lysines or arginines in the active sites, thus preventing water from being activated to form 2nd tetrahedral intermediate.

Chapter 5

Biodegradation of 1,3,5-triazine by *Acinetobacter* sp. trz

5.1 Summary

We identified how unsubsituted *s*-triazines [1,3,5-triazine] are degraded to formate and NH_3 through abiotic and biotic pathways using ^{1}H-NMR and HPLC analysis. 1,3,5-triazine is first degraded to formamidine, formamide, and formate non-enzymatically. Then formamidine and formamide is metabolized to NH_3 and formate by the newly isolated strain *Acinetobacter* sp. Trz.

5.2 Introduction

s-triazine derivatives have been used as herbicides, resin, and dyes. There have been a significant number of studies about how substituted *s*-triazines are degraded (2-4). For example, as stated in chapter 1, hydrolytic *s*-triazine pathways converge on cyanuric acid as a common intermediate (43). Then, cyanuric acid undergoes complete mineralization via three enzymes, AtzDEF, as found in *Pseudomonas* sp. ADP (46, 47). However, it is not known how the unsubstituted *s*-triazine ring might be metabolized. To explore this question, enrichment culturing was conducted with 1,3,5-triazine as the sole nitrogen source. A single bacterial isolate was obtained from the enrichment culture. The morphological and 16S rRNA gene sequencing analysis of the bacterium indicated that the isolate was an *Acinetobacter* and it was denoted as strain "Trz". ^{1}H-NMR and HPLC analysis revealed that 1,3,5-triazine first undergoes abiotic degradation to formamidine, formamide and ammonia. Then, formamidine is further degraded to formamide that is decomposed to NH_3 and formate, resulting in availability of nitrogen for *Acinetobacter* sp. Trz.

5.3 Materials and methods

5.2.1 Chemicals, media, strains, and culture conditions

All chemicals were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted. *Escherichia coli* MG1655 was obtained from Professor Michael J. Sadowsky. For R-medium preparation, R-salt containing (g/L) KH_2PO_4, 136.09; $MgSO_4.7H_2O$, 0.4; $FeSO_4.7H_2O$, 0.01; $CaCl_2.2H_2O$, 0.0294 was used with 1 ml of the trace element solution for R-medium. The trace element solution for R-medium consists of (mg/L) ZnSO4.7H2O, 0.01; $MnCl_2.4H_2O$, 0.0003; H_3BO_4, 0.003; $CoCl_2.6H_2O$, 0.002; $CuCl_2.2H_2O$, 0.0001; $NiCl_2.6H_2O$, 0.0002; $Na_2MoO_4.2H_2O$, 0.0003. For enrichment culture using R media (106), 0.2 % of glucose, sodium acetate, and disodium succinate were used as carbon sources. R media agar plates were prepared with R-medium containing 1.5% Noble agar. The preparation of Ammonium Mineral Salts medium (AMS) followed the protocol noted in Parales et al (107) with 0.2 %of sodium acetate as a carbon source and other nitrogen sources replacing ammonium sulfate as noted. The enrichment cultures and isolates were grown on a shaker at 28 °C and *E. coli* was grown on a shaker at 37 °C.

5.2.2 Enrichment cultures

Soil samples were taken from a campus garden. Initially 5g of soil was added to 50 ml of R minimal medium supplemented with each 0.2% of glucose, sodium acetate, and disodium succinate as carbon sources and 2mM 1,3,5-triazine as sole nitrogen source and shaken at 220 rpm at 28 °C. The 50ml culture was transferred every 2-3 days for total of 12 transfers. 0.2% of sodium acetate was chosen as the sole carbon source. The microorganism was isolated by streaking the last transfer culture on R-medium agar plate supplemented with 2 mM 1,3,5-triazine. Distinct colony types were streaked onto additional plates to obtain homogeneous size and shape of colonies on a plate. Pure cultures were used to inoculate aliquots of the enrichment medium to verify the ability to grow on 1,3,5-triazine as the sole nitrogen source.

5.2.3 Identification of the strain by 16S rRNA gene sequencing

To identify the genus of the enrichment isolate, 1460nt of the 16S rRNA gene was amplified with primers 27F and 1525R (108). The parameters for PCR reaction using Phusion polymerase (NEB, Ipswich, MA) was as follows: pre-heating at 94°C for 5 min, 32 cycles of denaturation at 94°C for 1 min, annealing at 57°C for 1 min and extension at 72°C for 2 min, and a final extension for 10 min at 72°C. Amplified DNA was purified with a Gel Extraction kit (QIAGEN, Valencia, CA) and sequenced by ACGT, Inc (Wheeling, IL) with the PCR primers and primers 342R and 1114F (109). The sequences were assembled into a single contig, which was aligned against type strain sequences on the Ribosomal Database Project (RDP) server (110). A phylogenetic tree was generated on the RDP servers using the weighted neighbor-joining tree building algorithm (111).

5.2.4 Identification of pathway intermediates

Proton nuclear magnetic resonance spectroscopy (NMR) spectra were acquired using a Varian Unity Inova 400 MHz NMR with VnmrJ 2.2D software and a Nalorac 4 Nuc probe. Samples were analyzed in PBS buffer containing (g/L) NaCl, 8; KCl, 0.2; Na_2HPO_4, 1.44; KH_2PO_4, 0.24 plus 10% deuterated water (D_2O).

High-pressure liquid chromatography (HPLC) analysis was performed on a Hewlett-Packard HP 1100 series chromatographic system equipped with a diode array detector and interfaced with an HP Chemstation platform. Compounds were separated on an Agilent (Santa Clara, CA) Zorbax Eclipse Plus C18 column (4.6 by 250 mm, 5 μm particle size) with an isocratic mobile phase of 95:5 phosphate buffer (5mM, pH 2.5): acetonitrile at a flow rate of 1 ml/min with the detector absorbance set at 200 nm. Retention times were the following: formamidine, 2.3 min; formamide, 2.6 min; formic acid, 3.1 min; and acetic acid, 3.9 min.

5.2.5 Degradation assays

For resting cell degradation assays and crude extract activity assays, *Acinetobacter* sp. and *E. coli* cells were grown in AMS medium with different N-sources or in LB to

an OD600 nm of 0.5. Cells were washed once with filter sterilized PBS buffer. Cells for resting cell assays were resuspended in PBS buffer supplemented with 10mM formamidine to an OD600 nm of 1.0-2.0 and shaken at 225 rpm at 28 °C or at 37 °C in case of *E. coli.* Supernatants were filtered with 0.2μm filters (Pall life science, Port Washington, NY) and stored at -80 °C before analysis. Samples were thawed and analyzed by NMR or diluted ten-fold with PBS buffer for HPLC analysis.

Crude cell extracts were prepared by resuspending cells harvested from 50 ml cultures in 1 ml PBS buffer. Then the cell suspension was sonicated gently using a sonicator (Branson, Danbury, CT) equipped with a microtip for 30 seconds total pulsed 5 seconds on and 5 seconds off at 20% power. The sonicated cells were centrifuged centrifuged at 14,000 rpm for 30 minutes at 4 °C and supernatants were filtered with 0.2 μm filters. Total protein concentration was determined using the Bio-Rad (Hercules, CA) protein assay based on the Bradford method (112). Activity assays were initiated by adding 0.5mg of extract to 1ml PBS buffer containing 10mM formamidine. The reactions were incubated at room temp for 3 days. Samples were filtered and diluted for HPLC analysis.

5.4 Result

5.3.1 Isolation, identification and growth conditions of strain Acinetobactor

After 11 serial cultures in R-salt minimal medium with 1 g of soil from the garden, a single strain was isolated that grew alone in media with 1,3,5-triazine as the sole nitrogen source. Alignment of the isolate's partial 16S rRNA gene placed it in genus *Acinetobacter* and phylogenetic analysis conducted with *Acinetobacter* type strains showed highest sequence identity to *A. lwoffii* DSM 2403 (98%, 1436/1460 nt identical) (Figure 5-1). The isolate was subsequently named *Acinetobacter* sp. Trz.

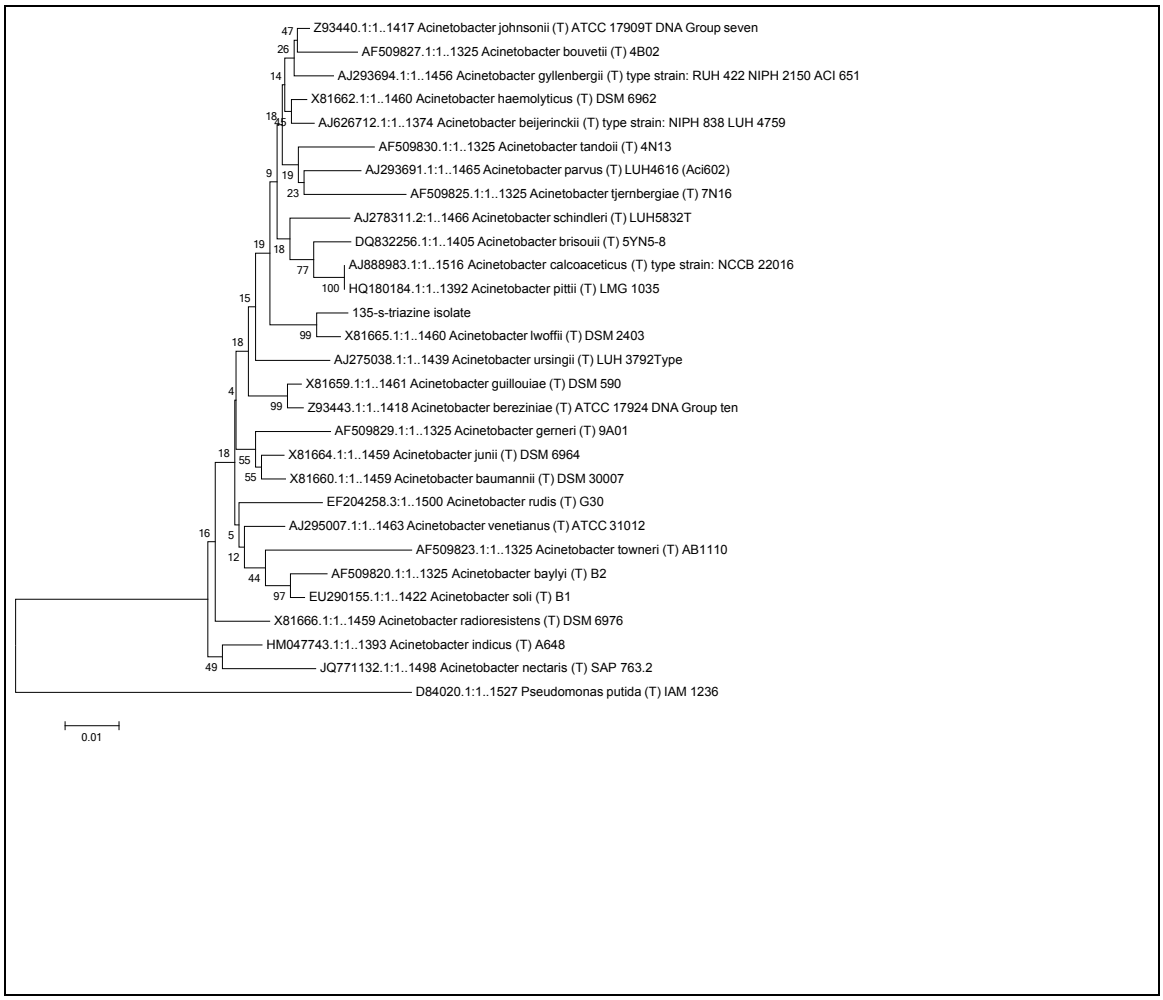


Figure 5-1. Phylogenetic tree for *Acinetobacter* sp. Trz based on 16S rRNA gene sequencing

5.3.2 Abiotic degradation of 1,3,5-triazine

1,3,5-triazine was assumed to be as stable in solution as are substituted *s*-triazines and the fact that pyrimidine and pyridine are highly stable. However, analysis of a day old 1,3,5-triazine solution by UV spectroscopy showed consistent decrease of the maximal peak absorbance at 217 nm. 1,3,5-triazine, 1,3,5-triazine degradation rates were tested in different pH to try and stabilize the compound. The degradation rate of 1,3,5-triazine was slowest in pH 7.0 (Figure 5-2). 1,3,5-triazine was degraded 50% and 80% within 2 hours at pH 8.0 and pH 8.5 respectively.

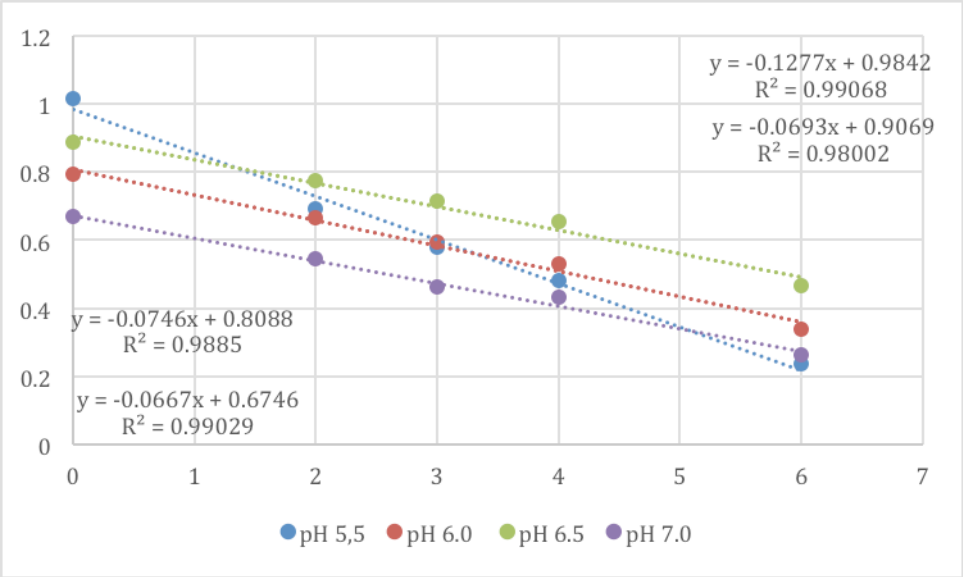


Figure 5-2. Degradation rate of 1,3,5-triazine at different pH

To confirm its instability and determine degradation products, a proton NMR spectrum of a freshly prepared 1,3,5-triazine in PBS buffer was monitored over time. The chemical shifts revealed that 1,3,5-triazine in solution spontaneously decomposed to three one-carbon compounds, which were identified as formamide (7.88 ppm), formic acid (8.27 ppm), and formamidine (7.62 ppm) by analyzing commercial standards of those compounds after 22 hours. There were additional chemical shifts at 8.73 and 8.15 ppm besides the chemical shifts for formamidine, formic acid, and formamide after 2 hours of incubation of 1,3,5-triazine solution. Hence, those two chemical shifts may represent N-formyl formamidine or other two carbon reaction intermediates.

5.3.3 Use of 1,3,5-triazine degradation products as nitrogen sources for growth by Acinetobacter sp. Trz

The growth yields of *Acinetobacter* sp. Trz were tested with different concentrations of 1,3,5-triazine and NH_4Cl. Because 1,3,5-triazine contains 3 nitrogens and NH_4Cl has 1 nitrogen in their structures, 6 mM NH_4Cl on the graph equals to 2 mM s-triazine (Figure 5-3). The isolate OD values in both 1,3,5-triazine and NH_4Cl cultures increases as more N sources are available in the R-medium. As a result the growth of *Acinetobacter* sp. Trz was due to metabolism of one or both of the nitrogen-containing decomposition products, formamidine and formamide.

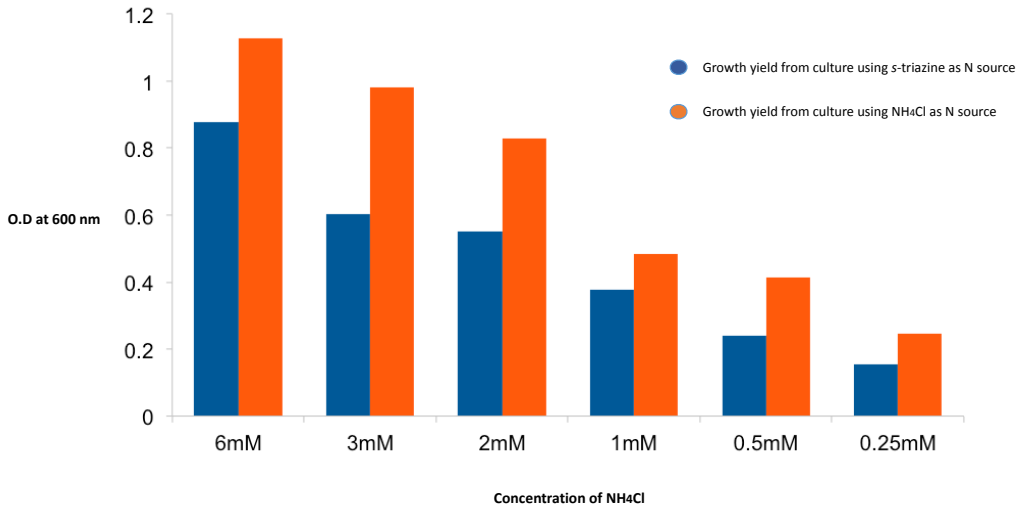


Figure 5-3. The growth yields of *Acinetobacter* sp. Trz at different concentrations of nitrogen sources

5.3.4 Formamidine degradation acceleration in the presence of Acinetobacter sp. Trz *or E. coli MG1655 resting cells*

To determine if formamidine hydrolysis occurred faster in the presence of cells, resting cells of *Acinetobacter* sp. Trz or *E. coli* MG1655 were incubated with formamidine and analyzed by HPLC. Formamidine degradation rates were much faster in treatments with cells as compared to controls without cells. After 48 h, strain

Acinetobacter sp. Trz grown on LB or in minimal medium with ammonium reduced the formamidine peak area to 19% or 25% of time = 0 h peak areas, respectively, and *E. coli* MG1655 cells grown on LB reduced the formamidine peak area to 15% of the time = 0 h peak area (Figure 5-4). The formamidine peak area without cells was still 85% of the time = 0 h peak area after 48 h. These results indicate that formamidine hydrolysis was faster in the presence of cells. It is interesting that *E. coli* MG1655 also has foramidine degradation activity while the cell could not grow on AMS medium supplemented with foramidine.

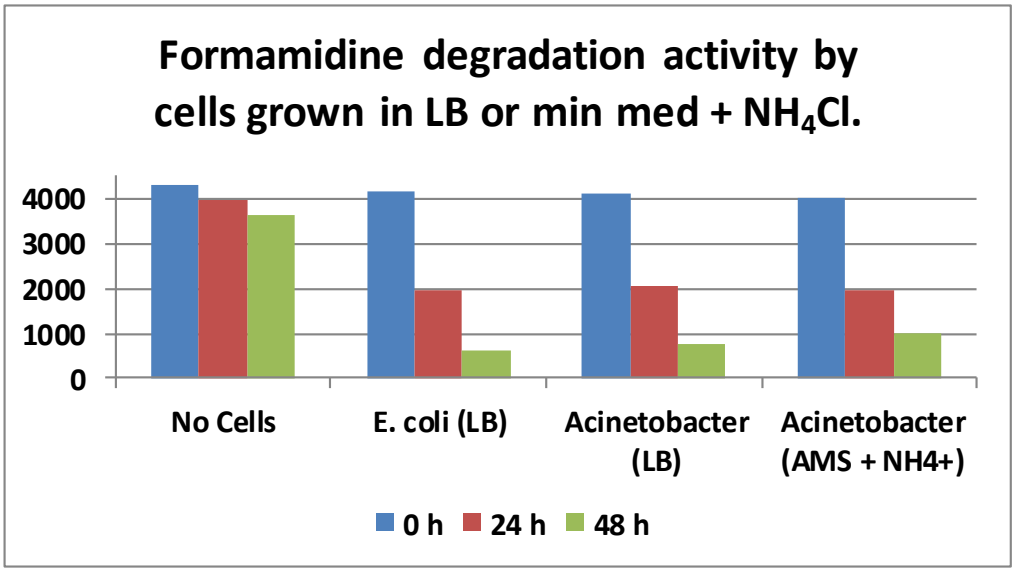


Figure 5-4. Formamidine degradation activity by the cells

To fully take advantage of 3 nitrogens in 1,3,5-triazine, both formamidine and formamide degradation are required by *Acinetobacter sp.* Trz. Hence, resting cell samples containing the each cell and 2 mM formamide or formamidine were analyzed with HPLC to see if formamide is degraded by *Acinetobacter sp.* Trz. The results show that formamidine spontaneouly degrades to formamide at both 25 ºC and 37 ºC. However formamidine degradation rate at 37 ºC was higher than the degradation rate at 25 ºC. Formamide was stable enough at both 25 ºC and 37 ºC to track formamide

degradation by both bacteria. The 2mM formamidine and formamide solutions incubated with the *Acinetobacter* only produced formate that means the *Acinetobacter* consumes formamidine for growth through formamidine to formamide conversion. On the other hand formamide may not be a good nitrogen source for *E. coli* MG1655 because formamide degradation in 2mM formamide plus *E. coli* MG1655 sample was not significant compared to the control sample, 2mM formamide solution without *E.coli*. However, formamidine to formamide conversion was still observed in 2mM formamidine sample with *E. coli* MG1655 which suggests that an *E. coli* MG1655 culture can accelerate the conversion of formamidine to formamide (Table 5-1).

	FAD peak area	FA peak area	FT peak area	OD at 600nm
2mM FAD at 25 ºC	6997.8	358.1	Not detectable	0.0005
2mM FAD + *Aci*	6403.7	Not detectable	39.5	0.4931
2mM FA at 25 ºC	--	1259.2	Not detectable	0.0019
2mM FA + *Aci*	--	Not detectable	50.5	0.4808
2mM FAD at 37 ºC	4933.6	709.7	Not detectable	0.0001
2mM FAD + *E.coli*	3793.7	904.9	Not detectable	0.2564
2mM FA at 37 ºC	Not detectable	1233.1	Not detectable	0.0019
2mM FA + *E.coli*	Not detectable	1177.3	Not detectable	0.0423

Table 5-1. HPLC result for formamidine (FAD) and formamide (FA) degradation analysis

The ^{1}H-NMR samples that contain *Acinetobacter sp.* Trz and 10 mM formamidine were incubated at different time points (time 0, day 3 incubation) together with the control without *Acinetobacter sp.* Trz. In a ^{1}H-NMR experiment, only the sample containing *Acinetobacter sp.* Trz showed formamidine peak decrease in concert with increasing formamide peak. Hence, the ^{1}H-NMR results support the HPLC results

indicating formamidine conversion to formamide is mediated by *Acinetobacter sp.* Trz.

5.3.5 Testing strain Acinetobacter sp. Trz *crude cell extracts for formamidine hydrolase activity*

Formamidine hydrolysis was found to be accelerated in the presence of cells. However, it was unknown if this activity was due to an enzymatic process. Degradation rates by cells grown on LB or minimal medium with ammonium chloride were similar, indicating no induction of a formamidine hydrolase activity by growth on rich or minimal medium. To test if formamidine hydrolase activity could be measured, crude cell extracts were prepared from cells grown on LB or AMS medium with ammonium chloride, added to 10 mM formamidine solutions, incubated, and then analyzed by HPLC. The results showed no apparent change in formamidine peak area after 3 days of incubation. The extract did introduce potentially interfering peaks, and more definitive results could possibly be obtained via NMR.

5.3.6 Characteristics of the isolate growth

The Table 5-2 shows the O.D values at 600 nm of *Acinetobacter sp.* Trz O.D in AMS medium supplemented with different N sources for 2 days. The growth test result shows that acetamide, guanine, and urea are better N sources than NH_4Cl (Table 5-2)

Nitrogen sources	O.D at 600 nm
No N	0.050±0.009
NH_4Cl	1.078±0.014
Acetamide	1.105±0.005
Acetylurea	0.027±0.001
Biuret	0.272±0.002
2-Cyanoacetamide	0.059±0.001
Dicyandiamide	0.023±0.001
Formylurea	0.798±0.025
Guanidine	1.950±0.009
Oxamide	0.045±0.001
Semicarbazide	0.018±0.001
Sulfamide	0.100±0.013
Urea	1.332±0.110

Table 5-2. The O.D values at 600 nm from *Acinetobacter sp.* Trz culture with different N sources

5.5 Discussion

Acinetobacter sp. Trz was isolated in the minimal medium containing 1,3,5-triazine as sole nitrogen source. Although 1,3,5-triazine was assumed to be stable due to its stable ring structure, UV spectroscopy results showed that 1,3,5-triazine is degraded to formamidine, formamide, and formate abiotically after the intermediates first forms. *Acinetobacter sp.* Trz could use both formamidine and formamide as sole nitrogen sources for growth. This result is supported by the HPLC and H-NMR results indicating that *Acinetobacter sp.* Trz degrades formamidine to formate via formamide formation to get NH_3 in the pathway. Those results propose the degradation pathway of 1,3,5-triazine via abiotic and biotic pathways (Figure 5-5).

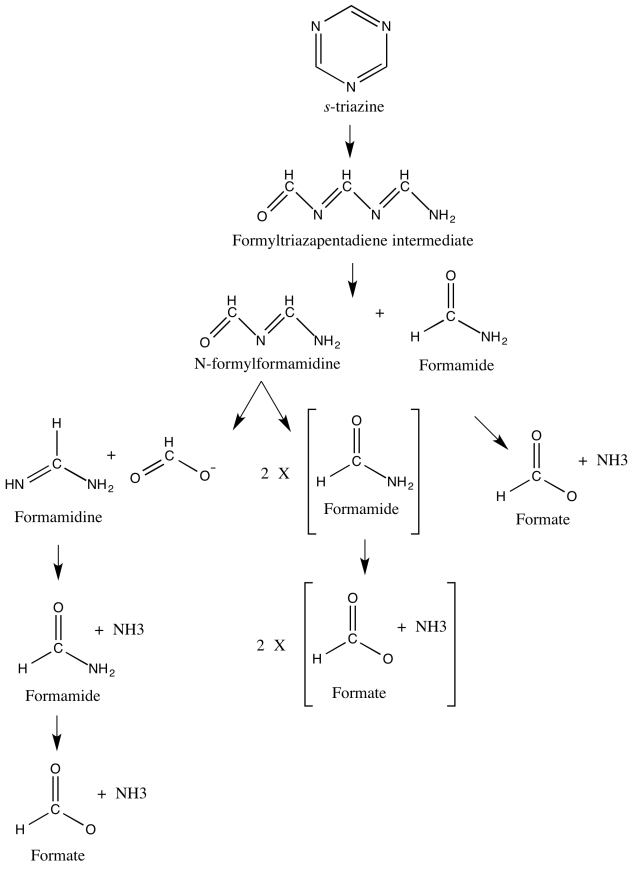


Figure 5-5. Degradation pathway of 1,3,5-triazine via abiotic and biotic pathways

To determine the formamidine degradation is enzymatic, cell extract samples collected at different time points such as time 0, 1 day and 3 days containing 10 mM formamidine were analyzed with HPLC and H-NMR with controls. Results from HPLC and H-NMR did not show significant decrease of the formamidine peak. Therefore currently it is difficult to suggest formamidine degradation is enzymatic.

REFERENCES CITED

1. Pavlova, M. (2009) Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate, *Nature Chemical Biology 5(10)*, 727-733.
2. Belluck, D.A., Benjamin, S.L., and Dawson, T. (1991) Groundwater contamination by atrazine and its metabolites, *Pesticide Transformation Products, ACS Symposium Series Vol 459*, 254-273.
3. Seiler, A., Brenneisen, P., and Green, D. (1992) Benefits and risks of plant protection products-possibilities of protecting drinking water: case atrazine, *Water Supply*.
4. Siripattanakul, S., Wirojanagud, W., McEvoy, J., Limpiyakorn, T., and Khan, E. (2009) Atrazine degradation by stable mixed cultures enriched from agricultural soil and their characterization, *Journal of Applied Microbiology 106(3)*, 986-992.
5. Ackerman, F. (2007) The economics of atrazine, *International Journal of Occupational and Environmental Health 13(4)*, 437-445.
6. Swanton, C. J., Gulden, R. H., and Chandler, K. (2007) A rationale for atrazine stewardship in corn, *Weed Science 55(1)*, 75-81.
7. Mitchell, P. D. (2014) Market‐level assessment of the economic benefits of atrazine in the United States, *Pest Management Science 70(11)*, 1684-1696.
8. Krämer, W., and Schirmer, U. (2007) Modern crop protection compounds*, Wiley.*
9. Appleby, A. P., Müller, F., and Carpy, S. (2000) Weed control, *Ullmann's Encyclopedia of Industrial Chemistry, (Wiley-VCH Verlag GmbH & Co. KGaA)*.
10. Anderson, P. C., and Georgeson, M. (1989) Herbicide-tolerant mutants of corn, *Genome 31(2)*, 994-999.
11. Krueger, R. F., and Seiber, J. N. (1985) Treatment and disposal of pesticide wastes, *Royal Society of Chemisty, London, England.*
12. McPheron, T. (2007) Melamine and cyanuric acid interaction may play part in illness and death from recalled pet food, *Cell 773*, 494-5419.
13. Kello, D. (1989) WHO drinking water quality guidelines for selected herbicides, *Food Additives & Contaminants 6(S1)*, S79-S85.
14. Mandelbaum, R. T., Allan, D. L., and Wackett, L. P. (1995) Isolation and characterization of a Pseudomonas sp. that mineralizes the s-triazine herbicide atrazine, *Applied and Environmental Microbiology 61(4)*, 1451-1457.
15. Thurman, E. M., and Meyer, M. T. (1996) Herbicide metabolites in surface water and groundwater: Introduction and overview, *ACS Symposium Series (American Chemical Society) Vol 630*, 1-15.
16. Gavrilescu, M. (2005) Fate of pesticides in the environment and its bioremediation, *Engineering in Life Sciences 5(6)*, 497-526.

17. Wang, J. (2011) Isolation and characterization of an *Arthrobacter sp*. strain HB-5 that transforms atrazine, *Environmental Geochemistry and Health 33(3)*, 259-266.
18. Frank, R., Sirons, G., and Anderson, G. (1983) Atrazine: the impact of persistent residues in soil on susceptible crop species, *Canadian Journal of Soil Science 63(2)*, 315-325.
19. Sheets, T. (1970) Persistence of triazine herbicides in soils, single pesticide Volume: The Triazine Herbicides*, Springer*, 287-310.
20. Sheets, T., and Harris, C. (1965) Herbicide residues in soils and their phytotoxicities to crops grown in rotations, *Residue Reviews/Rückstandsberichte (Springer)*, 119-140.
21. Cook, A. M., and Huetter, R. (1981) s-Triazines as nitrogen sources for bacteria, *Journal of Agricultural and Food Chemistry 29(6)*, 1135-1143.
22. Cook, A. M., and Hütter, R. (1984) Deethylsimazine: bacterial dechlorination, deamination, and complete degradation, *Journal of Agricultural and Food Chemistry 32(3)*, 581-585.
23. Grossenbacher, H., Cook, A. M., and Hütter, R. (1984) Bacterial degradation of N-cyclopropylmelamine: the steps to ring cleavage, *Biochemical Journal 222(2)*, 315-320
24. Wackett, L., Sadowsky, M., Martinez, B., and Shapir, N. (2002) Biodegradation of atrazine and related s-triazine compounds: from enzymes to field studies, *Applied Microbiology and Biotechnology 58(1)*, 39-45.
25. Radosevich, M., Traina, S. J., Hao, Y. L., and Tuovinen, O. H. (1995) Degradation and mineralization of atrazine by a soil bacterial isolate, *Applied and Environmental Microbiology 61(1)*, 297-302.
26. Struthers, J., Jayachandran, K., and Moorman, T. (1998) Biodegradation of atrazine by *Agrobacterium radiobacter* J14a and use of this strain in bioremediation of contaminated soil, *Applied and Environmental Microbiology 64(9)*, 3368-3375.
27. Topp, E., Mulbry, W. M., Zhu, H., Nour, S. M., and Cuppels, D. (2000) Characterization of s-triazine herbicide metabolism by a *Nocardioides* sp. isolated from agricultural soils, *Applied and Environmental Microbiology 66(8)*, 3134-3141.
28. Topp, E. (2000) Characterization of an atrazine-degrading *pseudaminobacter* sp. isolated from Canadian and French agricultural soils, *Applied and Environmental Microbiology 66(7)*, 2773-2782.
29. Behki, R. M., amd Khan, S. U. (1986) Degradation of atrazine by *Pseudomonas*: N-dealkylation and dehalogenation of atrazine and its metabolites, *Journal of Agricultural and Food Chemistry 34(4)*, 746-749.
30. Assaf, N. A., and Turco, R. F. (1994) Accelerated biodegradation of atrazine by a microbial consortium is possible in culture and soil, *Biodegradation 5(1)*, 29-35.

31. Boundy-Mills, K. L., De Souza, M., Mandelbaum, R. T., Wackett, L. P., and Sadowsky, M. J. (1997) The atzB gene of *Pseudomonas* sp. strain ADP encodes the second enzyme of a novel atrazine degradation pathway, *Applied and Environmental Microbiology 63(3)*, 916-923.
32. de Souza, M., Wackett, L. P., Boundy-Mills, K. L., Mandelbaum, R. T., and Sadowsky, M. J. (1995) Cloning, characterization, and expression of a gene region from *Pseudomonas* sp. strain ADP involved in the dechlorination of atrazine, *Applied and Environmental Microbiology 61(9)*, 3373-3378.
33. Moscinski, J., Jayachandran, K., and Moorman, T. (1996) Mineralization of the herbicide atrazine by *Agrobacterium radiobacter*, abstr. Q-414, *96th General Meeting of the American Society for Microbiology*, 458.
34. Wang, J., Zhu, L., Wang, Q., Wang, J., and Xie, H. (2014) Isolation and characterization of atrazine mineralizing *Bacillus subtilis* strain HB-6, *PLoS One 9(9)*, e107270
35. Martinez, B., Tomkins, J., Wackett, L. P., Wing, R., and Sadowsky, M. J. (2001) Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonassp*. strain ADP, *Journal of Bacteriology 183(19)*, 5684-5697.
36. Sadowsky, M. J., and Wackett, L. P. (2001) Genetics of atrazine and s-triazine degradation by *Psedomonas* sp. strain ADP and other bacteria, *ACS Symposium Series (American Chemical Society) Chapter 15*, 268-282
37. de Souza, M. L., Sadowsky, M. J., and Wackett, L. P. (1996) Atrazine chlorohydrolase from *Pseudomonas* sp. strain ADP: gene sequence, enzyme purification, and protein characterization, *Journal of Bacteriology 178(16)*, 4894-4900.
38. Sadowsky, M.J., Tong, Z., de Souza, M., and Wackett, L. P. (1998) AtzC is a new member of the amidohydrolase protein superfamily and is homologous to other atrazine-metabolizing enzymes, *Journal of Bacteriology 180(1)*, 152-158.
39. Sajjaphan, K. (2004) *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 70(7)*, 4402-4407.
40. Shapir, N. (2006) TrzN from *Arthrobacter aurescens* TC1 is a zinc amidohydrolase, *Journal of Bacteriology 188(16)*, 5859-5864.
41. Seffernick, J. L., Reynold, R., Fedorov, A. A., Fedorov, E., Almo, S. C., Sadowsky, M. J., and Wackett, L. P. (2010) X-ray structure and mutational analysis of the atrazine chlorohydrolase TrzN, *Journal of Biological Chemistry 285(40)*, 30606-30614.
42. Peat, T. S., Newman, J., Balotra, S., Lucent, D., Warden, A. C., and Scott, C. (2015) The structure of the hexameric atrazine chlorohydrolase AtzA, *Acta Crystallographica Section D: Biological Crystallography 71(3)*, 710-720.
43. Dodge, A. G., Wackett, L. P., and Sadowsky, M. J. (2012) Plasmid localization and organization of melamine degradation genes in *Rhodococcus* sp. strain Mel, *Applied and Environmental Microbiology 78(5)*, 1397-1403.
44. Balotra, S., Newman, J., Cowieson, N. P., French, N. G., Campbell, P. M., Briggs, L. J., and Scott, C. (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(2)*, 470-480.
45. García-González, V., Govantes, F., Porrúa, O., and Santero, E. (2005) Regulation of the *Pseudomonas* sp. strain ADP cyanuric acid degradation operon, *Journal of Bacteriology 187(1)*, 155-167.
46. Fruchey, I., Shapir, N., Sadowsky, M. J., and Wackett, L. P. (2003) On the origins of cyanuric acid hydrolase: purification, substrates, and prevalence of AtzD from *Pseudomonas* sp. strain ADP, *Applied and Environmental Microbiology 69(6)*, 3653-3657.
47. Karns, J. S. (1999) Gene sequence and properties of an s-triazine ring-cleavage enzyme from *Pseudomonas* sp. strain NRRLB-12227, *Applied and Environmental Microbiology 65(8)*, 3512-3517.
48. Copley, S. D. (2009) Prediction of function in protein superfamilies, *F1000 Biology Reports 1*, 91
49. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies, *Current Opinion in Chemical Biology 10(5)*, 492-497.
50. Pegg, S., Brown, S., Ojha, S., Huang, C. C., Ferrin, T. E., and Babbit, P. C. (2005) Representing structure–function relationships in mechanistically diverse enzyme superfamilies, *Pacific Symposium on Biocomputing (World Scientific) Vol35*, 3588-369.
51. Seffernick, J. L., Erickson, J. S., Cameron, S. M., Cho, S., Dodge, A. G., Richman, J. E., Sadowsky, M. J., and Wackett, L. P. (2012) Defining sequence space and reaction products within the cyanuric acid hydrolase (AtzD)/barbiturase protein family. *Journal of Bacteriology 194(17)*, 4579-4588.
52. Stamper, D.M., Krzycki, J. A., Nicomrat, D., and Tuovinen, O. H. (2005) Ring-cleaving cyanuric acid amidohydrolase activity in the atrazine mineralizing *Ralstonia basilensis* M91-3, *Biocatalysis and Biotransformation 23(6)*, 387-396.
53. 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, ureidomalonase, *Biochemical and Biophysical Research Communications 286(1)*, 222-226.
54. Cho, S., Shi, K., Seffernick, J. L., Dodge, A. G., Wackett, L. P., and Aihara, H. (2014) Cyanuric acid hydrolase from *Azorhizobium caulinodans* ORS 571: crystal structure and insights into a new class of Ser-Lys dyad proteins, *PLoS One 9(6)*, e99349.
55. Peat, T. S., Balotra, S., Wilding, M., French, N. G., Briggs, L. J., Panjikar, S., Cowieson, N., Newman, J., and Scott, C. (2013) Cyanuric acid hydrolase: evolutionary innovation by structural concatenation, *Molecular Microbiology 88(6)*, 1149-1163.
56. Huthmacher, K., and Most, D. (2005) Cyanuric acid and cyanuric chloride, *Ullmann's Encyclopedia of Industrial Chemistry*.

57. Canelli, E. (1974) Chemical, bacteriological, and toxicological properties of cyanuric acid and chlorinated isocyanurates as applied to swimming pool disinfection: a review, *American Journal of Public Health 64(2)*, 155-162.
58. Yeom, S., Mutlu, B. R., Aksan, A., and Wackett, L. P. (2015) Bacterial cyanuric acid hydrolases for water treatment. *Applied and Environmental Microbiology*, 02175-02115.
59. Dobson, R. L., Motlagh, S., Quijano, M., Cambron, R. T., Baker, T. R., Pullen, A. M., Regg, B. T., and Daston, G. P. (2008) Identification and characterization of toxicity of contaminants in pet food leading to an outbreak of renal toxicity in cats and dogs, *Toxicological Sciences 106(1)*, 251-262.
60. Jutzi, K., Cook, A. M., and Hutter, R. (1982) The degradative pathway of the s-triazine melamine. The steps to ring cleavage, *Biochemical Journal 208*, 679-684.
61. Cheng, G., Shapir, N., Sadowsky, M. J., and Wackett, L. P. (2005) Allophanate hydrolase, not urease, functions in bacterial cyanuric acid metabolism, *Applied and Environmental Microbiology 71(8)*, 4437-4445.
62. Cameron, S.M., Durchschein, K., Richman, J. E., Sadowsky, M. J., and Wackett, L. P. (2011) New family of biuret hydrolases involved in s-triazine ring metabolism, *ACS Catalysis 1(9)*, 1075-1082.
63. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S.L., and Clardy, J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin, *Journal of Molecular Biology 229(1)*, 105-124.
64. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods in Enzymology (276)*, 307-326
65. Patton, C. J., and Crouch, S. (1977) Spectrophotometric and kinetics investigation of the Berthelot reaction for the determination of ammonia, *Analytical Chemistry 49(3)*, 464-469.
66. Weatherburn, M. (1967) Phenol-hypochlorite reaction for determination of ammonia, *Analytical Chemistry 39(8)*, 971-974.
67. Vagin, A., and Teplyakov, A. (2009) Molecular replacement with MOLREP, *Acta Crystallographica Section D: Biological Crystallography 66(1)*, 22-25.
68. Huthmacher, K., and Most, D. (2000) Cyanuric acid and cyanuric chloride, *Ullmann's Encyclopedia of Industrial Chemistry, (Wiley-VCH Verlag GmbH & Co. KGaA)*.
69. Dodge, A. G., Preiner, C. S., and Wackett, L. P. (2013) Expanding the cyanuric acid hydrolase protein family to the fungal kingdom, *Journal of Bacteriology 195(23)*, 5233-5241.
70. Cho, S., Shi, K., Wackett, L. P., and Aihara, H. (2013) Crystallization and preliminary X-ray diffraction studies of cyanuric acid hydrolase from *Azorhizobium caulinodans*, *Acta Crystallographica Section F: Structural Biology and Crystallization Communications 69(8)*, 880-883.
71. Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension, *Acta Crystallographica Section D: Biological Crystallography 59(1)*, 38-44.

72. Emsley, P and Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallographica Section D: Biological Crystallography 60(12)*, 2126-2132.
73. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *Journal of Applied Crystallography 26(2)*, 283-291.
74. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis, *Journal of Computational Chemistry 25(13)*, 1605-1612.
75. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions, *Protein Engineering 8(2)*, 127-134.
76. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: estimation of static accessibility, *Journal of Molecular Biology 55(3)*, 379-IN374.
77. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallographica Section D: Biological Crystallography 67(4)*, 235-242.
78. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallographica Section D: Biological Crystallography 54(5)*, 905-921.
79. Oelschlaeger, P., Mayo, S. L., and Pleiss, J. (2005) Impact of remote mutations on metallo‐β‐lactamase substrate specificity: Implications for the evolution of antibiotic resistance, *Protein Science 14(3)*, 765-774.
80. Saen-Oon, S., Ghanem, M., Schramm, V. L., and Schwartz, S. D. (2008) Remote mutations and active site dynamics correlate with catalytic properties of purine nucleoside phosphorylase, *Biophysical Journal 94(10)*, 4078-4088.
81. Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D, *Nucleic Acids Research 38(suppl 2)*, W545-W549.
82. Andreeva, A., Howorth, D., Chandonia, J. M., Brenner, S. E., Hubbard, T. J., Chothia, C., and Murzin, A. G. (2008) Data growth and its impact on the SCOP database: new developments, *Nucleic Acids Research 36(suppl 1)*, D419-D425.
83. Burman, J. D., Stevenson, C. E., 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 Structural Biology 7(1)*, 30.
84. Burgi, H., Dunitz, J., and Shefter, E. (1973) Geometrical reaction coordinates. II. Nucleophilic addition to a carbonyl group, *Journal of the American Chemical Society 95(15)*, 5065-5067.

85. Ekici, Ö. D., Paetzel, M., and Dalbey, R. E. (2008) Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration, *Protein Science 17(12)*, 2023-2037.
86. Matthews, B., Sigler, P., Henderson, R., and Blow, D. (1967) Three-dimensional structure of tosyl-α-chymotrypsin, *Nature 214*, 652-656.
87. Wright, C. S., Alden, R. A., and Kraut, J. (1969) Structure of subtilisin BPN'at 2.5 Å resolution, *Nature 221*, 235-242.
88. Henderson, R. (1970) Structure of crystalline α-chymotrypsin: IV. The structure of indoleacryloyl-α-chymotrypsin and its relevance to the hydrolytic mechanism of the enzyme, *Journal of Molecular Biology 54(2)*, 341-354.
89. Watson, H., Shotton, D., Cox, J. M., and Muirhead, H. (1970) Three-dimensional Fourier synthesis of tosyl-elastase at 3.5 Å resolution, *Nature 225*, 806-811.
90. Paetzel, M., Dalbey, R. E., and Strynadka, N. C. (1998) Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor, *Nature 396*, 186-190.
91. Lin, L. L., and Little, J. W. (1989) Autodigestion and RecA-dependent cleavage of Ind− mutant LexA proteins, *Journal of Molecular Biology 210(3)*, 439-452.
92. Dalbey, R. E., Lively, M. O., Bron, S., and van Dijl, J. M. (1997) The chemistry and enzymology of the type I signal peptidases, *Protein Science : A Publication of the Protein Society 6(6)*, 1129-1138.
93. Strynadka, N. C., Adachi, H., Jensen, S. E., Johns, K., Sielecki, A., Betzel, C., Kazuo, S., and James, M. N. (1992) Molecular structure of the acyl-enzyme intermediate in β-lactam hydrolysis at 1.7 Å resolution, *Nature 359*, 700-705.
94. Li, Q., Seffernick, J. L., Sadowsky, M. J., and Wackett, L. P. (2009) Thermostable cyanuric acid hydrolase from *Moorella thermoacetica* ATCC 39073, *Applied and Environmental Microbiology 75(22)*, 6986-6991.
95. Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I. A., Lesley, S. A., and Godzik, A. (2007) XtalPred: a web server for prediction of protein crystallizability, *Bioinformatics 23(24)*, 3403-3405.
96. Goldschmidt, L., Cooper, D. R., Derewenda, Z. S., and Eisenberg, D. (2007) Toward rational protein crystallization: a web server for the design of crystallizable protein variants, *Protein Science 16(8)*, 1569-1576.
97. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallographica Section D: Biological Crystallography 66(2)*, 213-221.
98. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *Journal of Applied Crystallography 40(4)*, 658-674.
99. Ramachandran, G., Ramakrishnan, C., and Sasisekharan, V. (1963) Stereochemistry of polypeptide chain configurations, *Journal of Molecular Biology 7(1)*, 95-99.

100. Alber, T., Petsko, G. A., and Tsernoglou, D. (1976) Crystal structure of elastase-substrate complex at -- 55 degrees C, *Nature 263*, 297-300.
101. Ding, X., Rasmussen, B. F., Petsko, G. A., and Ringe, D. (1994) Direct structural observation of an acyl-enzyme intermediate in the hydrolysis of an ester substrate by elastase, *Biochemistry 33(31)*, 9285-9293.
102. Wilmouth, R. C., Westwood, N. J., Anderson, K., Brownlee, W., Claridge, T. D., Clifton, I. J., Pritchard, G. J., Aplin, R. T., and Schofield, C. J. (1998) Inhibition of elastase by N-sulfonylaryl β-lactams: Anatomy of a stable acyl-enzyme complex, *Biochemistry 37(50)*,17506-17513.
103. Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R., and Schofield, C. J. (2002) X-ray structure of a serine protease acyl-enzyme complex at 0.95-Å resolution, *Journal of Biological Chemistry 277(24)*, 21962-21970.
104. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) Structure of a specific acyl-enzyme complex formed between β-casomorphin-7 and porcine pancreatic elastase, *Nature Structural and Molecular Biology 4(6)*, 456-462.
105. Chung, I. Y. W., and Paetzel, M. (2011) Crystal structure of a viral protease intramolecular acyl-enzyme complex: insight into cis-cleavage at the VP4/VP3 junction of Tellina birna, *Journal of Biological Chemistry 286(14)*, 12475-12482.
106. Vassiliadis, P., Kalapothaki, V., Mavrommati, C., and Trichopoulos, D. (1984) A comparison of the original Rappaport medium (R medium) and the Rappaport-Vassiliadis medium (RV medium) in the isolation of salmonellae from meat products, *The Journal of Hygiene 93(1)*, 51-58.
107. Parales, R. E., Adamus, J. E., White, N., and May, H. D. (1994) Degradation of 1,4-dioxane by an actinomycete in pure culture, *Applied and Environmental Microbiology 60(12)*, 4527-4530.
108. Lane, D. (1991) 16S/23S rRNA sequencing, *Nucleic Acid Techniques in Bacterial Systematics*, 125-175.
109. Sly, L. I., Cox, T. L., and Beckenham, T. B. (1999) The phylogenetic relationships of *Caulobacter*, *Asticcacaulis* and *Brevundimonas* species and their taxonomic implications, *International Journal of Systematic Bacteriology 49(2)*, 483-488.
110. Cole, J. R., Wang, Q., Fish, J. A., Chai, B., McGarrell, D. M., Sun, Y., Brown, T., Porras-Alfaro, A., Kuske, C. R., and Tiedje, J. M. (2013) Ribosomal Database Project: data and tools for high throughput rRNA analysis, *Nucleic Acids Research*, gkt1244.
111. Bruno, W. J., Socci, N. D., and Halpern, A. L. (2000) Weighted neighbor joining: a likelihood-based approach to distance-based phylogeny reconstruction, *Molecular Biology and Evolution 17(1)*, 189-197.
112. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Analytical Biochemistry 72(1)*, 248-254.