MECHANISTIC INSIGHTS INTO THE DIVERGED ENZYMES OF THE AMIDOHYDROLASE SUPERFAMILY

A Dissertation

by

TINH T. NGUYEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

December 2009

Major Subject: Chemistry

MECHANISTIC INSIGHTS INTO THE DIVERGED ENZYMES OF THE AMIDOHYDROLASE SUPERFAMILY

A Dissertation

by

TINH T. NGUYEN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Frank M. Raushel
Committee Members,	Paul A. Lindahl
	Gregory D. Reinhart
	Coran M H. Watanabe
Head of Department,	David H. Russell

December 2009

Major Subject: Chemistry

ABSTRACT

Mechanistic Insights into the Diverged Enzymes of the Amidohydrolase Superfamily.

(December 2009)

Tinh T. Nguyen, B.S., Mary Washington College

Chair of Advisory Committee: Dr. Frank M. Raushel

The amidohydrolase superfamily is a functionally diverse set of enzymes that catalyzes predominantly hydrolysis reactions involving sugars, nucleic acids, amino acids, and organophosphate esters. A more divergent member of this superfamily, URI (uronate isomerase) from *Escherichia coli*, catalyzes the isomerization of D-glucuronate to D-fructuronate and D-galacturonate to D-tagaturonate. In *Bacillus halodurans*, two distinct operons were identified for the metabolism of D-glucuronate and D-galacturonate based on kinetics and genomic context. The canonical uronate isomerase is encoded by the gene Bh0705. A second URI in this organism, Bh0493, is the outlier of the group in terms of sequence similarity. Kinetic evidences indicate that Bh0705 is relatively specific for the isomerization of D-glucuronate, while Bh0493 is specific for the D-galacturonate pathway.

Bell-shaped pH-rate profiles were observed for the wild type URI from *Escherichia coli*. Primary isotope effects with [2-^{2}H]-D-glucuronate and solvent viscosity studies are consistent with product release as the rate limiting step. X-ray structure of Bh0493 was determined in the presence of D-glucuronate. A chemical

mechanism is proposed that utilizes a proton transfer from C-2 of D-glucuronate to C-1 that is initiated by the combined actions of Asp-355 and the C-5 hydroxyl of the substrate that is bound to the metal ion. The formation of the *cis*-enediol intermediate is further facilitated by the shuttling of the proton between the C-2 and C-1 oxygens by the conserved Tyr-50 and/or Arg-357.

Another divergent member of the AHS is the enzyme renal dipeptidase. Structural studies of the enzyme from *Streptomyces coelicolor* (Sco3058) demonstrate that the active site consists of a binuclear metal center. Bell-shaped pH-rate profiles are observed for both Zn^{2+} and Cd^{2+} enzymes. A chemical mechanism for renal dipeptidase is proposed based on structural analysis of the enzyme-inhibitor complex. The reaction is initiated by the polarization of the amide bond by the β-metal. Asp-320 activates the bridging hydroxide for nucleophilic attack at the peptide carbon center, forming a tetrahedral intermediate that is stabilized by the metal center and His-150. The protonated Asp-320 donates the proton to the α-amino group of the leaving group, causing the collapse of the tetrahedral intermediate and cleavage of the carbon-nitrogen bond.

DEDICATION

To the memory of Suhyung Park

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Frank Raushel, for his guidance and support throughout the course of this research. I also would like to thank my committee members: Dr. Paul Lindahl, Dr. Gregory Reinhart, and Dr. Coran Watanabe. I would like to extend my appreciation to Dr. David Barondeau for his generous time and effort in helping me with the crystallography of Sco3058.

Thanks to my family for their love and faith in me. Thanks to my best friends Thuy and Tuyet, who hold the key to my sanity. I also want to extend my gratitude to the past and present members of the Raushel group for their support and companionship. Thanks to the members of the Barondeau group for all of their help.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

FIGURE Page

LIST OF TABLES

CHAPTER I

INTRODUCTION

The increasing number of the available genome sequences, together with advances in the computational methods that reveals the relationships between evolutionarily related genes, have generate a greater understanding into the evolution of proteins, their functions, and their associated biological processes (*1*). However, one of the difficult and ongoing tasks is the identification and assignment of molecular activity to functionally unknown enzymes encoded within completely sequenced genomes. Many sequenced genes in both prokaryotes and eukaryotes have functions that are uncertain, unknown, or incorrectly annotated (*2*-*6*). The task of protein function discovery is facilitated by the emerging number of high-resolution protein structures. A complete understanding of the interplay of structure and function is essential for decoding substrate specificity for these proteins of unknown function and for the development of catalysts for new reactions using the active sites of preexisting templates of a unique protein fold (*7*).

The practice of incorporating genomic context (sequence families, structures, and functions) into the study of enzyme reaction and mechanism is referred to as genomic enzymology (*8*). One of the driving forces for the development of genomic enzymology

---

This dissertation follows the style of *Biochemistry*.

was the recognition that enzymes that are evolutionary diverged can catalyze different overall reactions. Enzymes that are evolutionarily related (homologous enzymes) are described in several ways. A family refers to a group of homologous enzymes that catalyze the same reaction in terms of mechanism and substrate specificity (*8*). A superfamily is a group of enzymes that catalyze either the same chemical reaction with differing substrate specificities or different overall reactions that share a common mechanistic feature such as a partial reaction, intermediate, or transition state. This mechanistic attribute is facilitated by conserved active site residues that play similar roles in catalysis for all members of the superfamily (*8*). A suprafamily refers to a group of homologous enzymes that utilize different mechanistic elements to catalyze different overall reactions. Classification of enzymes and consequently functional annotations for members of mechanistically diverse superfamilies and functionally distinct suprafamilies are not possible from just sequence data, rather it entails the integration of functional and structural characterization (*8*). The number of structural types is much less than the set of reactions that needs to be deciphered due to the limited number of protein domain folds (*9-10*).

The number of unique protein folds discovered to date is estimated to be far less than the number of proteins encoded by the human genome (*11-13*). Consequently, protein folds must be reused by divergent evolution or independently formulated by convergent evolution to accommodate the number of functions represented by proteins found in living organisms. Based on the available evidence, it is possible to picture three distinct strategies that nature uses for the divergent evolution of enzyme function. Each

of these strategies involves the initial copying of the gene encoding the protein to be evolved so that the original enzyme activity is preserved by the organism (*8*).

The first divergent evolution strategy involves enzymes that catalyze reactions in a biosynthetic pathway (*14-15*). When a substrate for an enzyme is depleted, a new enzyme is evolved to supply the substrate from an available precursor using the template from an enzyme that uses the substrate. Because the original and the newly evolved enzyme must share the ability to bind the same substrate/product, this evolution strategy is restricted to retain binding specificity but not reaction mechanisms. Thus the old and new enzymes would be members of functionally distinct suprafamilies. In contrast to this approach is the hypothesis that the protein for divergent evolution come from a pool of enzymes whose mechanisms provide the needed partial reaction or strategy for stabilization of reaction intermediates or transition states (*16-17*). Evolution alters certain mechanistic characteristics and/or substrate specificity so that the new reaction is catalyzed with enhanced proficiency. Perhaps the selected enzyme to be evolved is promiscuous with regard to the desired new reaction even at a very low rate, and subsequent evolution then provides enhanced proficiency as the new metabolic pathway is optimized (*18-20*). This strategy of evolution creates enzymes that are members of mechanistically diverse superfamilies. The third scheme of divergent evolution suggests that the active site architecture is dominant instead of substrate specificity or chemical mechanism. For this approach, an active site is capable of utilizing shared functional groups in different mechanistic and metabolic contexts to catalyze an alternate reaction.

Thus, the old and the newly evolved enzymes are members of functionally distinct suprafamilies (*8*).

The most frequently found protein fold in nature is the $(\beta/\alpha)_8$-barrel. According to SCOP databases, there are currently 33 identified $(\beta/\alpha)_8$-barrel superfamilies including the amidohydrolase superfamily (*8*,*21*), the enolase superfamily (*8*,*22-24*), the thiyl radical superfamily (*8*,*25-26*), and the crotonase superfamily (*8,27-28*). The first crystal structure of an enzyme that posesses this structural fold was triose phosphate isomerase (TIM) from chicken muscle (*29*); hence the $(\beta/\alpha)_8$-barrel is also called the TIM-barrel fold. The canonical TIM-barrel consists of eight β-strands forming the beta barrel that is surrounded by eight α-helices. The active site of the TIM-barrel enzymes generally formed at the C-termini of the β-strands and the flexible loops followed the β-strands. The $(\beta/\alpha)_8$ motif of the triose phosphate isomerase structure is illustrated in **Figure 1.1**. The presence of this $(\beta/\alpha)_8$ structural fold in so many enzymes demonstrates its versatility in the divergence of enzyme function. Recent studies have shown that the $(\beta/\alpha)_8$-barrel enzyme fold originates from the duplication and fusion of ancestral $(\beta/\alpha)_4$-half-barrel, and artificial $(\beta/\alpha)_8$-barrel protein has been constructed experimentally by fusing two copies of the C-terminal half-barrel model protein (*30*).

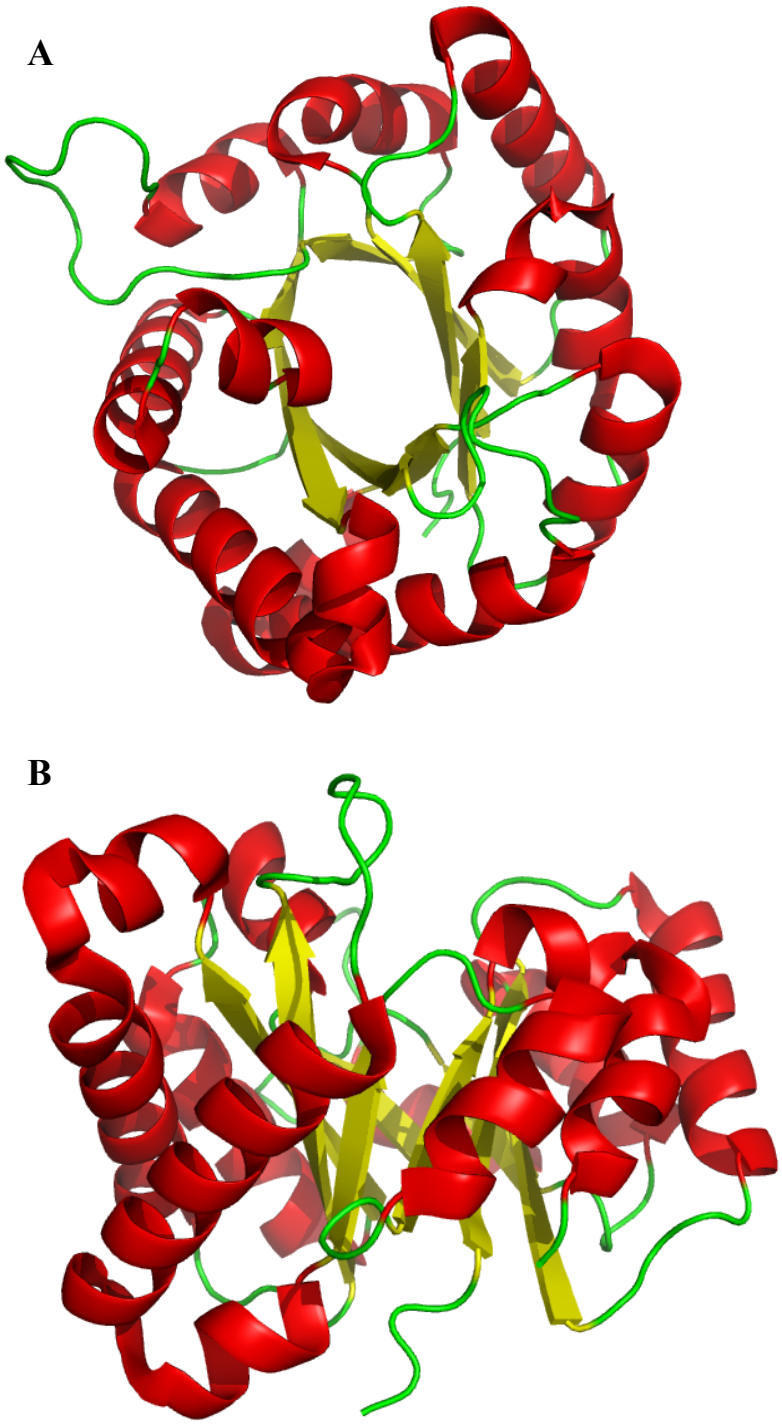


Figure 1.1: Crystal structure of triose phosphate isomerase from chicken muscle. **A)** Top view of TIM. **B)** Side view of TIM. (PDB: 1TIM)

One of the well-documented examples for the divergent evolution of protein architecture and function that can originate from a common ancestral precursor is the amidohydrolase superfamily (AHS) of enzymes. The AHS superfamily was first identified by Holm and Sander in 1997 based on the three-dimentional structural similarities of the protein folds and active sites of urease (URE), phosphotriesterase (PTE), and adenosine deaminase (ADA) (*21*). These enzymes, as mentioned earlier, all contain a $(\beta/\alpha)_8$ (TIM)-barrel structural domain. The most prominent structural landmark for this superfamily of enzymes is a mononuclear or binuclear metal center active site (*21,31*). The metals in a binuclear center are designated as α and β metal site. For the mononuclear enzymes, the metal is located at either the α or the β site. The divalent metal ions that have been found with these enzymes are zinc, cobalt, manganese, iron, and nickel. The metal center in the enzymes of the amidohydrolase superfamily has been shown to be essential for the expression of the overall catalytic activity (*31*).

X-ray crystal structures of several families of the amidohydrolase enzymes can be obtained from the Protein Data Bank (PDB). All of the amidohydrolase structures of the wild type and mutant enzymes (as of June 12, 2009) are shown in **Table 1.1**. From the 45 wild type structures, 29 have known catalytic function. These are urease (*32*) (URE, PDB code 2ubp), phosphotriesterase (*33*) (PTE, PDB code 1hzy), dihydroorotase (*34*) (DHO, PDB code 1j79), β-aspartyl dipeptidase (*35*) (IAD, PDB code 1onw), D and L-hydantoinases (*36-37*) (D-HYD, PDB code 1gkp and L-HYD, PDB code 1gkr), dihydropyrimidinase (*38*) (DHP, PDB code 1k1d), allantoinase (*39*) (ATase, PDB code

2e74), renal dipeptidase (*40*) (RDP, PDB code litq), *N*-acetyl glucosamine-6-phosphate deacetylase (*41-42*) (AGD, PDB code 1o12), uronate isomerase (*43-44*) (URI, PDB code 1j5s), D-amino acid deacetylase (*45*) (DAA, PDB code 1m7j), adenosine deaminase (*46*) (ADA, PDB code 1a4m), cytosine deaminase (*47*) (CDA, PDB code 1k6w), guanine deaminase (GDA, PDB code 2uz9), adenosine 5'-monophosphate deaminase (*48*) (AMPD, PDB code 2ael), 4-oxalomesaconate hydratase (PcmD, PDB code 2gwg), 2 pyrone 4,6-dicarboxylic acid hydratase (PcmC, PDB code 2qah), 2,6-dihydroxybenzoate decarboxylase (Rdc, PDB code 2dvt), amidozolonepropionase (*49*) (HutI, PDB code 2bb0), α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (*50*) (ACMSD, PDB code 2hbv), γ- and δ-lactonase (*4*) (Dr0930, PDB code 3fdk), phosphotriesterase-like-lactonase (*Sso*Pox, PDB code 2vc5) (*51*), two carboxypeptidases (*5-6*) (Cc0300 and Cc3125, PDB codes 2qs8 and 3be7, respectively), enamidase (*52*) (ENA, PDB code 2vun), N-isopropylammedile isopropylaminohydrolase (AtzC, PDB code 2qt3), a S-adenosylhomocysteine deaminase (7) (Tm0936, PDB code 1j6p), and histidinol-phosphate phosphatase from *T. thermophilus* (*53*) (HolPase, PDB code 2z4g). Sixteen of the reported structures remain with unknown biological function, these are the PTE homology protein (*54*) (PHP, PDB code 1bf6), three putative DNases from *Escherichia coli* (TatD, YjjV, and YcfH with PDB codes 1xwy, 1zzm, and 1yix, respectively), two of the three TatD-related protein from *Saccharomyces cerevisiae and Deinococcus radiodurans* (PDB codes 3e2v and 3gg7, respectively), and several uncharacterized amidohydrolase structures (PDB codes 3cjp, 2imr, 2i5g, 3h4u, 2paj, 2ics, 2p9b, 2r8c, and 3feg), solved by the efforts of the New York Structural Genomics Research

Consortium, and a TatD-related protein from *T. maritima* (PDB code 1j6o), solved by the efforts of the Joint Center for Structural Genomics.

Table 1.1: All of the structures belonging to the AHS that have been deposited in the Protein Data Bank (as of June 2009).

Deposition	PDB #	Enzyme	Organism	Active Site Ligands
9/10/08	3EGJ	NagA	*V. cholerae*	1 Ni (β)
4/21/09	3H4U	amidohydrolase	*Unidentified*	1 Zn (α)
11/30/08	3FEQ	eah89906	*Unidentified*	2 Zn
11/25/08	3FDK	unreleased hydrolase	*Unidentified*	Unknown
7/17/08	3DUG	arginine Carboxy-peptidase	*Unidentified*	2 Zn, Arginine
11/16/07	3BE7	arginine carboxy-peptidase	*Unidentified*	Arginine
9/10/07	2R8C	EAJ56179	*Unidentified*	2 Zn
5/21/07	2Q09	imidazolone-propionase	*Unidentified*	1 Fe, 3-(2,5-Dioxo-imidazolidin-4-yl)-propionic acid inhibitor
3/27/07	2PAJ	putative CDA / GuaD	*Unidentified*	1 Zn (α)
1/25/07	2OOF	imidazolone-propionase	*Unidentified*	1 Fe
8/20/01	1GKQ	D-hydantoinase	*Thermus sp.*	2 Zn
8/20/01	1GKP	D-hydantoinase	*Thermus sp.*	2 Zn
6/17/07	2Z4G	histidinol phosphate phosphatase	*T. thermophilus*	2 Fe, 1Zn
5/6/07	2Z00	DHO	*T. thermophilus*	2 Zn
5/3/07	2YZ5	histidinol phosphate phosphatase	*T. thermophilus*	2 Fe, 1Zn, Phosphate

Table 1.1: Continued.

4/26/07	2YXO	histidinol phosphate phosphatase	*T. thermophilus*	2 Fe, 1Zn, Sulfate
4/20/07	2PLM	SAH/TMA deaminase	*T. maritima*	1 Zn , inosylhomocysteine
8/11/07	2ANU	predicted phosphoesterase TM0559	*T. maritima*	4 Zn, 2Cl
4/12/03	1P1M	SAH/TMA deaminase	*T. maritima*	1 Ni, Methionine
10/15/02	1O12	NagA	*T. maritima*	1 Fe
7/9/02	1J6P	SAH/TMA deaminase	*T. maritima*	1 Ni
7/9/02	1J6O	TatD-related deoxy-ribonuclease	*T. maritima*	None
7/2/02	1J5S	uronate isomerase	*T. maritima*	None
9/18/07	2VC7	SsoPox lactonase, PTE	*S. solfataricus*	1 Fe, 1Co, (4S)-4-(decanoylamino)-5-hydroxy-3,4-dihydro-2H-thiophenium
9/18/07	2VC5	SsoPox lactonase, PTE	*S. solfataricus*	1 Fe, 1Co
6/15/07	2QAH	2-pyrone-4,6-dicarboxylic acid hydrolase	*S. paucimobilis*	None
6/3/08	3DC8	dihydro-pyrimidinase	*S. meliloti*	2 Zn
1/31/06	2FVM	dihydro-pyriminidase	*S. kluyveri*	2 Zn, N-carbamoyl-β-alanine
1/31/06	2FVK	dihydro-pyriminidase	*S. kluyveri*	2 Zn, Dihydrouracil
1/25/06	2FTY	dihydro-pyriminidase	*S. kluyveri*	2 Zn
8/6/08	3E2V	TatD like	*S. cerevisiae*	1 Mg
3/25/09	3GRI	DHO	*S. Aureus*	1 Zn (α)

Table 1.1: Continued.

5/12/06	2GZX	TatD deoxy-ribonuclease MW0446	*S. aureus*	2 Ni
8/1/06	2DVX	2,6-dihydroxy-benzoate decarboxylase	*Rhizobium sp.*	1 Zn, 2,3-dihydroxybenzaldehyde
8/1/06	2DVU	2,6-dihydroxy-benzoate decarboxylase	*Rhizobium sp.*	1 Zn, 2,6-dihydroxybenzoate
8/1/06	2DVT	2,6-dihydroxy-benzoate decarboxylase	*Rhizobium sp.*	1 Zn
11/25/08	3FDG	dipeptidase AC	*R. sphaeroides*	2 Mg
5/4/06	2GWG	4-oxalomesaconate hydratase (LigJ)	*R. palustris*	1 Zn
8/1/07	2QT3	N-isopropyl-ammelide isopropyl-amidohydrolase AtzC	*Pseudomonas sp adp*	1 Zn (α)
8/10/05	2AMX	ADA	*P. yeolii*	1 Co
8/8/07	2QVN	ADA	*P. vivax*	Guanosine-5'-monophosphate
4/10/07	2PGR	ADA	*P. vivax*	1 Zn, 2'-Deoxycoformycin
4/9/07	2PGF	ADA	*P. vivax*	1 Zn, Adenosine
12/19/05	2FFI	putative 2-pyrone-4,6-dicarboxylic acid hydrolase	*P. putida*	Phosphate
5/4/06	2GWN	DHO	*P. gingivalis*	2 Zn, β-mercaptoethnol, chloride, sulfate, glycerol, cacodylate ion
10/8/08	3ETK	amidohydro_3	*P. furiosus*	2 Mg

Table 1.1: Continued.

6/14/06	2HBX	α-amino-β-carboxymuconate-ε-semialdehyde-decarboxylase (ACMSD)	*P. flourescens*	1 Co
6/14/06	2HBV	α-amino-β-carboxymuconate-ε-semialdehyde-decarboxylase (ACMSD)	*P. flourescens*	1 Mg, 1 Zn
1/31/07	2OQL	PTE H254Q/H257F	*P. dimunita*	2 Zn, 2-[Bis-(2-hydroxy-ethyl)-amino]-2-hydroxymethyl-propane-1,3-diol
12/4/06	2O4Q	PTE G60A	*P. dimunita*	2 Zn, Dimethylarsenate
12/4/06	2O4M	PTE I106G/F132G/H257Y	*P. dimunita*	2 Zn, Dimethylarsenate, Acetate, Glycerol
5/9/00	1EZ2	PTE	*P. dimunita*	2 Zn, diisopropylmethyl phosphonate
5/9/00	1EYW	PTE	*P. dimunita*	2 Zn, triethylphosphate
10/23/07	3B40	putative peptidase K44Q, I62V, P64L, E69T, F73V, N79G, N195S, S203A, T215S, A216P, N314D	*P. aeruginosa*	None
8/24/06	2I5G	dipeptidase like T56A, K300E	*P. aeruginosa*	None
11/11/01	1KCX	dihydro-pyrimidinase related	*M. musculus*	None
1/31/98	1A4M	ADA	*M. musculus*	1 Zn, 6-hydroxy-1,6-dihydro purine nucleoside
1/31/98	1A4L	ADA	*M. musculus*	1 Zn, 2'-Deoxycoformycin

Table 1.1: Continued.

8/30/96	1UIP	ADA H238E	*M. musculus*	1 Zn, Purine riboside
8/30/96	1UIO	ADA H238A	*M. musculus*	1 Zn, 6-hydroxy-7,8-dihydro purine adenosine/ H238A
2/29/96	1FKX	ADA D296A	*M. musculus*	1 Zn, 6-hydroxy-1,6-dihydro purine adenosine
2/29/96	1FKW	ADA D295E	*M. musculus*	1 Zn, Purine riboside
12/2/94	2ADA	ADA	*M. musculus*	1 Zn, 6-hydroxy-7,8-dihydro purine adenosine
12/22/92	1ADD	ADA	*M. musculus*	1 Zn, 1-deaza-adenosine
11/29/05	2F6K	amidohydrolase II target Lp24	*L. plantarum*	1 Mn (α)
7/25/07	2QPX	putative metal dependent hydrolase	*L. casei atcc 334*	2 Zn
3/4/00	1EJW	urease at 298K	*K. aerogenes*	2 Ni
3/4/00	1EJX	urease at 100K	*K. aerogenes*	2 Ni
3/4/00	1EJV	urease H320Q	*K. aerogenes*	2 Ni
3/4/00	1EJU	urease H320N	*K. aerogenes*	2 Ni
3/4/00	1EJT	urease H219Q	*K. aerogenes*	2 Ni
3/4/00	1EJS	urease H219N	*K. aerogenes*	2 Ni
3/4/00	1EJR	urease D221A	*K. aerogenes*	2 Ni
2/5/00	1EF2	urease	*K. aerogenes*	2 Mn
2/17/98	1A5L	urease K217C	*K. aerogenes*	None
2/17/98	1A5O	urease K217C	*K. aerogenes*	2 Ni, Formate
2/17/98	1A5N	urease K217A	*K. aerogenes*	2 Ni, Formate
2/17/98	1A5M	urease K217A	*K. aerogenes*	None
2/17/98	1A5K	urease K217E	*K. aerogenes*	None
4/23/97	1FWJ	urease	*K. aerogenes*	2 Ni
4/23/97	1FWI	urease H134A	*K. aerogenes*	1 Ni (β)
4/23/97	1FWH	urease C319Y	*K. aerogenes*	2 Ni
4/23/97	1FWG	urease C319S	*K. aerogenes*	2 Ni
4/23/97	1FWF	urease C319D	*K. aerogenes*	2 Ni
4/23/97	1FWE	urease C319A	*K. aerogenes*	2 Ni, Acetohydroxamate

Table 1.1: Continued.

4/23/97	1FWD	urease C319A pH 9.4	*K. aerogenes*	2 Ni
4/23/97	1FWC	urease C319A pH 8.5	*K. aerogenes*	2 Ni
4/23/97	1FWB	urease C319A pH 6.5	*K. aerogenes*	2 Ni
4/23/97	1FWA	urease C319A pH 7.5	*K. aerogenes*	2 Ni, Carbonate
6/20/95	1KRC	urease H320A	*K. aerogenes*	2 Ni, Carbon Dioxide
6/20/95	1KRB	urease H219A	*K. aerogenes*	2 Ni
6/20/95	1KRA	urease	*K. aerogenes*	None
2/16/95	2KAU	urease	*K. aerogenes*	2 Ni
1/3/05	2BGN	dipeptidyl dipeptidase IV and ADA	*H. sapiens and B. taurus*	Tat N-terminal nonapeptide and Zn
N/A	1m7m	ADA	*H. sapiens*	Theoretical model
7/31/08	3E0L	designed ammelide deaminase	*H. sapiens*	1 Zn (α)
3/25/08	2VR2	dihydro-pyrimidinase	*H. sapiens*	2 Zn
1/24/08	2VM8	Crmp-2 D-HYD like	*H. sapiens*	None
4/26/07	2UZ9	GuaD	*H. sapiens*	1 Zn, Xanthine
4/26/06	2GSE	human dihydro-pyriminidase like	*H. sapiens*	None
2/3/02	1ITU	renal dipeptidase	*H. sapiens*	2 Zn, Cilastatin
2/2/02	1ITQ	renal dipeptidase	*H. sapiens*	2 Zn
11/1/00	1E9Z	urease	*H. pylori*	2 Ni
11/1/00	1E9Y	urease	*H. pylori*	2 Ni, Acetohydroxamate
4/29/03	1P6C	PTE H254G/H257W/L303T	*Flavobacterium sp.*	2 Zn, Diisopropyl methyl phosphonate, Diethyl-4-methylbenzyl phosphonate

Table 1.1: Continued.

4/29/03	1P6B	PTE H254G/H257W/ L303T	*Flavobacterium sp.*	3 Zn, Diethyl-4-methylbenzylphosphate, ethyl dihydrogenphosphate
9/13/06	2ICS	put. adenine deaminase	*E. faecalis*	1 Zn, Adenine
1/21/05	1YMY	NagA	*E. coli*	None
8/17/08	3E75	allantoinase	*E. coli*	2 Zn
8/17/08	3E74	allantoinase	*E. coli*	2 Fe
5/17/07	2Z2B	DHO deletion 107-116	*E. coli*	2 Zn
5/17/07	2Z2A	DHO T109G	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate
5/17/07	2Z29	DHO T109A	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydro-pyrimidine-4-carboxylate
5/17/07	2Z28	DHO T109V	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate
5/17/07	2Z27	DHO T109S	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate
5/17/07	2Z26	DHO T110A	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate
5/17/07	2Z25	DHO T110V	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate

Table 1.1: Continued.

5/17/07	2Z24	DHO T110S	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate
3/14/07	2P53	NagA D273N	*E. coli*	1 Zn, *N*-methylphosphonamidate-glucosamine-6-phosphate inhibitor
3/14/07	2P50	NagA	*E. coli*	1 Zn
2/28/07	2EG8	DHO	*E. coli*	2 Zn, 5-flouroorotate
2/28/07	2EG7	DHO	*E. coli*	2 Zn, 2-oxo-1,2,3,6-tetrahydropyrimidine-4,6-dicarboxylate
2/28/07	2EG6	DHO	*E. coli*	2 Zn
11/8/06	2E25	DHO T109S	*E. coli*	2 Zn, 5-Flouroorotate
8/18/05	2AQV	IAD Y137F	*E. coli*	2 Zn
8/18/05	2AQO	IAD E77Q	*E. coli*	2 Zn
6/14/05	1ZZM	TatD homolog	*E. coli*	1 Zn, 3,6,9,12,15,18-hexaoxaicosane-1,20-diol
2/4/05	1YRR	NagA	*E. coli*	Phosphate
1/13/05	1YIX	YcfH TatD homolog	*E. coli*	1 Zn
12/21/04	1YBQ	isoasparatyl dipeptidase D285N	*E. coli*	2 Zn, L-β-aspartylhistidine
11/2/04	1XWY	TatD deoxy-ribonuclease	*E. coli*	1 Zn
9/17/04	1XGE	DHO	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, 4(s)-2,6-dioxohexahydropyrimidine-4-carboxylate

Table 1.1: Continued.

10/31/03	1RAK	CDA D314S	*E. coli*	Fe, 5-fluoro-4-(S)-hydroxy-3,4-dihydropyrimidine
10/31/03	1RA5	CDA D314A	*E. coli*	1 Fe, 5-fluoro-4-(S)-hydroxy-3,4-dihydropyrimidine
10/31/03	1RA0	CDA D314G	*E. coli*	1 Fe, 5-fluoro-4-(S)-hydroxy-3,4-dihydropyrimidine
10/31/03	1R9Z	CDA D314S	*E. coli*	1 Fe
10/31/03	1R9Y	CDA D314A	*E. coli*	1 Fe
10/31/03	1R9X	CDA D314G	*E. coli*	1 Fe
6/15/03	1POK	isoaspartyl dipeptidase	*E. coli*	2 Zn, Asparagine
6/15/03	1POJ	isoaspartyl dipeptidase	*E. coli*	2 Zn, 2-{[[(1S-1-amino-2-carboxyethyl](dihydroxy)phosphoranyl]methyl}-4-methylpentanoic acid
6/15/03	1PO9	isoaspartyl dipeptidase	*E. coli*	2 Zn
5/14/03	1PB0	YCDX protein	*E. coli*	3 Zn
3/2/03	1ONX	isoaspartyl dipeptidase	*E. coli*	2 Zn, Aspartate
3/2/03	1ONW	isoaspartyl dipeptidase	*E. coli*	2 Zn
7/14/02	1M68	YCDX protein	*E. coli*	3 Zn, 1 Sulfate
7/12/02	1M65	YCDX protein	*E. coli*	1 Zn, 2 Na
10/17/01	1K70	CDA	*E. coli*	1 Fe, 4-hydroxy-3,4-dihydro-1H-pyrimidin-2-one inhibitor
10/17/01	1K6W	CDA	*E. coli*	1 Fe
5/16/01	1J79	DHO	*E. coli*	2 Zn, N-Carbamoyl-L-aspartate, dihydroorotate

Table 1.1: Continued.

5/27/98	1BF6	PTE homology protein	*E. coli*	2 Zn, Glycerol, (4S)-2-Methyl-2,4-pentanediol
5/27/08	2VUN	enamidase	*E. barkeri*	1 Fe, 1 Zn, 1Cl
2/27/09	3GG7	TatD like	*D. radiodurans*	1 Mn (α)
11/2/07	2ZC1	PTE	*D. radiodurans*	2 Co
10/4/06	2IMR	SAH/TMA deaminase DR_0824	*D. radiodurans*	1 Zn
1/25/06	2FTW	dihydro-pyriminidase	*D. discoideum*	2 Zn
3/17/09	3GNH	L-lysine/L-arginine carboxypeptidase	*C. vibrioides*	2 Zn, Methyl phosphonated L-Arginine
9/14/07	2RAG	Putative dipeptidase	*C. cresentus*	2 Zn, Cl
5/18/07	2Q01	Uronate isomerase	*C. cresentus*	None
9/6/06	2I9U	GuaD	*C. acetobutylium*	1 Fe, Guanine
3/13/08	3CJP	CAC3332	*C. Aceto-butylicum*	2 Zn
1/26/05	1YNY	D-hydantoinase	*Bacillus sp.*	2 Mn
2/28/09	3GGM	Bt9727_2919	*B. thuringiensis*	None
8/20/07	2Z7G	ADA	*B. taurus*	1 Zn (α), (2S,3R)-3-(6-amino-9H-purin-9-yl)nonan-2- ol
10/30/06	2E1W	ADA	*B. taurus*	1 Zn, 1-{(1R,2S)-2-Hydroxy-1-[2-(1-naphthyl)ethyl]propyl}-1H-imidazole-4-carboxamide

Table 1.1: Continued.

2/2/05	1WXZ	ADA	*B. taurus*	1 Zn, 1-((1R,2S)-1-{2-[2-(4-chlorophenyl)-1,3-benzoxazol- 7-yl]ethyl}-2-hydroxypropyl)-1H-imidazole- 4-carboxamide
2/2/05	1WXY	ADA	*B. taurus*	1 Zn, N-[4,5-Bis(4-hydroxyphenyl)-1,3-thiazol-2- yl]hexanamide
6/22/04	1W1I	ADA and dipeptidyl peptidase	*B. taurus*	1 Zn, Dipeptidyl peptidase (IV)
4/16/04	1VFL	ADA	*B. taurus*	1 Zn
12/14/03	1V7A	ADA	*B. taurus*	1 Zn, 1-{(1R,2S)-2-Hydroxy-1-[2-(2-naphthyloxy)ethyl]propyl}- 1H-imidazone-4-carboxamide
12/14/03	1V79	ADA	*B. taurus*	1 Zn, 1-{(1R,2S)-1-[2-(2,3,-Dichlorophenyl)ethyl]-2-hydroxypropyl}-1H-imidazole-4-carboxamide
10/5/03	1O5R	ADA	*B. taurus*	1 Zn, 1-[(1R)-3-(6-{[(Benzylamino)carbonyl]amino}- 1H-indol-1-yl)-1-(hydroxymethyl)propyl]-1H- imidazole-4-carboxamide
10/3/03	1UML	ADA	*B. taurus*	1 Zn, 1-((1R)-1-(hydroxymethyl)-3-{6-[(3-phenylpropanoyl)amino]- 1H-indol-1-yl}propyl)-1H-imidazole-4-carboxamide

Table 1.1: Continued.

9/8/03	1QXL	ADA	*B. taurus*	1 Zn, 1-((1R)-1-(hydroxymethyl)-3-(1-naphthyl)propyl)- 1H-imidazole-4-carboxamide
12/9/02	1NDZ	ADA	*B. taurus*	1 Zn, 1-((1R)-1-(hydroxymethyl)-3-(6-((3-(1-methyl- 1H-benzimidazol-2-yl)propanoyl)amino)-1H- indol-1-yl)propyl)-1H-IMIDAZOLE-4-carboxamide
12/9/02	1NDY	ADA	*B. taurus*	1 Zn, 1-((1R)-1-(hydroxymethyl)-3-(1-naphthyl)propyl)- 1H-imidazole-4-carboxamide
12/9/02	1NDW	ADA	*B. taurus*	1 Zn, 1-((1R)-1-(hydroxymethyl)-3-phenylpropyl)- 1H-imidazole-4-carboxamide
12/9/02	1NDV	ADA	*B. taurus*	1 Zn, N''-(4-(5-((1H-benzimidazol-2-ylamino)methyl)- 2-thienyl)-1,3-thiazol-2-yl)guanidine
1/10/02	1KRM	ADA	*B. taurus*	1 Zn, 6-hydroxyl-1,6-dihydropurine riboside
11/22/07	2VHL	NagA	*B. subtilis*	2 Fe, glucosamine-6-phosphate,
11/22/07	2VHL	NagA	*B. subtilis*	2 Fe, Glucosamine-6-phosphate product
2/19/06	2G3F	imidazolone-propionase	*B. subtilis*	1 Zn, substrate analogue imidazole-4-acetate
10/16/05	2BB0	imidazolone-propionase	*B. subtilis*	1 Zn, acetate
9/25/01	1K1D	D-hydantoinase	*B. stereo-thermophilus*	2 Zn

Table 1.1: Continued.

12/14/02	1NFG	D-hydantoinase	*B. pickettii*	2 Zn
1/14/04	1S3T	urease	*B. pasteurii*	2 Ni, Sulfate, Borate
4/9/01	1IE7	urease	*B. pasteurii*	2 Ni, Phosphate
2/25/99	4UBP	urease	*B. pasteurii*	2 Ni, Acetohydroxamate
12/16/98	3UBP	urease	*B. pasteurii*	2 Ni, Diamidophosphate
11/4/98	2UBP	urease	*B. pasteurii*	2 Ni, Sulfate
1/21/98	1UBP	urease	*B. pasteurii*	2 Ni, β-mercaptoethanol
3/24/07	2P9B	putative prolidase	*B. longum*	None
1/25/07	2OOD	GuaD	*B. japonicum*	1 Zn, Guanine
6/25/07	2QEE	BH0493	*B. halodurans*	1 Zn (α)
6/5/07	2Q6E	uronate isomerase	*B. halodurans*	1 Zn (α)
5/21/07	2Q08	uronate isomerase	*B. halodurans*	1 Zn (α)
4/24/07	2PNK	BH0493	*B. halodurans*	None
8/7/08	3E3H	PTE H254R, H275F	*B. dimunita*	2 Co, Diethyl 4-methylbenzyl-phosphonate
12/18/06	2OB3	PTE H257Y/L303T	*B. dimunita*	2 Zn
9/1/03	1QW7	PTE	*B. dimunita*	2 Co, Diethyl-4-methylbenzyl-phosphonate
6/26/01	1JGM	PTE	*B. dimunita*	2 Cd, Ethanol, Formate, 2-phenyl-ethanol
1/29/01	1I0D	PTE	*B. dimunita*	Zn, Cd, Ethanediol, Formate, 2-phenyl-ethanol
1/29/01	1I0B	PTE	*B. dimunita*	2 Mn, Ethanediol, Formate, 2-phenyl-ethanol
1/26/01	1HZY	PTE	*B. dimunita*	2 Zn, Ethanediol, Formate, 2-phenyl-ethanol
2/13/96	1DPM	PTE	*B. dimunita*	2 Zn, Diethyl-4-methylbenzyl-phosphonate
4/25/95	1PSC	PTE	*B. dimunita*	2 Cd, Diethyl-4-methylbenzyl-phosphonate

Table 1.1: Continued.

7/7/94	1PTA	PTE	*B. dimunita*	None
4/8/08	3CS2	PTE G60A	*B. diminuta*	2 Co, Dimethylarsenate (cacodylate)
2/20/08	3CAK	PTE	*B. diminuta*	2 Co, Ethylphosphate
7/31/08	3E0F	putative PTE	*B. adolescentis*	2 Fe, 1Zn, acetate, phosphate
2/10/08	3C86	OpdA PTE S92A	*A. tumefaciens*	1 Co, 1 Fe, Diethyl thiophosphate product, 1,2-ethanediol
8/23/07	2R1P	OpdA PTE S92A, K185R	*A. tumefaciens*	1 Co, 1 Fe , Diethyl thiophosphate product, 1,2-ethanediol
8/23/07	2R1O	OpdA PTE S92A	*A. tumefaciens*	1 Co, 1 Fe, Diethyl thiophosphate product, 1,2-ethanediol
8/23/07	2R1N	OpdA PTE S92A, N265D	*A. tumefaciens*	1 Co, 1 Fe, Diethyl 4-methoxyphenyl phosphate slow substrate,
8/23/07	2R1M	OpdA PTE S92A, N265D	*A. tumefaciens*	1 Co, 1 Fe, Diethyl phosphate product, 1,2-ethanediol
8/23/07	2R1L	OpdA PTE S92A, N265D	*A. tumefaciens*	1 Co, 1 Fe, Diethyl thiophosphate, 1,2-ethanediol
8/23/07	2R1K	OpdA PTE S92A, N265D	*A. tumefaciens*	1 Co, 1 Fe, Diethyl phosphate, 1,2-ethanediol
5/9/07	2PUZ	imidazolonepropionase	*A. tumefaciens*	1 Fe, N-formimino-L-glutamate product
1/5/07	2OGJ	DHO-like	*A. tumefaciens*	2 Zn, Imidazole
4/13/06	2GOK	imidazolonepropionase	*A. tumefaciens*	1 Fe
9/8/05	2D2J	PTE	*A. tumefaciens*	2 Co
9/8/05	2D2H	PTE	*A. tumefaciens*	2 Co, Trimethylphosphate

Table 1.1: Continued.

9/8/05	2D2G	PTE	*A. tumefaciens*	2 Co, Dimethylthiophosphate
6/25/05	2A3L	AMP deaminase	*A. thaliana*	1 Zn, Coformycin 5'-phosphate
7/30/07	2QS8	Xaa-Pro dipeptidase	*A. macleodii*	1 Mg (α) , Methionine
3/30/09	3GUW	TatD like	*A. Fulgidus*	2 Zn
11/20/03	1V51	D-aminoacylase	*A. faecalis*	2 Zn, Acetate
11/20/03	1V4Y	D-aminoacylase H220A	*A. faecalis*	1 Zn, AcetateLigand free/ Mutant
11/20/03	1RK6	D-aminoacylase	*A. faecalis*	1Zn, 1 Cd, Acetate
11/20/03	1RK5	D-aminoacylase D366A	*A. faecalis*	1 Zn, 1 Cu, Acetate
11/20/03	1RJR	D-aminoacylase D366A	*A. faecalis*	2 Zn, Acetate
11/20/03	1RJQ	D-aminoacylase D366A	*A. faecalis*	1 Zn, Acetate
11/20/03	1RJP	D-aminoacylase	*A. faecalis*	1 Zn, 1 Cu, Acetate
7/22/02	1M7J	D-aminoacylase	*A. faecalis*	2 Zn, Acetate
8/20/01	1GKR	L-hydantoinase	*A. aurescens*	2 Zn
5/20/08	3D6N	DHO	*A. Aeolicus*	1 Zn, Citrate
10/15/04	1XRT	DHO	*A. aeolicus*	2 Zn
10/14/04	1XRF	DHO	*A. aeolicus*	2 Zn, Sulfate

Based on the structurally characterized members of the AHS, there are several identified variations of the divalent metal centers, as depicted in **Figure 1.2**. The most common metal center is the binuclear metal center found in enzymes such as phosphotriesterase, dihydroorotase (DHO), isoaspartyl dipeptidase (IAD), urease, and hydantoinase (**Figure 1.2A**). In all of these proteins, the two metal ions are ligated to the protein via six amino acid side chains. According to the crystal structures, M_α, the more buried metal ion, is coordinated to the two histidines in the HxH motif at the end of β-strand 1 and an aspartate from β-strand 8. The more solvent exposed metal, M_β, is

ligated to the imidazole side chain of two histidines at the end of β-strands 5 and 6. The two metal ions are bridged by a solvent hydroxide as well as a carbamate functional group of a posttranslational modification of a carboxylated lysine from strand 4 (*55*). . One variation of this metal center subtype is the replacement of the bridging carboxylated lysine from β-strand 4 to a glutamate (**Figure 1.2A)**. This is seen in the structure of the PTE homology protein (PHP) from *E. coli*. Another variation of this metal center subtype is seen with the crystal structure of an organophosphorus hydrolase from *Deinococcus radiodurans* in which the β-metal ion is coordinated to a tyrosine from β-strand 3 in addition to the other four ligands (**Figure 1.2B)**. In the crystal structure of the enzyme *N*-acetyl glucosamine-6-phosphate deacetylase (AGD) from *Bacillus subtilis*, the binuclear metal center ligands are conserved except the bridging residue glutamate which derives from strand 3 instead of strand 4 (**Figure 1.2C**) (*42*). A more drastic variation to the binuclear center is seen with the enzyme renal dipeptidase (RDP) from humans (**Figure 1.2D**). According to the crystal structure, the M_α ligand motif from strand 1 is HxD instead of HxH, and the residue that bridges the two metals is a glutamate at the end of β-strand 3 instead of strand 4. Another significant difference in the structure of RDP is that the conserved aspartate from strand 8 is no longer a metal ligand to M_α, instead it interacts with the hydrolytic water molecule via hydrogen bonding (*40*).

Another subtype of metal center among the AHS enzymes is found in adenosine deaminase (ADA) and cytosine deaminase (CDA) (**Figure 1.2E**). These enzymes contain a single metal ion in the active site at the position that corresponds to the M_α

position of the binuclear center. According to the structures, the metal ion is coordinated to the two histidines from strand 1 as well as the conserved histidine from strand 5. The histidine from strand 6 is not a metal ligand but appears to play a role in catalysis. The residue that houses the bridging carboxylate from strand 4 is missing. In addition to the protein, the metal is also ligated to a solvent water molecule (*56-57*). One of the newest additions to this subtype is the protein α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) from *Pseudomonas fluorescens*, a member of the amidohydrolase that catalyzes a nonhydrolytic reaction. It has been suggested that this enzyme utilizes similar chemistry to that employed by ADA, CDA, and other M_{α} mononuclear enzymes of the AHS in its catalytic mechanism (*50,58*). One variation of this subtype metal center is seen in the structure of γ-resorcylate decarboxylase in which the HxH motif at the end of β-strand 1 has been replaced by an ExH motif (**Figure 1.2F)**. Another variation of the α-metal center is the replacement of the aspartate from β-strand 8 with a glutamate, as seen in the structure of 4-oxalomesaconate hydratase (PcmD) (**Figure 1.2G)**. A structure of a TatD-related DNase from *D. radiodurans* (PDB code 3GG7) indicates another metal center variation in which the histidine from strand 5 is present but no longer a metal ligand, instead a glutamate from strand 4 is now coordinating to the metal (**Figure 1.2H**). Another structure of an enzyme belonging to the TatD-related DNase family from *Saccharomyces cerevisiae* (PDB code 3e2v) contains a magnesium ion in which the metal is coordinated to an aspartate from strand 8, a histidine from strand 6, and three solvent water molecules. The histidine from strand 5 is conserved but it is not a metal ligand (**Figure 1.2I**).

The final variation of the α-metal center is represented by the uronate isomerase (URI) family. The crystal structures for three members of this family have been deposited in the Protein Data Bank (PDB codes 1j5s, 2q01, and 2q6e). From the structure of the *Thermotoga maritima* and *Caulobacter crescentus*, no metal ion is found in the active site. However, metal ligands for the M_α are fully present, suggesting that the proteins are capable of binding a divalent cation in the M_α site (**Figure 1.2J**). In addition, the crystal structure of a newly discovered second uronate isomerase from *Bacillus halodurans* (Bh0493) reveals the presence of one zinc ion in the active site. According to the structure, the zinc is coordinated to the two histidines in the HxH motif at the end of β-strand 1 and an evolutionary conserved aspartate from β-strand 8. Sequence and structural alignment indicate that the histidine from strand 5 is conserved among the URI from *E. coli*, *T. maritima*, and *C. crescentus*, but is no longer present in the Bh0493 enzyme (**Figure 1.2K**). Moreover, the histidine from strand 6 is also not conserved among the uronate isomerase family (*42-43*). A slight variation from this is seen with the structure of Dr0824 (PDB code 3imr), a hypothetical protein from *D. radiodurans* (**Figure 1.2L**). According to the structure, the enzyme contains a zinc ion in the M_α position where it is coordinating to the HxH from strand 1, a histidine from strand 5, and a water molecule. The histidine from strand 6 is present but does not take part in the metal coordination. The aspartate from strand 8 is also conserved but it too is not a metal ligand, instead it interacts with the water molecule that is coordinating to the metal via hydrogen bonding.

A variation of the mononuclear metal center is with the enzyme AGD from *T. maritima* (**Figure 1.2M**). This enzyme has all the metal ligands found in the PTE subgroup with the exception of the bridging carboxylate from strand 4. Instead, a glutamate from strand 3 is positioned as a potential bridging residue. Though the crystal structure only shows a single metal ion in the active site, the presence of all the conserved metal ligands suggests that the enzyme from *T. maritima* is capable of binding to two metal ions under the right conditions, but only a single metal ion is required for catalytic activity. In the crystal structure of AGD from *E. coli*, electron density corresponds to a single metal ion in the active site. In addition, the two histidines from strand 1 have been replaced with a glutamine and an asparagine. Unlike ADA and CDA, the metal ion in the active site of AGD from *T. maritima* and *E. coli* occupies the M_β site instead of the M_α site. From the structure, the metal ion is coordinated by the imidazole side chain of the two histidines from strand 5 and 6, a glutamate at the end of strand 3, and a solvent water molecule (**Figure 1.2N**) (*42*).

Another variation of a single divalent cation center in the M_β positon is seen with the enzyme D-amino acid deacetylase (DAA), as illustrated in **Figure 1.2O**. From the crystal structure of DAA, the histidines from strand 1 do not interact with the metal. The metal is coordinated by the imidazole side chain of the histidines from strand 5, the imidazole side chain of the histidine from strand 6, and an unusual cysteine residue from strand 2. The aspartate from strand 8 is positioned to act as an activator of the hydrolytic water. It has been shown that the second metal can be forced into the active site of DAA, but its presence is not necessary for enzymatic activity (*45,59*). Another variation of the

β-metal center subtype is seen in the structure of TatD (**Figure 1.2P**). According to the crystal structure, the metal is coordinated to a cysteine and a histidine from strand 6 as well as a glutamate in the ExD motif from strand 8.

Final variations of metal center in the AHS are with the structures of the histidinol-phosphate phosphatase family (PHP). According to the structure of PHP from *T. thermophilus* HB8 (**Figure 1.2Q**), the enzyme contains a trinuclear metal cluster in which there are two irons and one zinc (*53*). The distances between Fe_1-Fe_2, Fe_1-Zn, and Fe_2-Zn are 3.47, 4.70, and 5.84 Å, respectively. The two irons represent the binuclear center and their metal coordination is similar to the ones seen in the *B. subtilis* AGD subtype. The zinc, designated as M_γ, represents a novel subtype of metal center with distinct metal ligands (a histidine from loop 1, a histidine from strand 3, and a histidine from strand 8). Another structure belonging to the PHP family contains up to 4 zinc ions in the active site (**Figure 1.2R**). The metals in the tetranuclear metal cluster are designated as M_α, M_β, M_γ and M_δ sites. The M_α and M_β sites resemble the binuclear center in which the M_α is liganded to the HxH from strand 1, the aspartate from strand 8, and is bridged to the M_β by a glutamate from strand 3. The M_β, however, is coordinated to a histidine from strand 4 in addition to the glutamate from strand 3. The M_γ metal interacts with an aspartate from the loop after strand 1, a histidine from strand 2, a histidine from strand 8, and a water molecule. The metal M_δ of the tetranuclear cluster is bridged to the M_α by the aspartate from strand 8 as well s coordinating to an aspartate from strand 5. The structures of HolPase-family, therefore, represent extreme cases of structural diversity within the AHS.

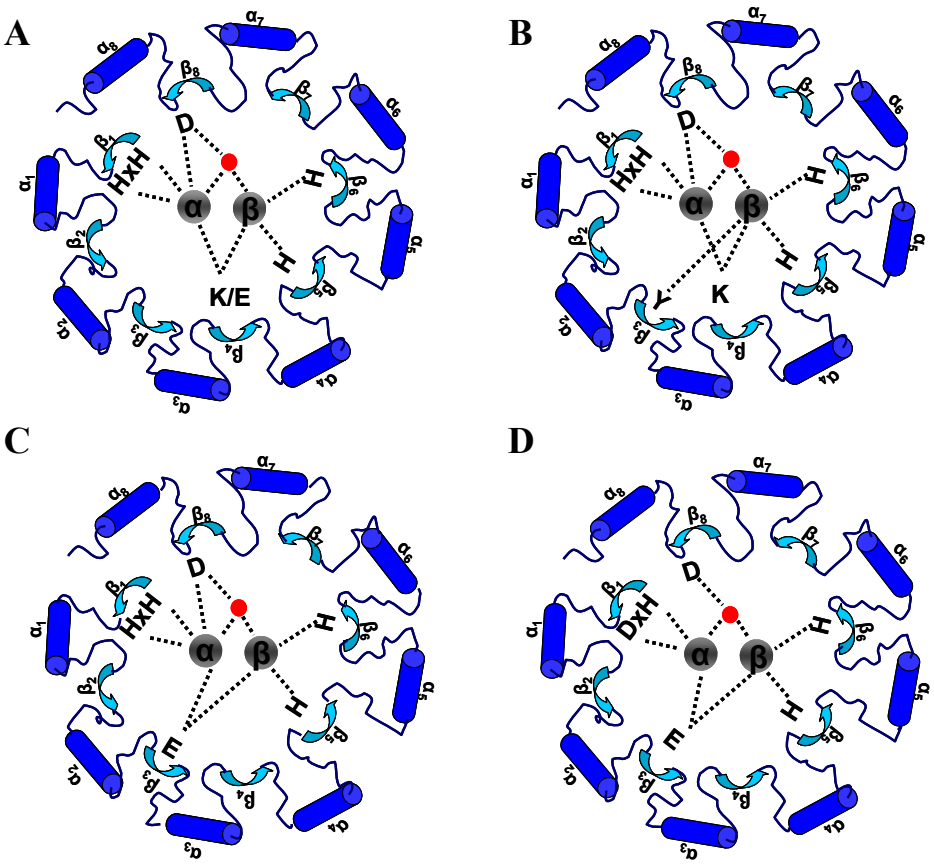


Figure 1.2: Variations of binuclear and mononuclear metal centers in the AHS. The metal ions are in grey and the water molecules are in red. Representations of the different variations are PTE from *P. diminuta* (A), PTE from *D. radiodurans* (B), AGD from *B. subtilus* (C), hRDP (D), CDA (E), rdc (F), PcmD (G), a TatD-related DNase from *D. radiodurans* (H), an unknown from *S. cerevisiae* (I), URI (J), URI from *B. halodurans* (Bh0493) (K), Dr0824 (L), AGD from *T. maritima* (M), AGD from *E. coli* (N), DAA (O), TatD (P), and HolPase-family (Q,R).

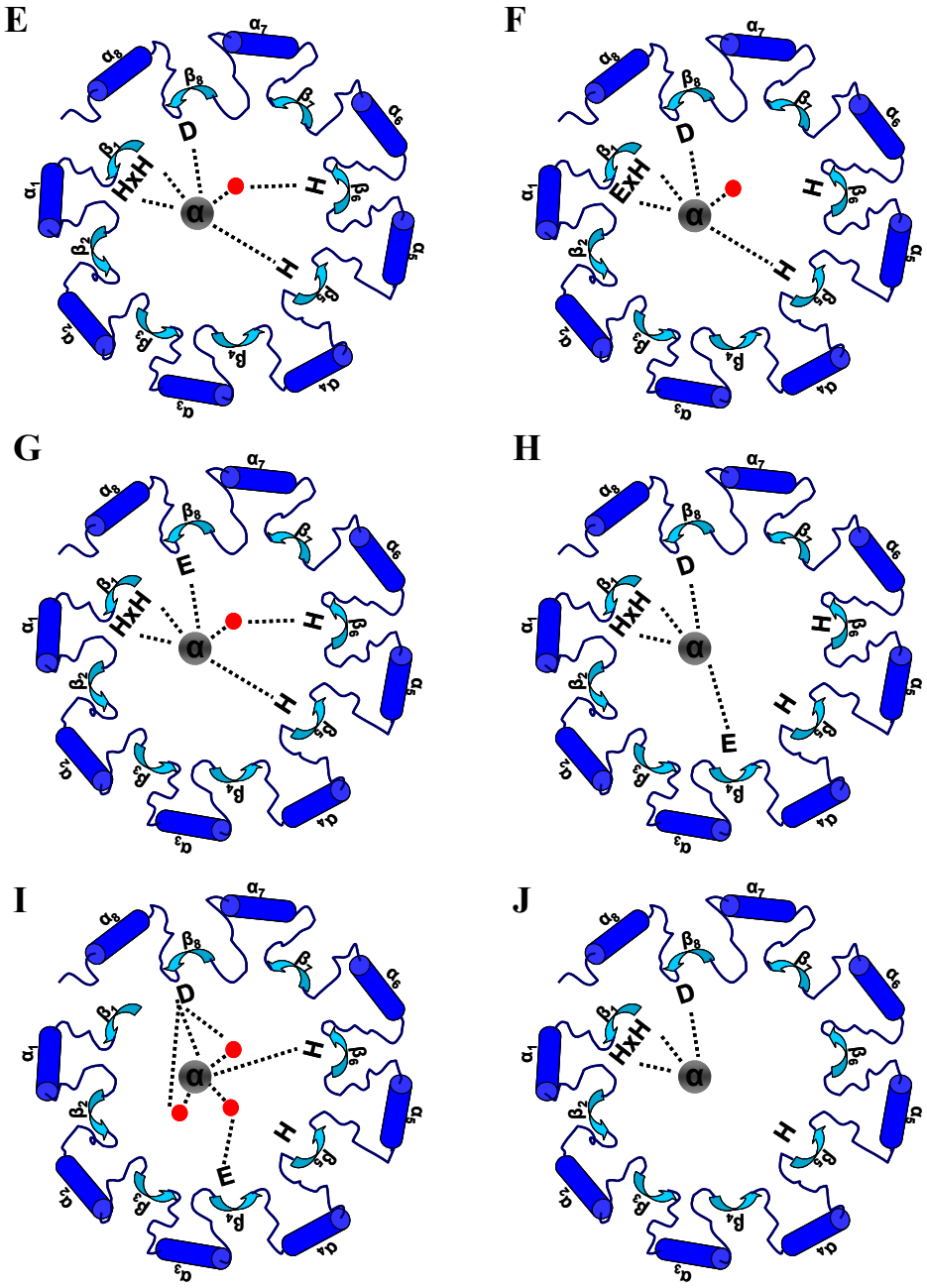


Figure 1.2: Continued.

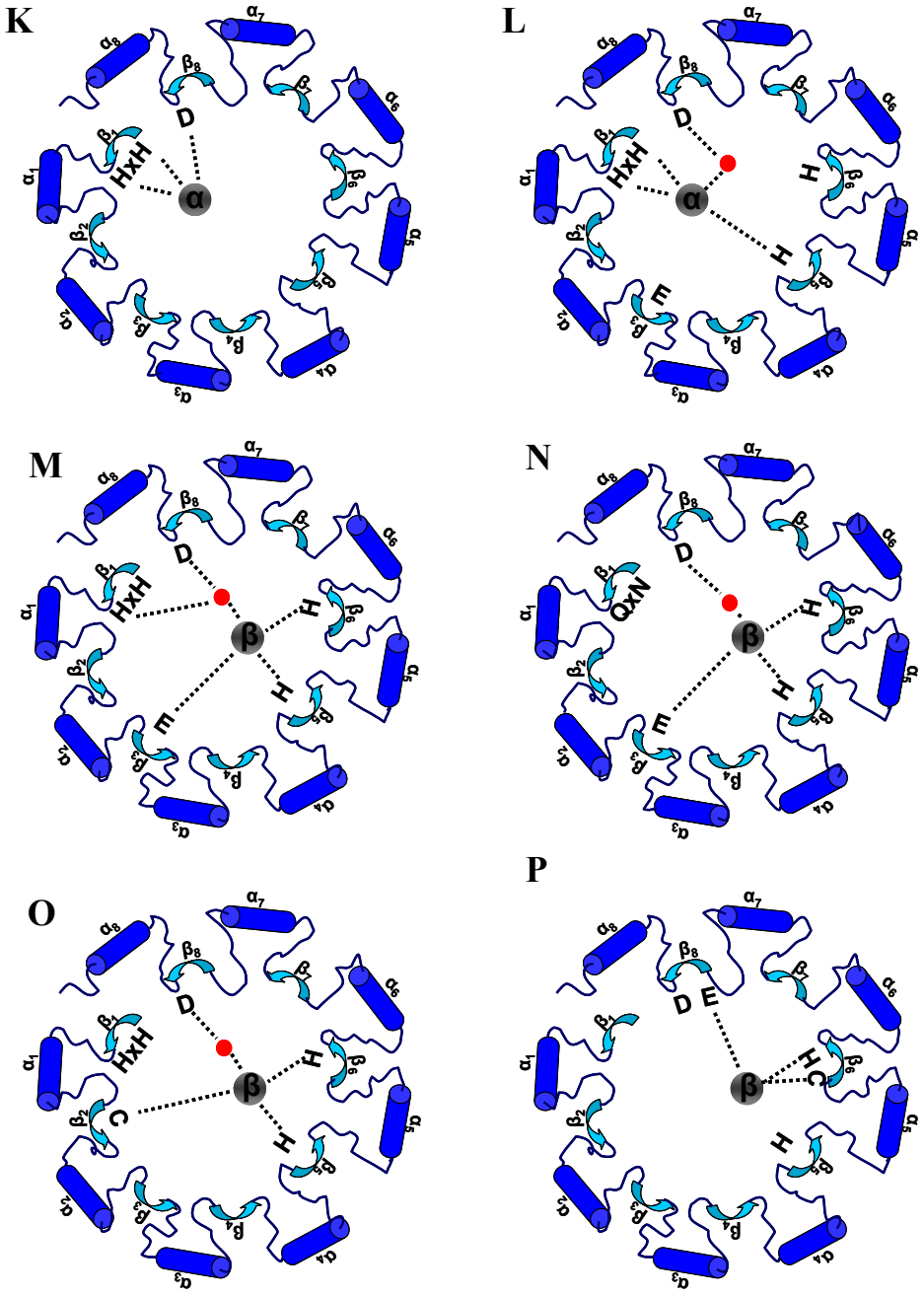


Figure 1.2: Continued.

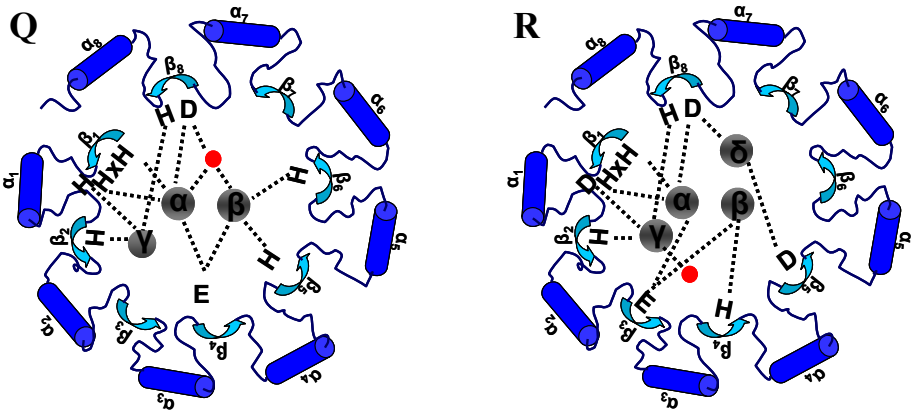
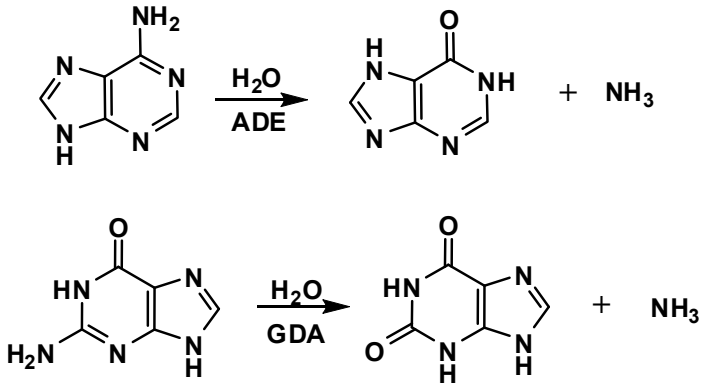


Figure 1.2: Continued.

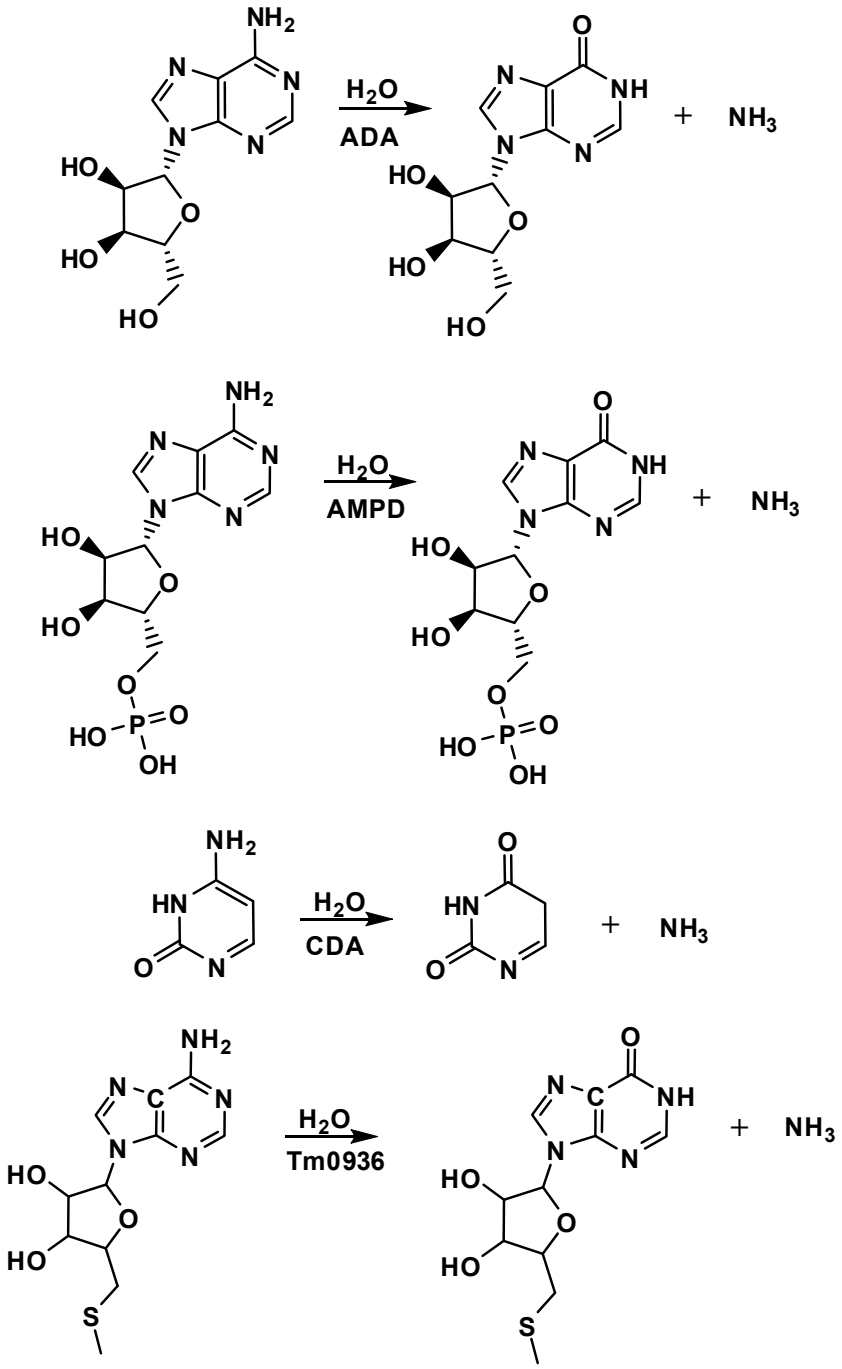
The amidohydrolase superfamily is a functionally diverse group of enzymes, and members of this superfamily have been found in every organism sequenced to date (*31*). These enzymes are found in the pathways of pyrimidine metabolism (CDA, DHO, DHP, and HYD) (*,36-38,47,60*), purine metabolism (ADA, ADE, AMPD, Tm0936, GDA, URE and ATase) (*7,32,39,46-48,61-62*), atrazine degradation (AtzA, AtzN, AtzB, AtzC, trzA, and TriA) (*63-67*), organophosphate degradation (PTE and adpB) (*68-69*), histidine metabolism (HolPase, HutI and HutF) (*3,49,53*), tryptophan catabolism (ACMSD) (*50*), sugars metabolism (AGD and URI) (*70,44*), amino acid metabolism (Cc3125, Cc0300, RDP, IAD, D-ANase, D-Amase, D-AAase, and D-AGase) (*5-6,40,71-74*), nicotinate catabolism (ENA) (*52*), phthalate catabolism (PcmC and PcmD) (*75*), hydroxybenzoates biodegradation (rdc) (*76*), lignin-related biphenyl structure degration (ligW and ligY) (*77-78*), and thymidine salvage (IDCase) (*79*). Since its

initial discovery, there are approximately forty families that have been functionally characterized and the number of sequences available on the Structure-Function Linkage Database (SFLD) now exceeds 10,000 (*10,80*). The reactions catalyzed by the enzymes of the AHS are predominantly hydrolytic reactions of the C-N amide bond of amino acids, nucleic acids, and sugars. These enzymes also catalyze reactions involing C-O and P-O ester bond cleavage. The mechanism of hydrolysis for this enzyme superfamily involves the activation of a water molecule by the metal center for nucleophilic attack. The metal ion also plays a role in the activation of the scissile bond of the substrate for bond cleavage and takes part in the stabilization of the tetrahedral intermediate. In addition to hydrolytic reactions, a few enzymes of this superfamily also catalyze nonhydrolytic reactions such as nonoxidative decarboxylation and isomerization reactions. Examples of reactions catalyzed by AHS enzymes (as of June 12, 2009) are catalogued in **Scheme 1.1**.

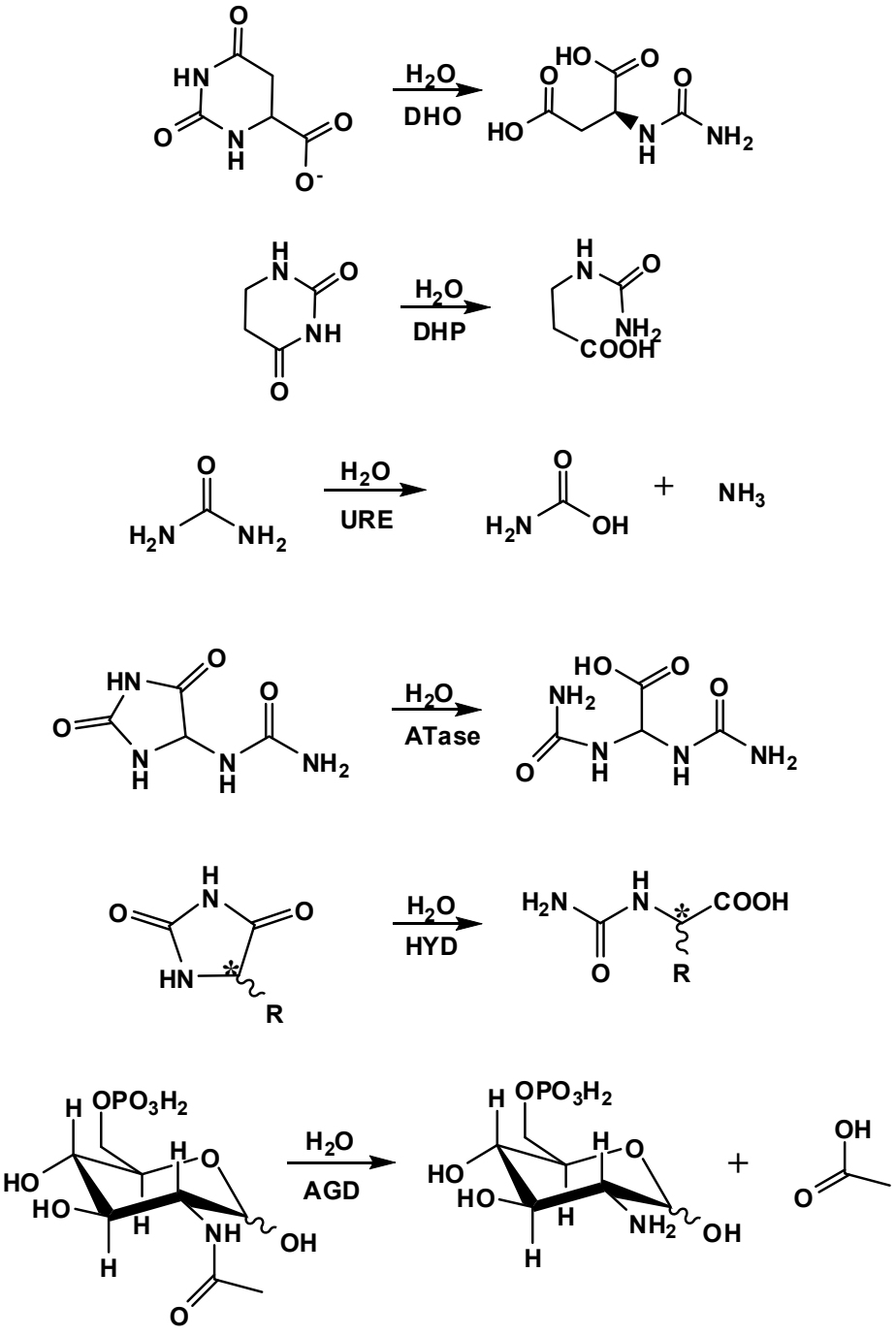
Scheme 1.1: Examples of reactions catalyze by enzymes of the AHS.



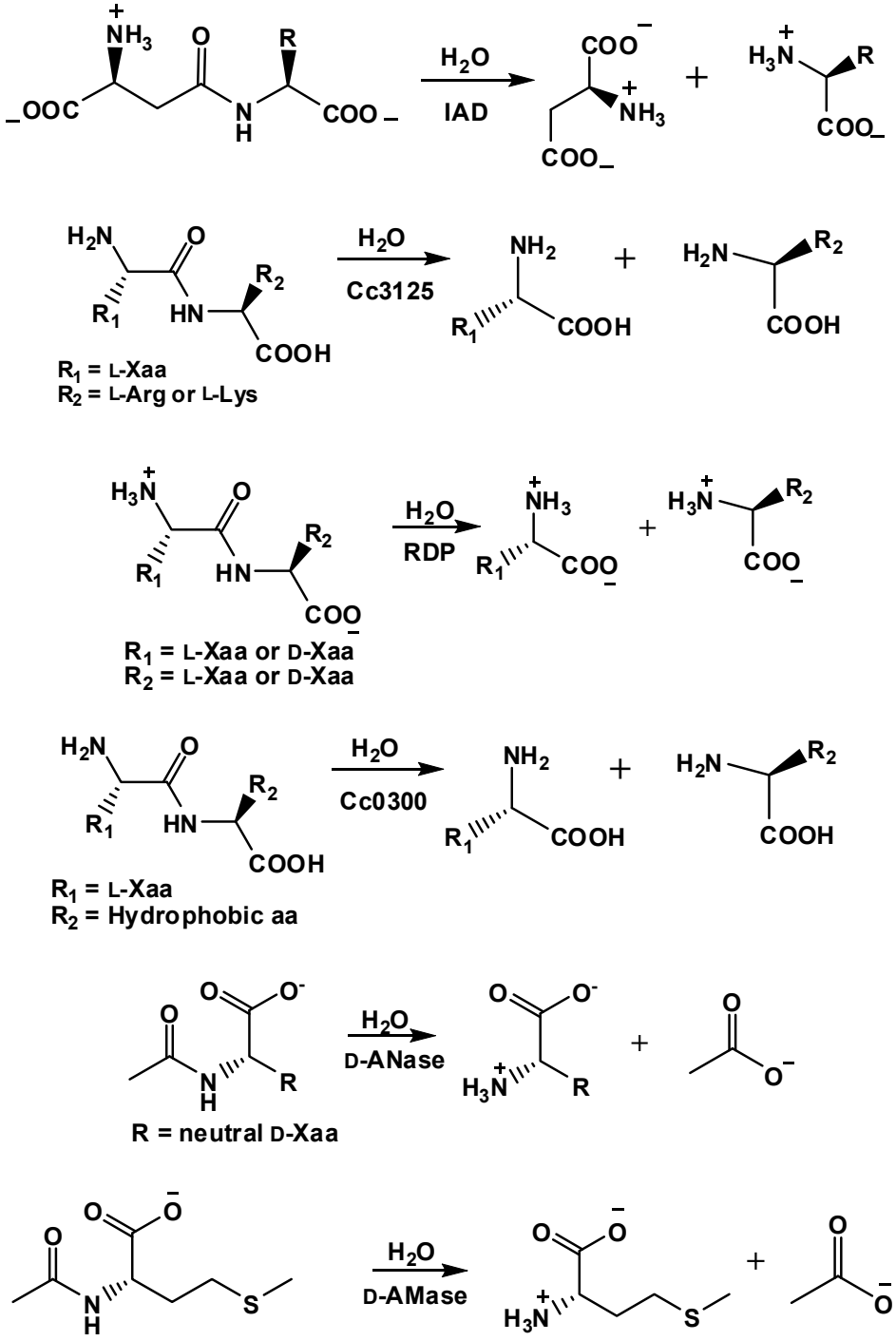
Scheme 1.1: Continued.



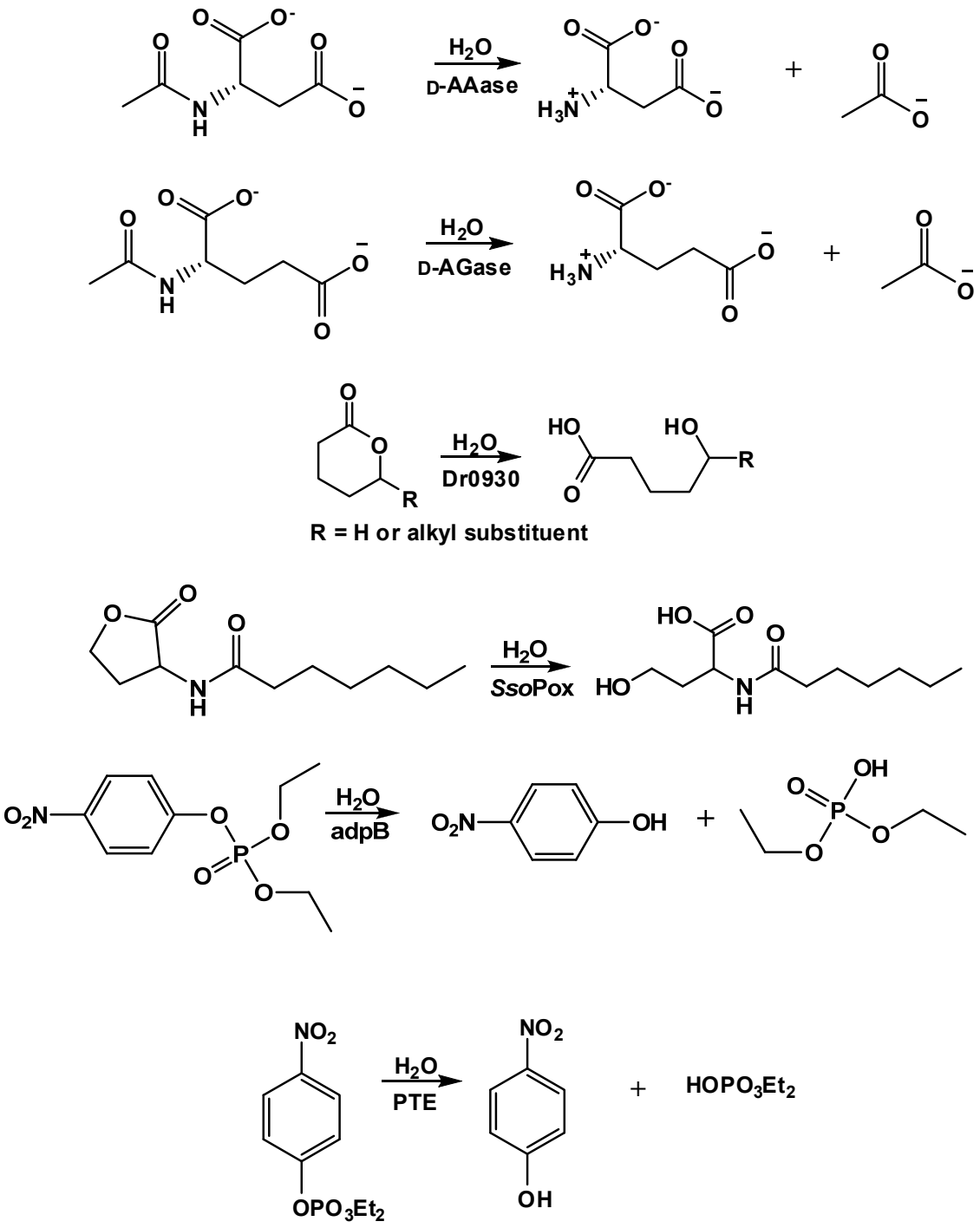
Scheme 1.1: Continued.



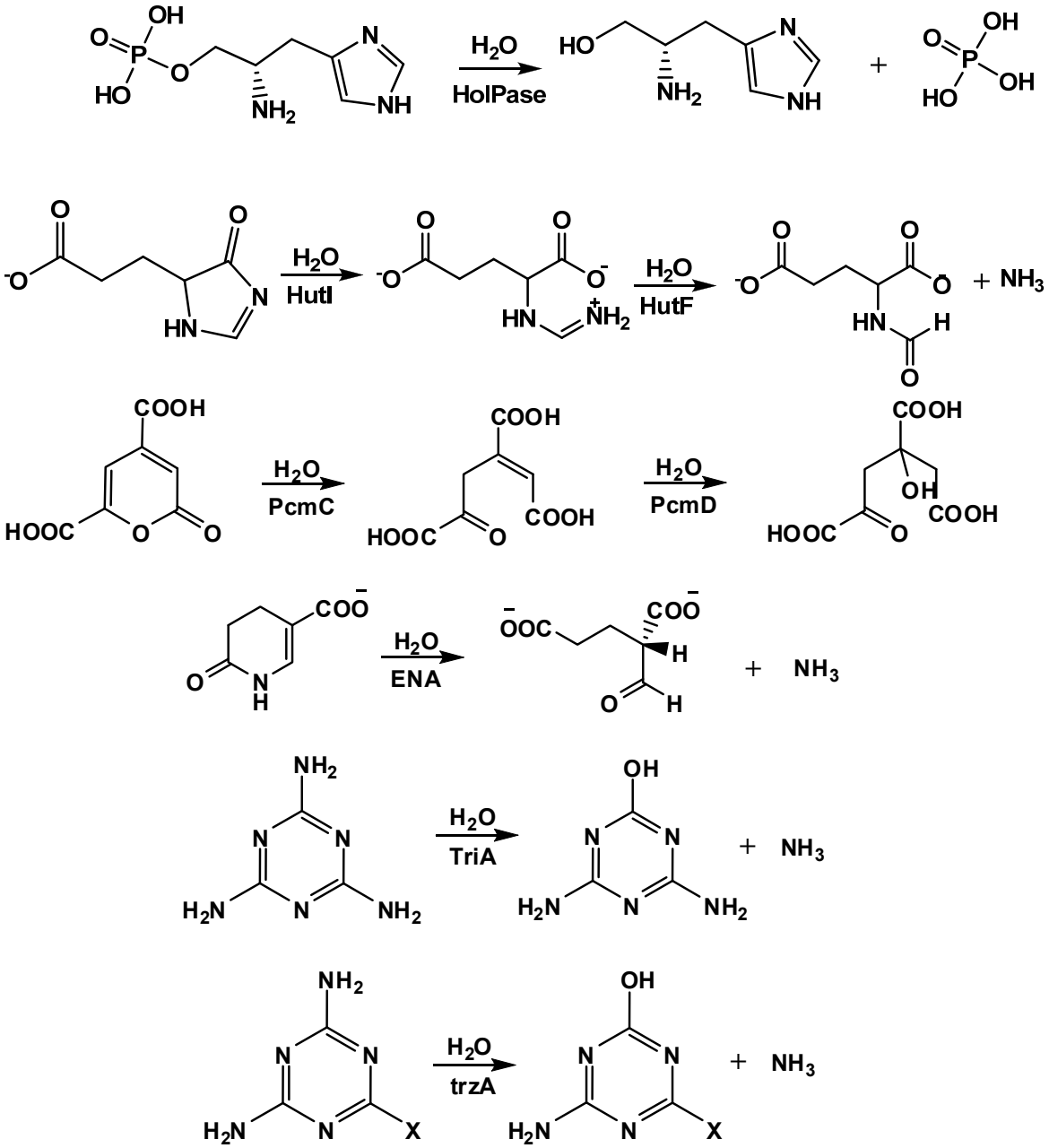
Scheme 1.1: Continued.



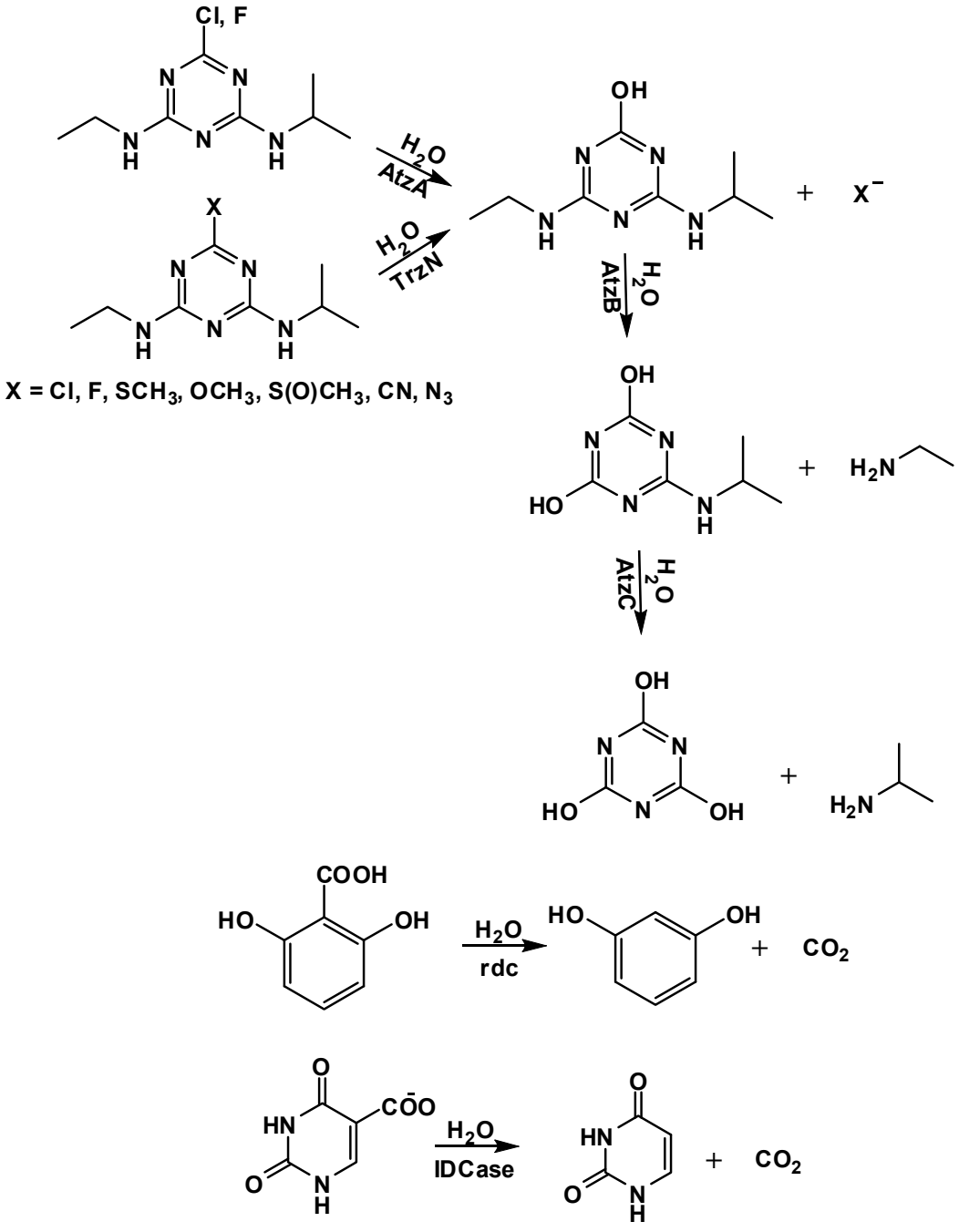
Scheme 1.1: Continued.



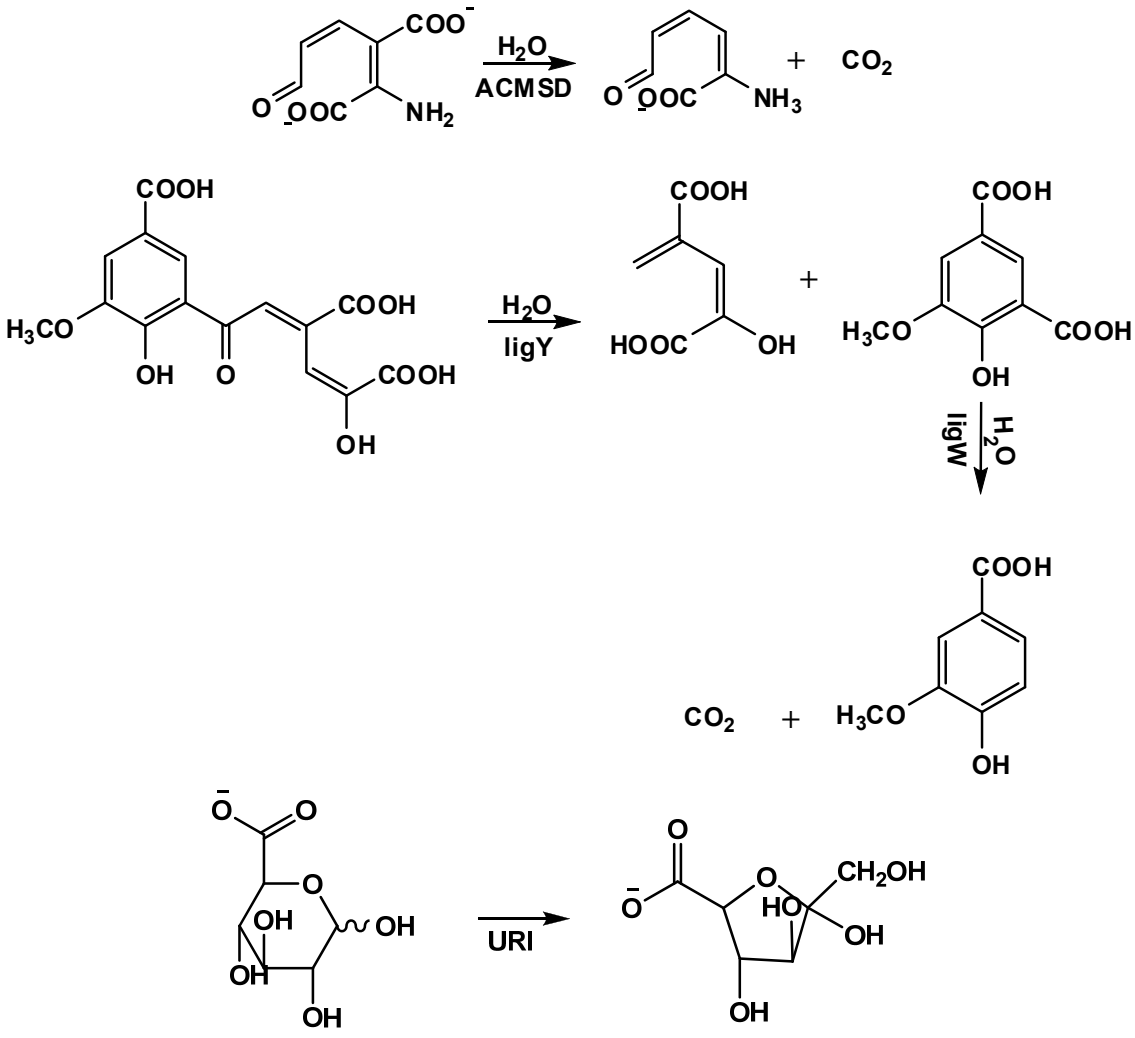
Scheme 1.1: Continued.



Scheme 1.1: Continued.



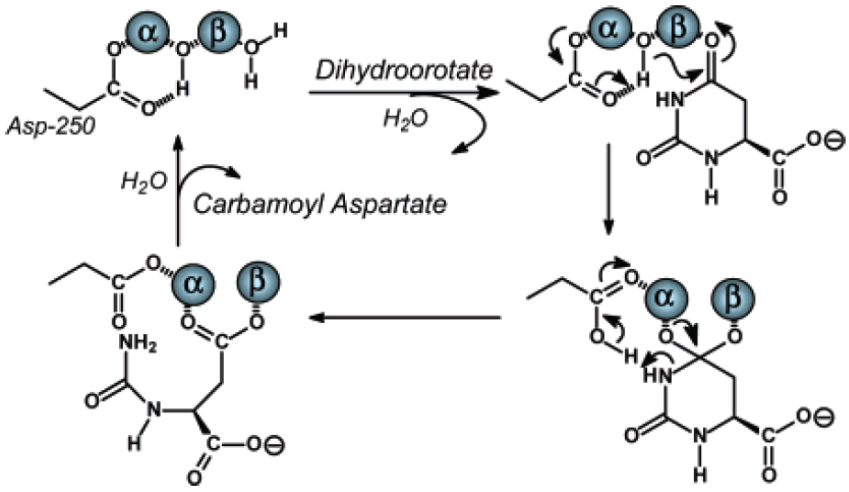
Scheme 1.1: Continued.



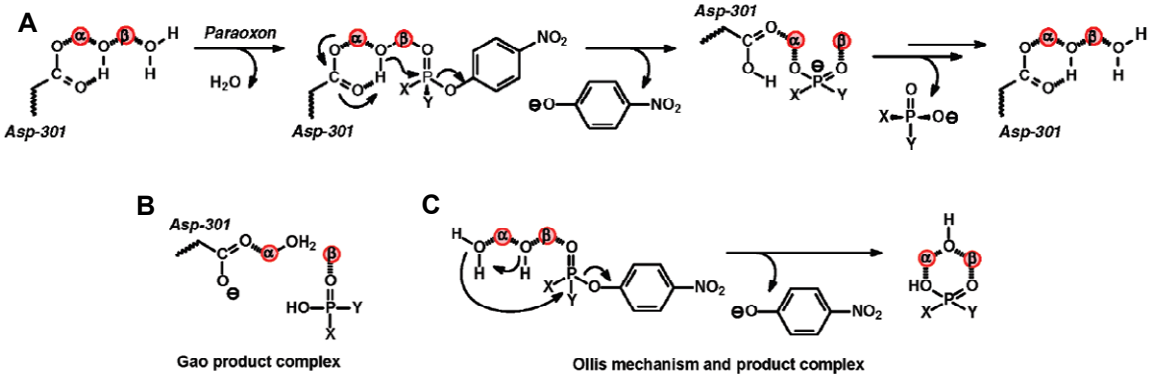
The detailed reaction for some of the enzymes of the AHS has been examined in depth. The best understood case is with the enzyme dihydroorotase from *E. coli* in which the crystal structure has been solved with substrate and product bound to separate monomeric subunits within the dimeric protein (*34*). The specific interactions within the enzyme active site before and immediately after catalysis are therefore manifested. Dihydroorotase catalyzes the reversible hydrolysis of dihydroorotate to carbamoyl aspartate. In the structure with dihydroorotate bound in the active site, the carbonyl group of the scissile amide bond is shown to be interacting with the M_β metal ion. This interaction helps polarize the amide bond and enhances the electrophilicity of the carbonyl carbon. The bridging hydroxide is hydrogen bonded to the aspartate from β-strand 8 and is positioned for the nucleophilic attack on the *re*-face of the carbonyl substrate. In the structure of DHO complexed with carbomoyl aspartate, the forming carboxylate group bridges the two metal ions in the active site. This indicates that the hydrolytic water is the bridging hydroxide. The conserved aspartate from strand 8 is positioned to protonate the amide nitrogen, suggesting its role as a proton shuttle between substrate and product (*34,60*). An illustration of the mechanism of DHO is shown in **Scheme 1.2** (*31*). Similar mechanisms have been proposed for isoaspartyl dipeptidase (*71*) and phosphotriesterase (*68*). Alternate mechanisms have also been proposed for PTE. In the PTE computational-based mechanism proposed by Wong and Gao, as the reaction advances from the enzyme-substrate complex to the enzyme-product complex, the diethyl phosphate product coordinates to the β-metal ion in a mononuclear manner. This coordination scheme is coupled with a lengthening of the

metal-metal distance from 3.6 to 5.3 Å (*81*). Another proposed mechanism for PTE has been postulated by Ollis et al. In this mechanism, Ollis and colleagues proposed that the bridging hydroxide acts as a base and abstracts a proton from a water molecule that is appararently loosely coordinated to the M_{α}. This mechanism is distinctly different from the previous two mechanisms where the bridging hydroxide is proposed as the nucleophile (*82*). The three PTE mechanisms are presented in **Scheme 1.3**.

Scheme 1.2: Proposed mechanism for the reversible hydrolysis of dihydroorotate by DHO (*31*).

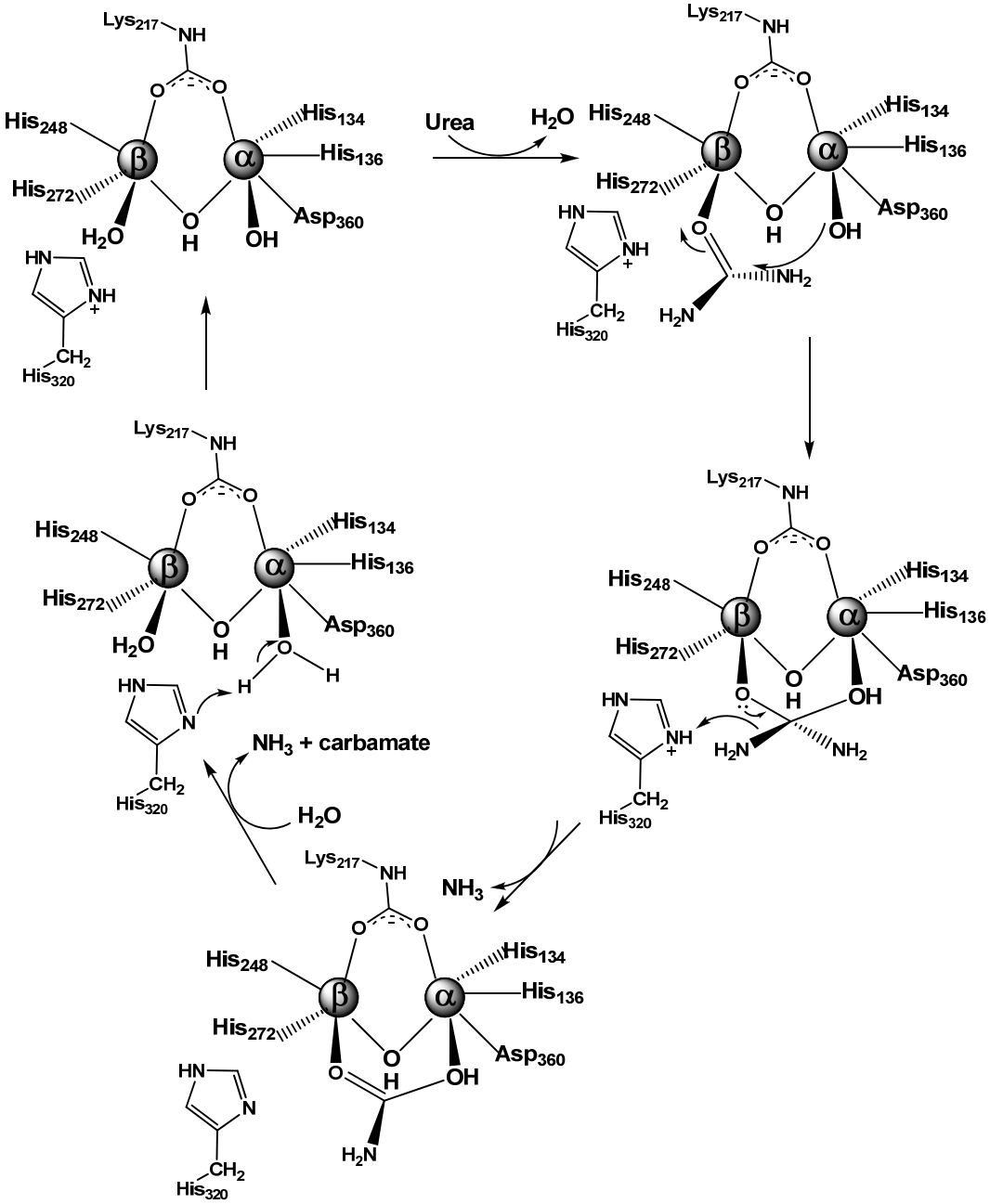


Scheme 1.3: The three mechanisms proposed for phosphotriesterase (*83*). A) A slightly modified version of the original mechanism proposed by Aubert et al. The bridging hydroxide nucleophilically attacks the phosphorus center of the substrate via a S_N2-like reaction. The bond to the phenol breaks and the resulting diethyl phosphate is bound within the active site as a bridging ligand between the two divalent cations. B) The Gao product complex where the diethyl phosphate is coordinated to the β-metal ion in a monodentate manner. C) The mechanism proposed by Ollis et al. The bridging hydroxide acts as a base that abstracts a proton from the nucleophilic M_α-bound water molecule. This deprotonation activates the attack on the phosphorus center of the substrate.

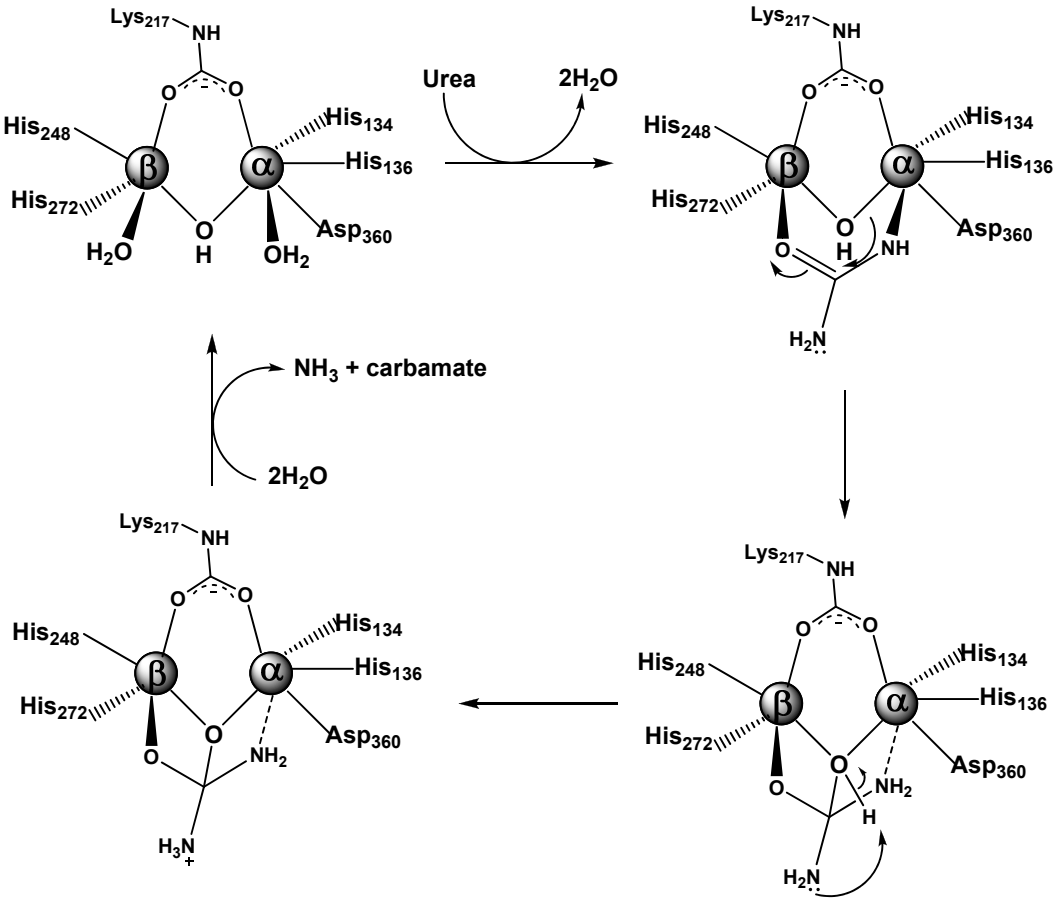


Another binuclear enzyme in which the mechanism has also been under great controversial debate is urease. There are three main reaction pathways that have been proposed for the enzyme-catalyzed decomposition of urea: the hydrolytic mechanism that involves the formation of a carbamate intermediate, the carbon dioxide mechanism in which urea is directly hydrolyzed to carbon dioxide and ammonia, and the elimination mechanism that involves the formation of a cyanate intermediate (*84*). The most commonly accepted mechanism, the carbamate mechanism, has been extracted from studies of bacterial urease from *K. aerogenes* and *B. pasteurii*. Two different variations of the carbamate mechanism have been proposed: the Karplus mechanism and the Benini mechanism. In the Karplus and Hausinger mechanism, urea binds to the active site and replaces a water molecule that is coordinating to M_β, while the water molecule (hydroxide) in the M_α site is retained. This Ni_α-bound OH acts as a nucleophile and attacks the carbonyl carbon of the urea molecule, which is further polarized by the Ni center. The reaction proceeds through a tetrahedral intermediate to form carbamic acid after the release of ammonia, which is assisted by proton transfer from the protonated His320 (*85*). A detailed illustration of this mechanism is shown in **Scheme 1.4**. An alternate mechanism for urease has been proposed by Benini et al, shown in **Scheme 1.5**. In this mechanism, urea replaces the two water molecules in the M_α and M_β sites and binds to the Ni active site center in a bidentate manner. The bridging hydroxide then acts as a nucleophile and attacks the urea carbonyl carbon atom, at the same time provides the proton that facilitates the release of ammonia (*32*).

Scheme 1.4: The mechanism of urease proposed by Karplus and Hausinger.

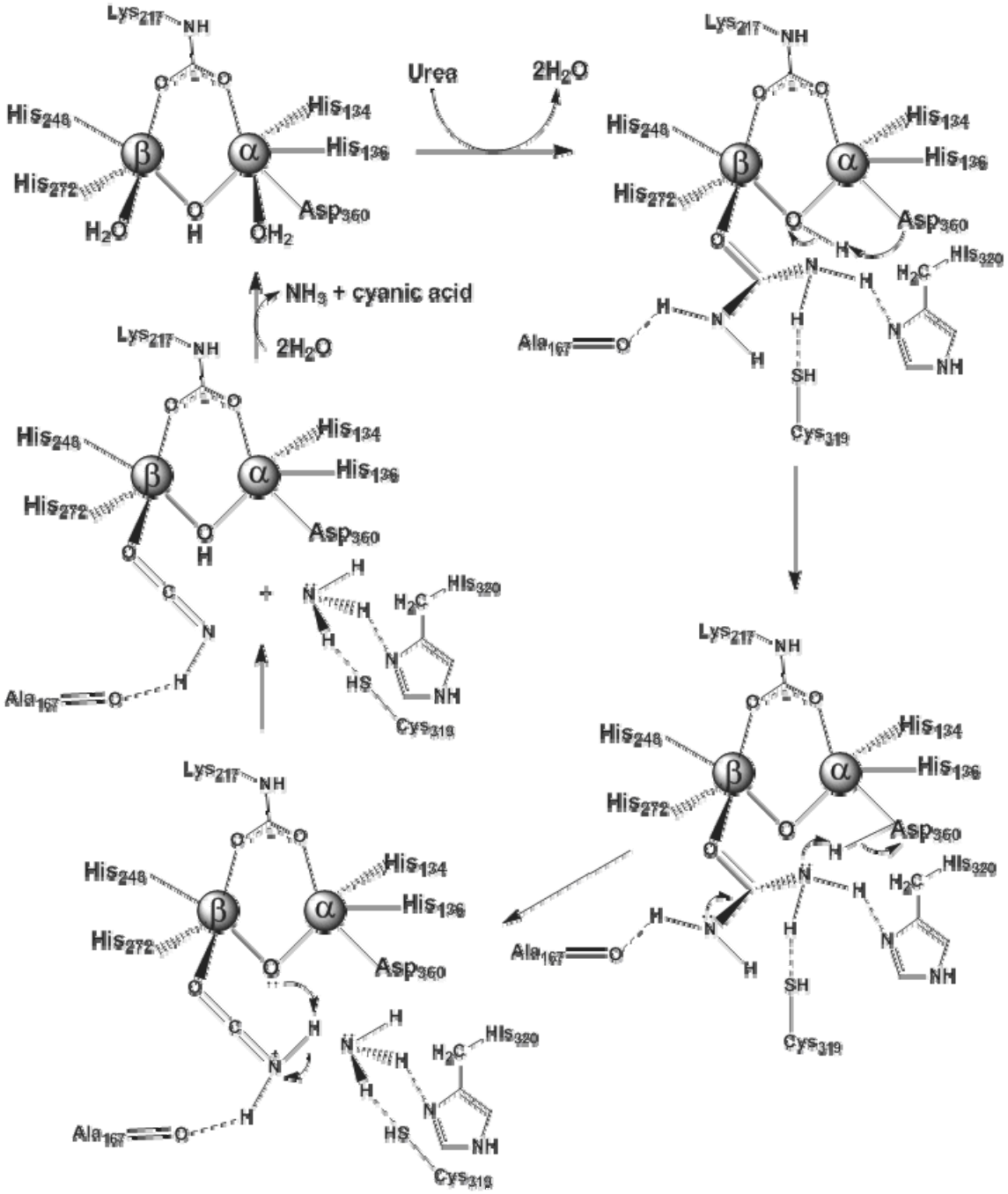


Scheme 1.5: The hydrolytic mechanism for urease proposed by Benini et al.



Computational methods have postulated the feasibility of a cyanic acid mechanism of urea decomposition, and this mechanism is also supported by kinetic studies of biomimetic complexes. The elimination mechanism is referred to as a "protein-assisted proton transfer" mechanism due to the major role of the protein in helping with the transfer of a proton between both NH_2 ends of urea. The first proposed protein-assisted proton transfer mechanism centers on Asp221 as the main residue relevant for catalysis. The second protein-assisted proton transfer, which can be described as "Asp-mediated proton transfer," involves the transfer of a proton from the bridging hydroxide to an NH_2 of urea. This process is facilitated by an active site Asp residue. The reactive nucleophilic di-Ni-coordinated oxygen is then capable of attacking the carbonyl carbon of urea to form a tetrahedral intermediate, as seen with the hydrolytic pathway. The result of this quantum mechanical analysis suggests that the active site of urease can undergo competitive elimination and hydrolytic mechanisms, leading to isocyanic acid or carbamic acid, respectively (*86*). A schematic example of a protein-assisted proton transfer mechanism is shown in **Scheme 1.6.**

Scheme 1.6: Schematic representation of a protein-assisted proton transfer mechanism in the elimination pathway of urea degradation.

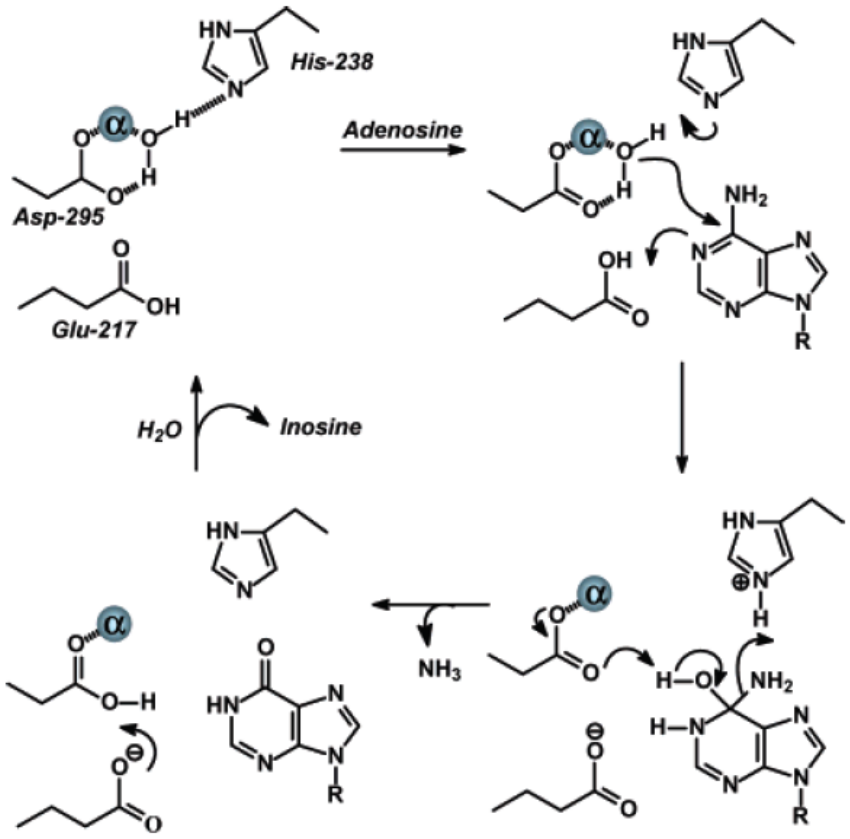


One other example of a binuclear metal center mechanism within the AHS is the mechanism proposed for the enzyme human renal dipeptidase (hRDP). The hRDP metal center is distinctly different from the binuclear centers described previously in that the metal ligands HxH from β-strand 1 is now HxD, and the conserved aspartate from strand 8 is no longer coordinating to the α-metal ion. Based on a crystal structure of the enzyme-cilastatin complex, Nitanai and colleagues suggested that the hydrolysis is initiated by a proton abstraction from the bridging water by the aspartate from strand 8. This deprotonation facilitates a nucleophilic attack on the carbonyl carbon of the amide bond by water, forming a tetrahedral adduct. The resulting oxyanion of the adduct is stabilized by the positive charge of the conserved His152 (*40*). This mechanism of intermediate stabilization is distinctive from other mechanisms proposed for enzymes of the AHS that have a binuclear metal center.

The reaction mechanisms for mononuclear enzymes of the AHS have been investigated to a great extent for adenosine deaminase (*46,56*) and *N*-acetyl-D-glucosamine-6-phosphate deacetylase (*70*). The reaction catalyzed by adenosine deaminase is an example of a nucleophilic aromatic substitution and the active site of ADA is a subtype of the mononuclear center where the metal ion is in the M_α position. In the mechanism proposed for ADA as outlined in **Scheme 1.7** (*31*), the histidine from strand 6 acts as a general base and abstracts the proton from the metal-bound water molecule. The resulting hydroxide attacks the aromatic ring forming a tetrahedral intermediate followed by a proton donation from a glutamate from strand 5 in a HxxE motif. Deprotonation of the carbinol intermediate by the conserved aspartate from

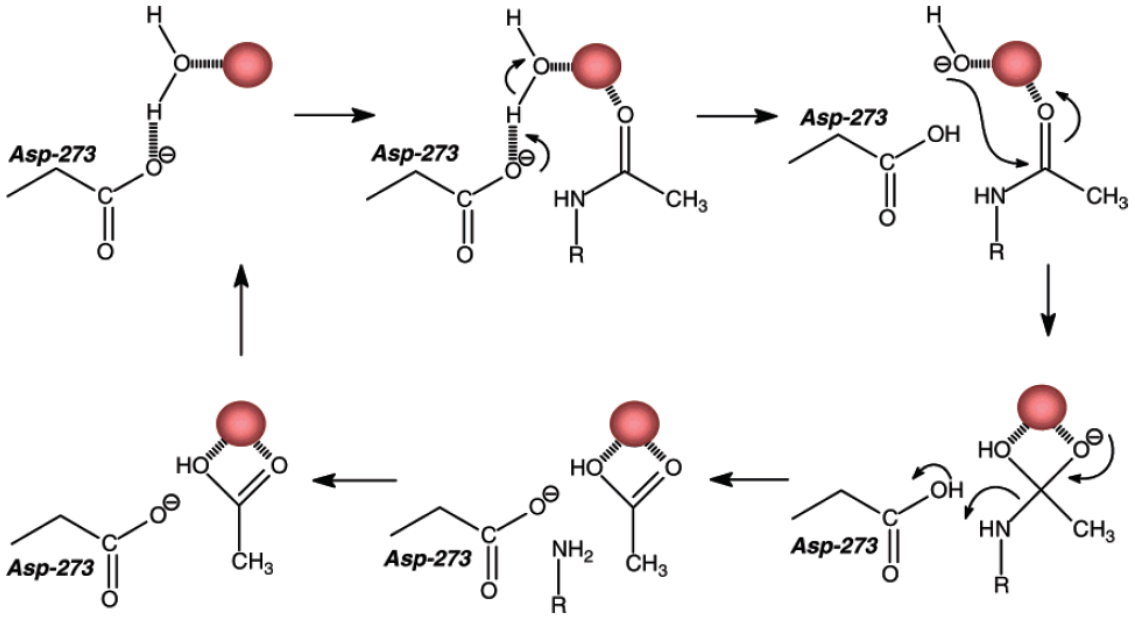
strand 8 and proton donation of the histidine from strand 6 assist the formation and release of the product (*46*).

Scheme 1.7: Proposed mechanism for the hydrolysis of adenine by ADA (*31*).



The structure and mechanism of AGD from *E. coli* was elucidated recently (*42*). AGD from *E. coli* catalyzes the deacetylation of *N*-acetyl-D-glucosamine-6-phosphate using a single metal ion positioned at the M_β site. The mechanism of AGD (**Scheme 1.8**), proposed by Hall et al (*70*), is initiated by the proton abstraction of the active site hydrolytic water (which is ligated to the metal ion) by the invariant aspartate from strand 8. The resulting hydroxide nucleophilicly attacks the carbonyl carbon of the substrate that is being polarized by the metal ion, forming a tetrahedral intermediate. The collapse of the tetrahedral intermediate is followed by a proton transfer from the aspartate of β-strand 8 to the amine leaving group of the product and ultimately the release of products. Apparently the mechanism utilized by NagA from *Bacillus subtilis* is different than that of the *E. coli* enzyme since it contains a binuclear metal center (*41*). These features highlight the significant diversity for the evolution of function within the amidohydrolase superfamily.

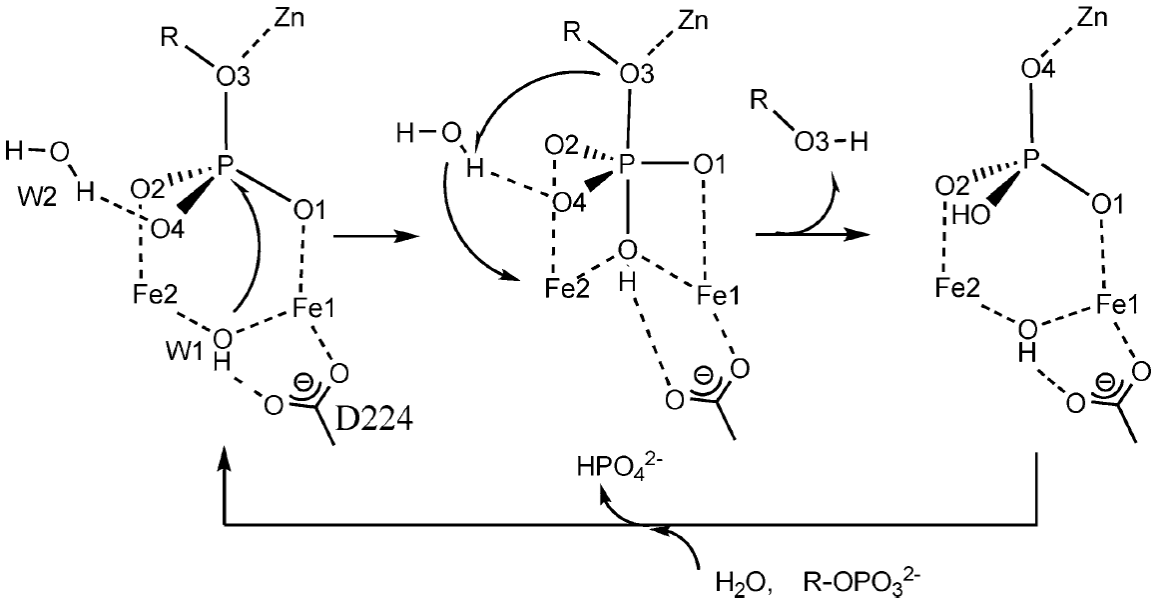
Scheme 1.8: Proposed mechanism for the hydrolysis of *N*-acetyl-glucosamine-6-phosphate deacetylase, an enzyme with a mononuclear metal center at the β-site (*70*).



The reaction mechanism for one of the most diverged members of the AHS, HolPase, has been proposed by Omi et al based on the crystal structure of the enzyme and an enzyme-substrate complex model (**Scheme 1.9**) (*53*). According to the mechanism, the substrate displaces a water molecule within the trinuclear cluster, and the substrate phosphate coordinates to the three metals. These interactions enhance the electrophilicity of the phosphorus atom. The bridging hydroxide acts as a nucleophile and makes an attack on the phosphorus center to form a trigonal pyramidal transition state, which is stabilized through the maintenance of the interactions of the phosphate and the metal center. This attack is facilitated by interactions between the hydroxide and

the two irons as well as between the hydroxide and Asp224. The geometrical constraint of the transition state makes the phosphorus atom move slightly toward the two irons. The P-O3 bond, which is being polarized by the Zn ion, loosens and the O3 gets protonated by a water molecule (W2 in **Scheme 1.9**). The histidinol is released from the active site, and another substrate enters to liberate the phosphate ion and start another round of catalysis.

Scheme 1.9: Reaction mechanism of histidinol-phosphate phosphatase from *T. thermophilus* (*53*). The formation and stabilization of a trigonal bipyramidal transition state is facilitated by interactions between the phosphate group of the substrate and the trinuclear metal center.

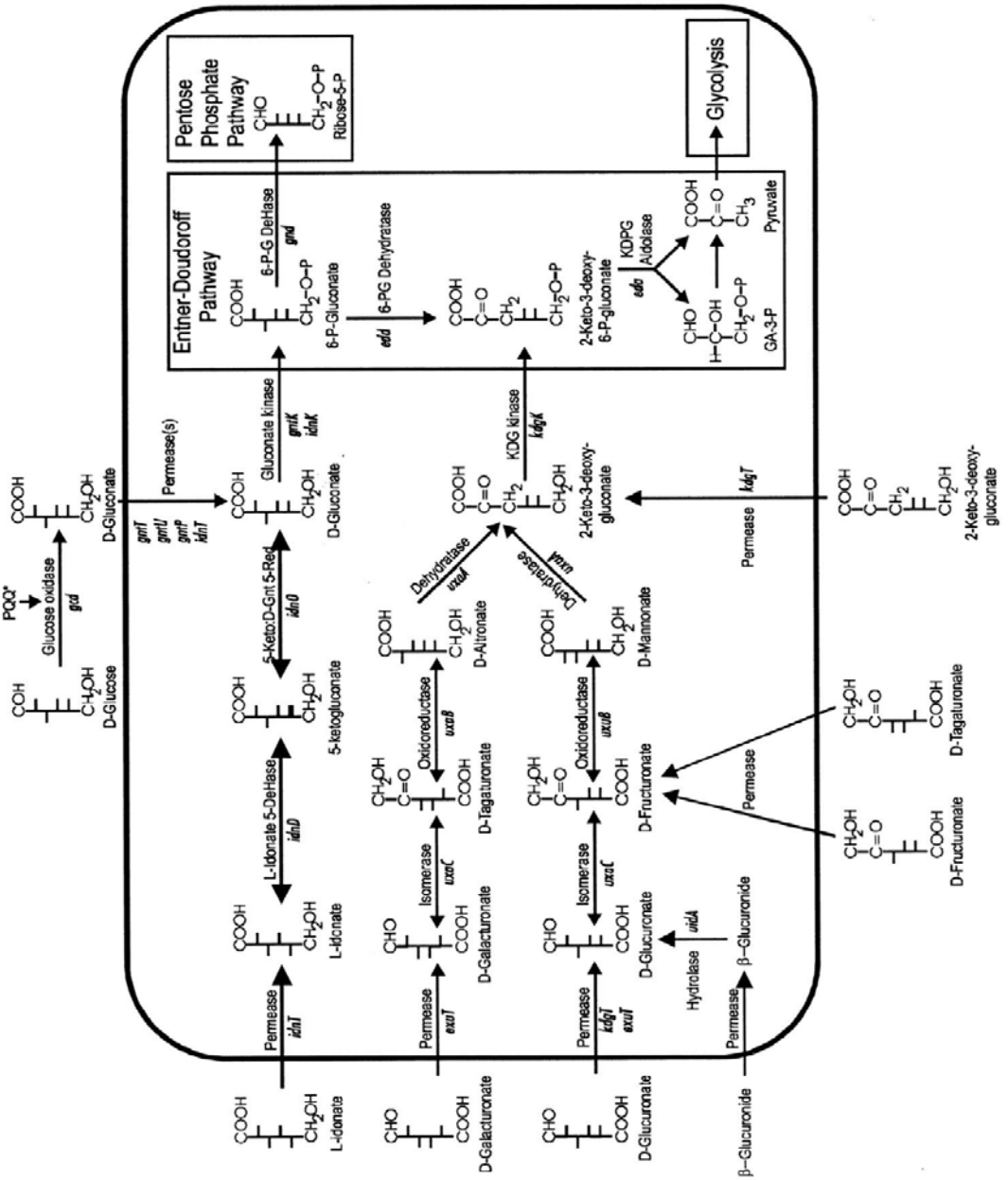


In this dissertation, the mechanisms of two members of the amidohydrolase superfamily were characterized and examined in detail using kinetics and X-ray crystallography. In the first part of the dissertation, the functional discovery of a second uronate isomerase (galacturonate isomerase) from *Bacillus halodurans* will be addressed. Unlike *E. coli* and *T. maritima*, the genomic context of the organism *B. halodurans* suggests that the pathways for the utilization of D-glucuronate and D-galacturonate are organized separately and involved separate enzymes for the isomerization of the respective hexauronate, the first reaction in the pathway. The X-ray crystal structure of galacturonate isomerase has also been solved, by the efforts of our collaborator Dr. S. Almo at the Albert Einstein College of Medicine, in the presence of substrate and intermediate mimics for the elucidation of the isomerization reaction mechanism utilized by enzymes of this family. The second part of the dissertation focuses on the mechanistic discussion of the enzyme uronate isomerase (URI) from *E. coli*. This enzyme is a unique member of the amidohydrolase superfamily due to its rare nonhydrolytic reactivity. Therefore characterization of this enzyme offers a striking example for the divergence of chemical reactivity of enzymes that share a common structural ancestor. The last part of this dissertation will be focused on the mechanistic details of a renal dipeptidase-like protein from *Streptomyces coelicolor*. The mechanism

for the hydrolysis of dipeptides by this enzyme will be presented, and kinetic and structural characterizations will be used for the corroboration of the proposed mechanism.

Uronate isomerase is considered an amidohydrolase enzyme based on sequence alignment. A crystal structure of the *T. maritima* URI has been obtained by the Joint Center for Structural Genomics (*43*). However, the structure contains a water molecule in place of the putative metal ion in the active site. The metal is predicted to be a zinc ion that is coordinating to the two histidines from β-strand 1 and the aspartate from strand 8. A sequence alignment of uronate isomerases from 36 different organisms reveals that uronate isomerases are all missing the histidine from strand 6. This enzyme represents the only well-characterized AH enzyme that does not contain a histidine at this location. Instead, it contains a conserved tryptophan at the end of strand seven. This enzyme is also unusual in that it is one of the enzymes from the AH superfamily that does not catalyze a hydrolysis reaction (*31*).

Uronate isomerase facilitates the conversion of D-glucuronate to D-fructuronate and of D-galacturonate to D-tagaturonate, the first step in the catabolism of hexuronate in *E. coli* (*87*). The parallel pathways for the break down of D-glucuronic and D-galacturonic acids converge with the formation of 2-keto-3-deoxy-gluconate. This intermediate is then phosphorylated and enters the Entner-Doudoroff pathway where it is further metabolized into glyceraldehyde-3-phosphate and pyruvate, two important metabolites of glycolysis. The pathways for the catabolism of hexuronate are outlined in **Scheme 1.10**. It was determined that the 72 *E. coli* strains isolated all possess the Entner-Doudoroff pathway and are capable of utilizing gluconate and glucuronate as energy sources (*87*). In humans, D-glucuronate is used for drug and toxin conjugation and is excreted in the urine and bile (*88*). D-glucuronate and D-galacturonate are found in high concentration in the intestinal tract and are believed to be substrates to support the growth of *E. coli* in the gut (*87*). Hence the activity of uronate isomerase in *E. coli* is essential for its survival in intestinal invironments.

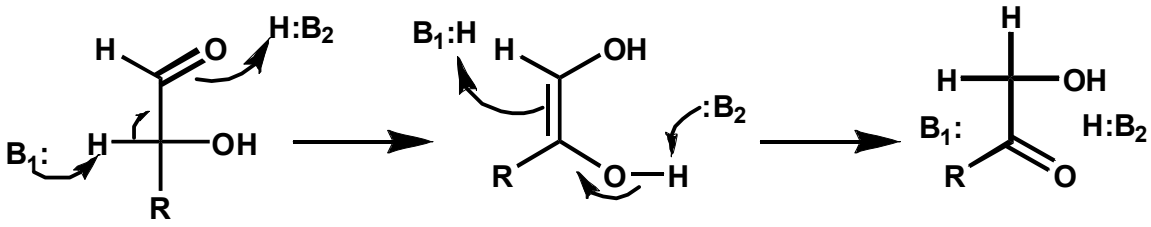


Scheme 1.10: Catabolic pathway of hexuronate in *E. coli* (87).

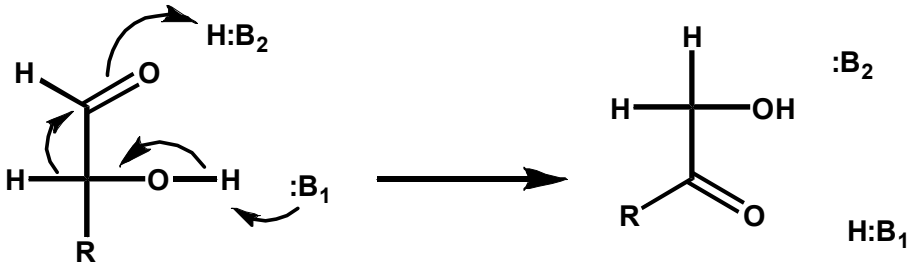
The reaction catalyzed by URI is that of the aldose-ketose isomerization reaction via a 1,2-hydrogen transfer. These enzymatic reactions generally proceed either through a proton transfer mechanism (**Scheme 1.11A**) such as phosphoglucose isomerase of the aldose-ketose isomerase family of enzymes (*89*) and triosephosphate isomerase (*90-94*) or a less common hydride transfer mechanism (**Scheme 1.11B**) as proposed for the metalloenzyme xylose isomerase (*95-97*). In the proton transfer mechanism of triosephosphate isomerase (TIM), Glu-165 abstracts the *pro-R* proton from C-1 of dihydroxyacetone phosphate (DHAP) and transfers it to the C-2 of glyceraldehyde 3-phosphate, while His-95 catalyzes the proton transfer between O1 and O2, as suggested by X-ray crystallography and mutagenesis studies (*90-94*). The mechanism of xylose isomerase involves two divalent cations that mediate the hydride transfer reaction. Evidence suggests that the enzyme binds the α form of the pyranose ring, and His-53 is involved in the protonation of O5 for ring opening. Following the ring-opening step, O2 is deprotonated by a metal-bound hydroxide ion accompanied by a shift in the position of the second metal (M2) towards the substrate. At a liganding distance to the substrate, M2 acts as an electron withdrawing force on the C1-O bond of substrate. This polarization results in a partial positive character on C-1 and thereby a more readily hydride transfer (*95-97*). Recently, the mechanism of URI from *E. coli* has been established as a proton transfer process based on the detection of substrate/solvent hydrogen exchange upon incubation with enzyme. The rate of deuterium incorporation at C2 of D-glucuronate was determined to be 0.018 s^{-1} compared to 196 s^{-1} for product formation (*98*).

Scheme 1.11: Aldo-keto isomerization of sugars. **A**) Proton transfer mechanism. **B**) Hydride transfer mechanism.

A



B



The isomerization of D-glucuronate to D-fructuronate requires a proton transfer from C2 of D-glucuronate to C1 at either the *re*- or *si*-face of D-fructuronate. The stereospecificity of this transfer was established by utilizing the enzymes D-mannonate dehydrogenase and D-mannonate dehydratase. The deuterium transferred from C2 of D-glucuronate is ultimately found at C6 of 2-keto-3-deoxy-D-gluconic acid. The detection of the deuterium at the *pro*-R position of 2-keto-3-deoxy-D-gluconic acid by NMR spectroscopy analyses confirmed that the proton is transferred to the *pro*-R position at

C1 of D-fructuronate (*98*). In chapter II of this dissertation, characterization of the mechanism of URI from *E. coli* will be discussed in detail using enzyme kinetics.

In the pursuit of protein function for unknown members of the AHS, one of the proteins in *B. halodurans* with an uncertain function is Bh0493. The sequence of Bh0493 is one of the most divergent protein sequences identified to date. A search of the DNA sequences available from NCBI reveals that the most closely related protein with a characterized catalytic function is Tm0064 from *Thermotoga maritima.* This protein is a uronate isomerase and a crystal structure of this enzyme has been deposited in the PDB (1J5S). However, the overall amino acid sequence identity between Bh0493 and Tm0064 is rather low. When these two proteins are compared to one another the sequence identity is less than 19%. Nevertheless, a sequence alignment between Bh0493 and other enzymes known to be uronate isomerases reveals that the residues suspected to be critical for the functioning of the isomerization reaction are conserved. These residues include the conserved HxH motif from β-strand 1, the aspartate from β-strand 8, and a WWF motif from β-strand 7. In addition, residues that are critical for functioning in the *E. coli* enzyme are conserved in the sequence of Bh0493 based on the sequence alignment of Bh0493 with other authentic uronate isomerases. In this part of the dissertation, cloning, expression, structure and functional characterization of Bh0493 will be reported. The X-ray structures of Bh0493 in the presence of substrate and intermediate mimics will aid in the elucidation of the isomerization mechanism employed by the uronate isomerase family of enzymes. The gene context within the genome of *B. halodurans* will also be examined to further obtain evidence for the

support of functional assignment of a gene encoding an enzyme in the catabolic pathways of hexuronate.

The third enzyme that will be characterized in this dissertation is a renal dipeptidase-like protein (Sco3058) from *Streptomyces coelicolor*. *Streptomyces* are recognized for their production of biologically active secondary metabolites and are responsible for almost two-thirds of all known natural antibiotics. The best genetically characterized strain is *Streptomyces coelicolor A3(2)*, which has become the preferred model organism in Streptomyces research (*99*).

The renal dipeptidase family of enzyme was first identified and characterized in the 1970s based on its physicochemical properties (*100*). In addition to a wide range of dipeptides, renal dipeptidase is also involved in the renal metabolism of glutathione and its conjugates and is the sole enzyme that is responsible for the hydrolysis of penem and carbapenem β-lactam antibiotics (*101-106*). The first crystal structure from the renal dipeptidase family is the enzyme from human (PDB: litq) (*40*). The quarternary structure of the hrDP is a homodimer with each polypeptide subunit consisting of 369 amino acid residues (42 kDa). The sequence of the hrDP shows a 75% identity to its mammalian homologs. From the crystal structure, the protein utilizes a HxD motif at strand 1 to bind to the M_{α} zinc ion instead of the conventional HxH. Moreover, the residue that bridges the two metals is a glutamate from β-strand 3 and not strand 4. The aspartate at the end of β-strand 8 is conserved, but it is no longer coordinating to the M_{α}. According to the X-ray structure, this aspartate residue appears to be interacting with the hydrolytic water via hydrogen bonding. This enzyme was also cocrystallized in the

presence of a known dipeptidyl inhibitor cilastatin (PDB entry code 1itu). The cilastatin-bound active site (**Figure 1.3**) suggested, as previously mentioned, that the hydrolytic water is activated by the conserved aspartate from β-strand 8. The deprotonated water then nucleophilically attacks the carbonyl carbon of the peptide bond to form a tedrahedral intermediate. It was also proposed that the conserved positively charged His-152 then acts as an oxyanion hole to stabilize the resulting negatively charged intermediate. This mechanism of hydrolysis is distintly different from the mechanism of characterized enzymes in the AHS in which the tetrahedral intermediate is stabilized by the solvent exposed M_β (*40*).

The first renal dipeptidase-like gene *acdp* from a prokaryote was cloned from *Acinebacter calcoaceticus* (*107*). This enzyme has a low sequence identity of 23% to the human renal dipeptidase, but the essential residues identified in the mammalian dipeptidases are all conserved. Unlike its mammalian homologs, the gene product of *acdp* could not hydrolyze the unsaturated dipeptide glycyldehydrophenylalanine; instead the protein showed substrate specificity towards various dipeptides with a preference for dipeptides contaiting a D-amino acid at the C-terminus (*107*). The protein Sco3058, on the other hand, contains a 45% and 22% sequence identities to the hrDP and *acdp*, respectively. Substrate specificity analysis of this protein indicates that the enzyme is promiscuous towards a wide range of dipeptides containing L- and D-amino acids with a preference for L-Xaa D-Xaa substrates. In this dissertation, kinetic and structural properties of the enzyme Sco3058 from *Streptomyces coelicolor* will be discussed in detail.

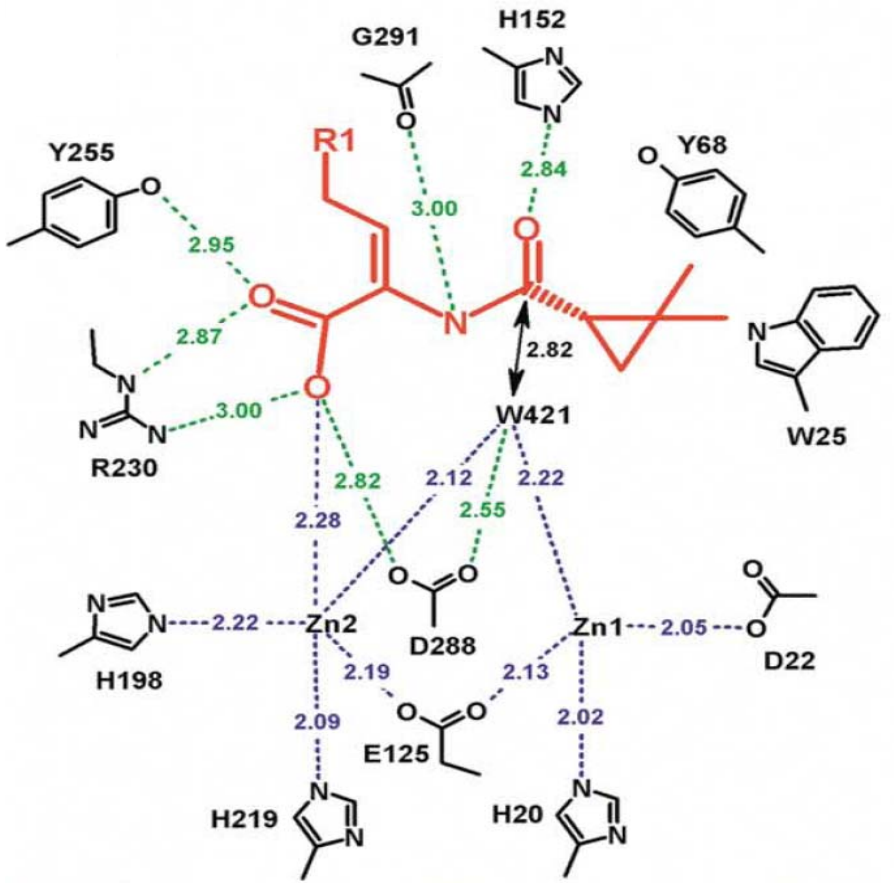


Figure 1.3: Active site of hRDP with bound cilastatin (in red) (*40*).

CHAPTER II

AT THE PERIPHERY OF THE AMIDOHYDROLASE SUPERFAMILY: BH0493 FROM *BACILLUS HADODURANS* CATALYZES THE ISOMERIZATION OF D-GALACTURONATE TO D-TAGATURONATE*

The identification and functional assignment of enzymes encoded within completely sequenced genomes is a difficult and demanding problem. According to some estimates, approximately 40% of newly sequenced genes have an *unknown*, *uncertain*, or *incorrect* functional assignment (*2*). The extent of the catalytic diversity inherent within metabolic and catabolic pathways will not be fully understood until the functional annotations of these newly sequenced genes have been addressed. The absence of a comprehensive understanding of the metabolic landscape has therefore motivated the development of new methodologies for the assignment of function for enzymes that catalyze reactions with ambiguous substrates. These emerging strategies include phylogenetic profiling 108, *109*), recognition of domain fusions and gene clusters (*87*,*110-113*), and other computational approaches (*88*). Most of the commonly used techniques require sequence and/or structural comparisons to well-characterized

*Reprinted with permission from "At the Periphery of the Amidohydrolase Superfamily: Bh0493 from Bacillus Halodurans Catalyzes the Isomerization of D-galacturonate to D-tagaturonate" by T. T. Nguyen, S. Brown, A. A. Fedorov, E. V. Fedorov, P. C. Babbitt, S. C. Almo, and F. M. Raushel, 2008. *Biochemistry*, 47, 1194-1206, Copyright 2008 by Journal of the American Chemical Society.

homologues to make high quality functional predictions. However, these attempts at functional characterization can be problematic when the sequence or structural similarity is remote (*2*).

Members of the amidohydrolase superfamily (AHS) have been identified in every organism sequenced to date (*20*, *31*). This enzyme superfamily catalyzes predominantly hydrolytic reactions where a water molecule is activated by one or two divalent metal ions embedded within the active site (*31*). All of the enzymes within the AHS adopt a $(\beta/\alpha)_8$-barrel structural fold with an active site that resides at the C-terminal end of the β-barrel (*21*, *31*). One of the most divergent members of the amidohydrolase superfamily is uronate isomerase (URI). This enzyme family exhibits few statistically significant sequence links to any other member of the AHS (**Figure 2.1**), and structure-based comparisons are required to link this family to other families in the superfamily.

The URI from *Escherichia coli* (UxaC) catalyzes the isomerization of D-glucuronate to D-fructuronate and of D-galacturonate to D-tagaturonate. These isomerization reactions represent the first step in the separate metabolic pathways for the utilization of D-glucuronate and D-galacturonate in bacteria. The isomerized products are subsequently reduced to D-mannonate and D-altronate by UxuB and UxaB, respectively, and then dehydrated by UxuA and UxaA to 2-keto-3-deoxy-D-gluconate. The entire metabolic pathway for the metabolism of these sugars in *E. coli* is summarized in **Scheme 2.1**.

We have begun to delineate the extent of the structural and functional diversity inherent within the amidohydrolase superfamily of enzymes. One of the most divergent

protein sequences identified to date from this investigation has emerged from the genome of the alkaliphilic bacterium *Bacillus halodurans* (gi:15613056). This protein (Bh0493) was tentatively assigned as a member of the AHS based upon a weak sequence similarity (<19%) with the structurally characterized uronate isomerase from *Thermotoga maritima* (Tm0064, gi: 15642839). Providing additional evidence for this assignment, the putative active site residues that originate from the ends of β-strands 1 and 8 are apparently conserved in the *B. halodurans* enzyme. In this chapter, we report the cloning, expression, structure and functional characterization of Bh0493. The X-ray structure of Bh0493, determined with a single zinc ion in the active site, confirms that this enzyme is a member of the AHS. This enzyme will catalyze the isomerization of D-galacturonate to D-tagaturonate in a metabolic pathway that has diverged from the isomerization of D-glucuronate. We have also functionally characterized Bh0705 from B. halodurans (gi: 15613268), the amino acid sequence that more closely resembles that of a canonical uronate isomerase. Crystal structures of Bh0493 in the presence substrate and intermediate mimics will aid in the elucidation of the isomerization mechanism utilized by this family of enzyme.

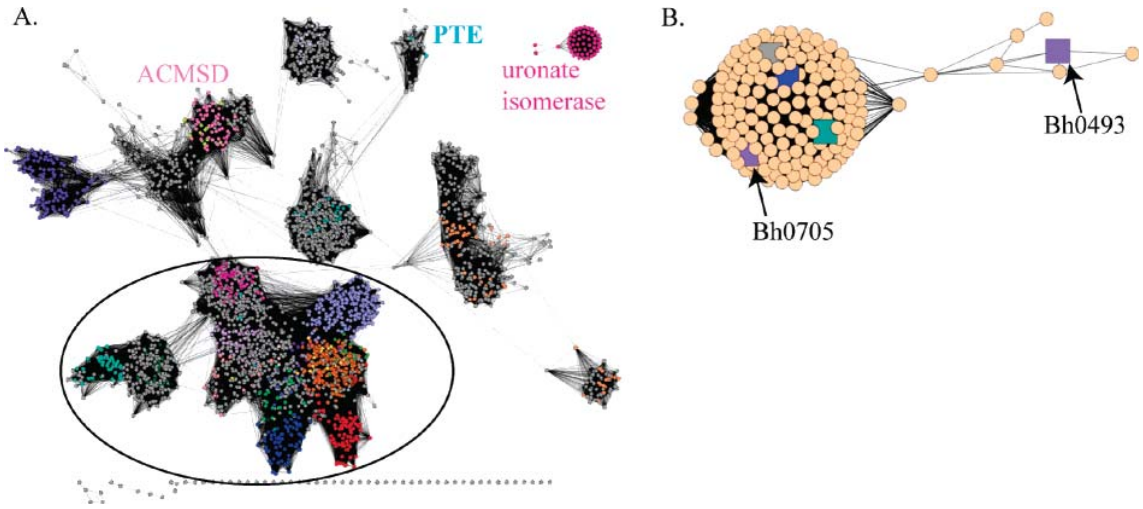
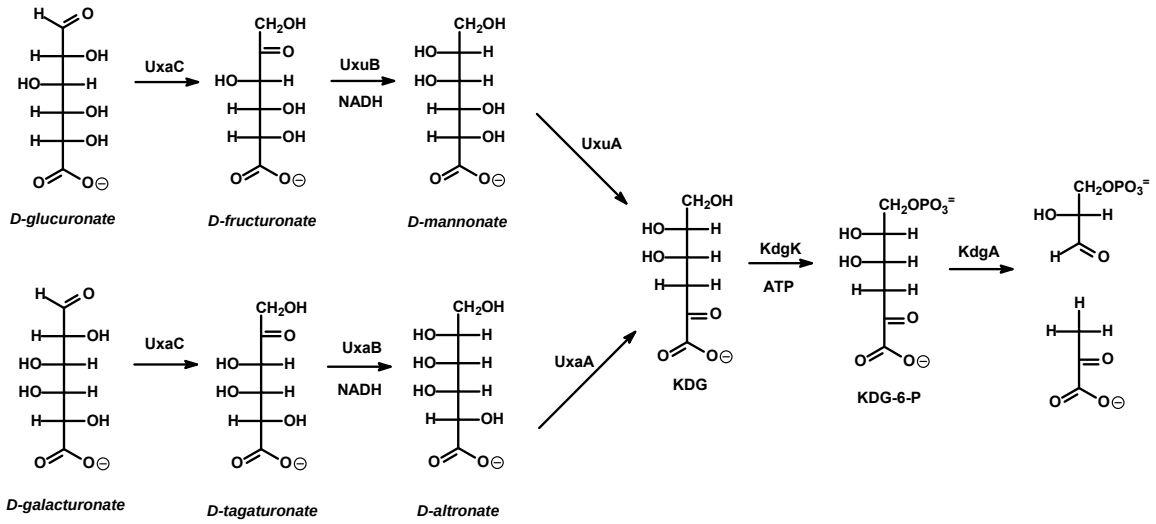


Figure 2.1: Network representation of the sequence relationships in the AHS. Each node in the network represents a single sequence in the uronate isomerase-like sequence set, and each edge represents the pairwise connection between two sequences with the most significant BLAST *E*-value (better than the cutoff) connecting the two sequences. Lengths of edges are not meaningful except that sequences in tightly clustered groups are relatively more similar to each other than sequences with few connections. (A) An 80% ID filtered set of AHS sequences from the SFLD showing connections with a BLAST *E*-value more significant than or equal to 10^{-4}. Sequences are colored by SFLD subgroup. Sequences with characterized functions originally identified as part of the AHS by Holm and Sander are contained within the circle (except for phosphotriesterase, (PTE)). Additional groups discussed in the text are indicated with labels. (B) An unfiltered set of uronate isomerase-like sequences at a BLAST *E*-value more significant than or equal to 10^{-10}. Sequences that have been experimentally characterized as uronate isomerase and/or have an X-ray crystal structure are colored as follows: gray, Tm0064; lavender, Bh0705 and Bh0493; blue Cc1490; and green, b3092.

Scheme 2.1: Metabolic pathways for the metabolism of D-glucuronate and D-galacturonate.



MATERIALS AND METHODS

Materials. D-Glucuronate and NADH were purchased from Sigma-Aldrich. D-galacturonate was obtained from Acros. The inhibitors, D-arabinaric acid (**1**) and the monohydroxamate derivative of D-arabinaric acid (**2**) were synthesized by Dr. Yingchun Li. The DNA sequencing reactions were conducted by the Gene Technologies Laboratory at Texas A&M University. Determination of the metal content of the purified proteins was done using ICP-MS. Gel filtration and anion exchange columns were purchased from GE HealthCare. Restriction enzymes and T4 DNA ligase were obtained from New England Biolabs. The pET30 vector was acquired from Novagen and the Platinum Pfx DNA polymerase was purchased from Invitrogen. The Wizard

Miniprep DNA purification kit was obtained from Promega and the genomic DNA of *B. halodurans* was acquired from ATCC (BAA-125D).

Cloning and Overexpression of Bh0705 and Bh0493. The gene *uxa*C (Bh0705, gi: 15613268) from *B. halodurans* was amplified by PCR with two primers, 5'-GCAG-GAGCATTAATGACGAATTTTCTTTCAGAAGACTTCCTCTTAATGAACGAAT-ACG-3'containing an *Ase*I restriction site and 5'-GCGCAAGCTTTCATCATAGTCG-AACAGTCTCCTTCATGTCG-3' containing a *Hin*dIII restriction site. The *uxa*C gene was inserted into a pET30a (+) vector that had been digested with *Ase*I and *Hin*dIII. The protein was expressed and purified from BL21(DE3) star cells (Novagen) that had been transformed with the pET30a (+) plasmid containing the *uxa*C gene from *B. halodurans*. The gene for Bh0493 (gi: 15613056) was cloned with two primers, 5'-GCAGGAGCCA-TATGTCCATAAACAGTAGGGAAGTGTTACGGG-3' and 5'-CGCGGAATTCTTA-TTACGTTTGCTGCTCCACCTTCACTGATGTGACG-3' containing the restriction sites for *Nde*I and *Eco*RI, respectively. The gene product was ligated into a pET30a (+) vector and expressed in BL21(DE3) star cells. The same protocol was utilized to purify Bh0705 and Bh0493. A single colony of the cells containing the gene of interest was inoculated and incubated overnight in 5 mL of LB medium containing 50 μg/mL kanamycin at 37 ºC. The overnight culture was used to inoculate 1 L of LB medium containing 50 μg/mL kanamycin. The cells were grown at 30 ºC and induced with 0.4 mM IPTG when the cell density at A_{600} reached ~0.4-0.6. After 24 hours, the cells were harvested at 4,000*g*. The cell pellet was resuspended in 5 mL of 50 mM HEPES, pH 8.0, (buffer A) for every gram of cell paste and disrupted by sonication. The nucleic

acids were removed by the dropwise addition of protamine sulfate to a final concentration of 2% w/v and subsequent centrifugation at 14,000*g*. The protein was precipitated between 50-70% saturation of ammonium sulfate, and the protein pellet was resuspended in a minimum amount of buffer A. The protein was loaded onto a pre-equilibrated Superdex 200 gel filtration column and eluted with buffer A. The fractions containing the protein of interest were pooled based on activity, loaded onto a Resource Q column, and eluted with a 0-30% gradient of 1 M NaCl in 20 mM HEPES, pH 8.0 (buffer B). The fractions containing the desired protein were pooled and the purity of the protein was determined by SDS-gel electrophoresis. The purified enzyme was sterile-filtered and stored at 4º C.

Cloning, Expression, and Purification of UxaB. The *uxa*B gene from *E. coli.* (gi: 49176119) was amplified by PCR with two primers, 5'-GCAGGAGCCATATGAAA-ACACTAAATCGTCGCGATTTTCCCGGTGC-3' and 5'-GCGCAAGCTTTTATTA-GCACAACGGACGTACAGCTTCGCGCATCCCTTTTTCG-3', containing *Nde*I and *Hin*dIII sites, respectively, and the resulting fragment was inserted into the *Nde*I and *Hin*dIII sites of the pET30 vector. The protein was expressed in the BL21(DE3) strain of *E. coli*. A single colony was used to inoculate a 5-mL overnight culture of LB medium containing 50 μg/mL kanamycin. The 5 mL overnight culture was inoculated into 1 L of LB containing the same concentration of kanamycin. The cells were grown at 30 ºC, induced with 0.4 mM IPTG at an A_{600} ~0.6, and incubated overnight. The cells were collected by centrifugation at 4,000*g*, and the cell pellet was resuspended in buffer

A. The cells were disrupted by sonication and centrifuged at 14,000*g*. The protein was soluble and the supernatant solution was used directly in the enzymatic assays.

Cloning, Expression, and Purification UxuB. The *uxu*B gene (gi: 26111644) from *E. coli* that encodes D-mannonate dehydrogenase (MDH) was cloned into a pET30 expression vector. The gene was transformed into BL21(DE3) and the cells grown in TB broth at 37 °C. D-Mannonate dehydrogenase was purified using the same methods previously described for the isolation of uronate isomerase, but no metal was added to the purification buffers. The cell extract was made 40-60% of saturation with ammonium sulfate, and the resuspended pellet was subjected to fractionation with the aid of a Superdex 200 gel filtration column. Fractions containing D-mannonate dehydrogenase were pooled, quick-frozen, and stored at -80 °C.

Sequence Analysis. The uronate isomerase sequence clusters were organized and created in collaboration with Patricia C. Babbitt at the School of Pharmacy, University of California, San Francisco, CA.

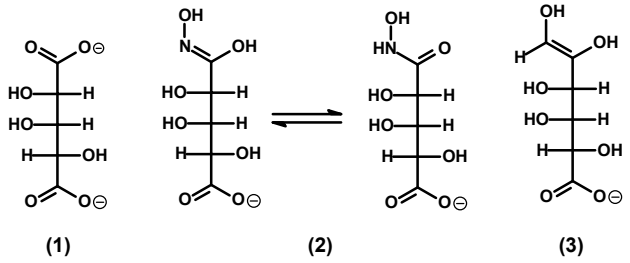
Crystallization of Bh0493. The structure of Bh0493 was solved in collaboration with the Steven C. Almo group at the Department of Biochemistry, Albert Einstein College of Medicine, at Bronx, New York. The PDB entry codes are 2q6e for the Se-Met-substituted enzyme and 2qo8 for the wild-type complexed with Zn.

Enzyme Activity and Substrate Specificity. The isomerase activities of Bh0705 and Bh0493 were measured with D-glucuronate and D-galacturonate. The enzymatic reactions with D-glucuronate and D-galacturonate were coupled to the reduction of D-fructuronate and D-tagaturonate with the appropriate dehydrogenase. These assays

monitor the oxidation of NADH to NAD^+ and the change in absorbance was measured spectrophotometrically at 340 nm. The standard assays contained 50 mM HEPES, pH 8.0, 0.2 mM NADH, ~ 2 μM of either D-mannonate dehydrogenase or D-tagaturonate dehydrogenase, and various substrate concentrations in a final volume of 250 μL.

Inhibition Studies. The inhibition of the reactions catalyzed by Bh0705 and Bh0493 was quantified using two compounds that have previously been shown to inhibit the isomerization of D-glucuronate by uronate isomerase from *E. coli*. D-Arabinaric acid (**1**) and (2S,3R,4S)-4-(hydroxycarbamoyl)-2,3,4-trihydroxybutanoate (**2**) were tested as inhibitors using standard assay conditions with D-glucuronate as the substrate and D-mannonate dehydrogenase as the coupling enzyme. The structures of these compounds are presented in **Scheme 2.2**.

Scheme 2.2: Structures of D-arabinarate (1), hydroxamate derivative (2) and *cis*-enediol intermediate.



Data Analysis. The kinetic parameters, k_{cat} and k_{cat}/K_m, for the enzyme Bh0705 were determined by fitting the initial velocity data to equation 2.1, where v is the initial velocity, E_t is the total enzyme concentration, k_{cat} is the turnover number, A is the substrate concentration, and K_a is the Michaelis constant. The double reciprocal plots for Bh0493 were nonlinear (concave down) and fitted to equation 2.2 (*114*). In this equation k_1 is the maximum velocity at low substrate concentration and the sum of k_1 and k_2 is the maximum velocity at high substrate concentration. The apparent Michaelis constants at low and high substrate concentrations are K_1 and K_2, respectively. Competitive inhibition patterns by the compounds that mimic the putative *cis*-enediol intermediate were fit to equation 2.3, where I is the inhibitor concentration and K_i is the slope inhibition constant.

$$v / E_t = k_{cat}[A] / (K_a + [A]) \quad (2.1)$$

$$v / E_t = (k_1[A] / (K_1 + [A])) + (k_2[A]) / (K_2 + [A])) \quad (2.2)$$

$$v / E_t = (k_{cat}[A]) / (K_a(1 + (I / K_i)) + [A]) \quad (2.3)$$

RESULTS

Bioinformatics. As of July 2009, approximately 190 uronate isomerase like sequences were identified in the NCBI nr database. All, except two archaeal sequences, are from bacteria, reflecting their role in bacterial glucuronic and galacturonic acid metabolism. As shown in **Figure 2.2**, these sequences cluster into three groups. Group 1 contains one of the two functionally characterized uronate isomerases from *B. halodurans* (Bh0705). Group 2 contains the structurally characterized uronate isomerase from *C. crescentus*. Group 3 contains the second functionally characterized uronate isomerase from *B. halodurans* (Bh0493).

Purification and Properties of Bh0705. The gene for the enzyme that encodes for *Bh0705* was expressed in BL21(DE3) cells. The protein Bh0705 was sparingly soluble after the cells were disrupted by sonication. Approximately 14 mg of purified protein was obtained from 14 grams of wet cell paste. The metal content of the purified protein was determined to contain 0.9 equivalents of Zn^{2+} per subunit using ICP-MS. SDS-PAGE analysis revealed the presence of a single band at approximately 55 kDa for Bh0705. This value is in agreement with the reported gene sequence.

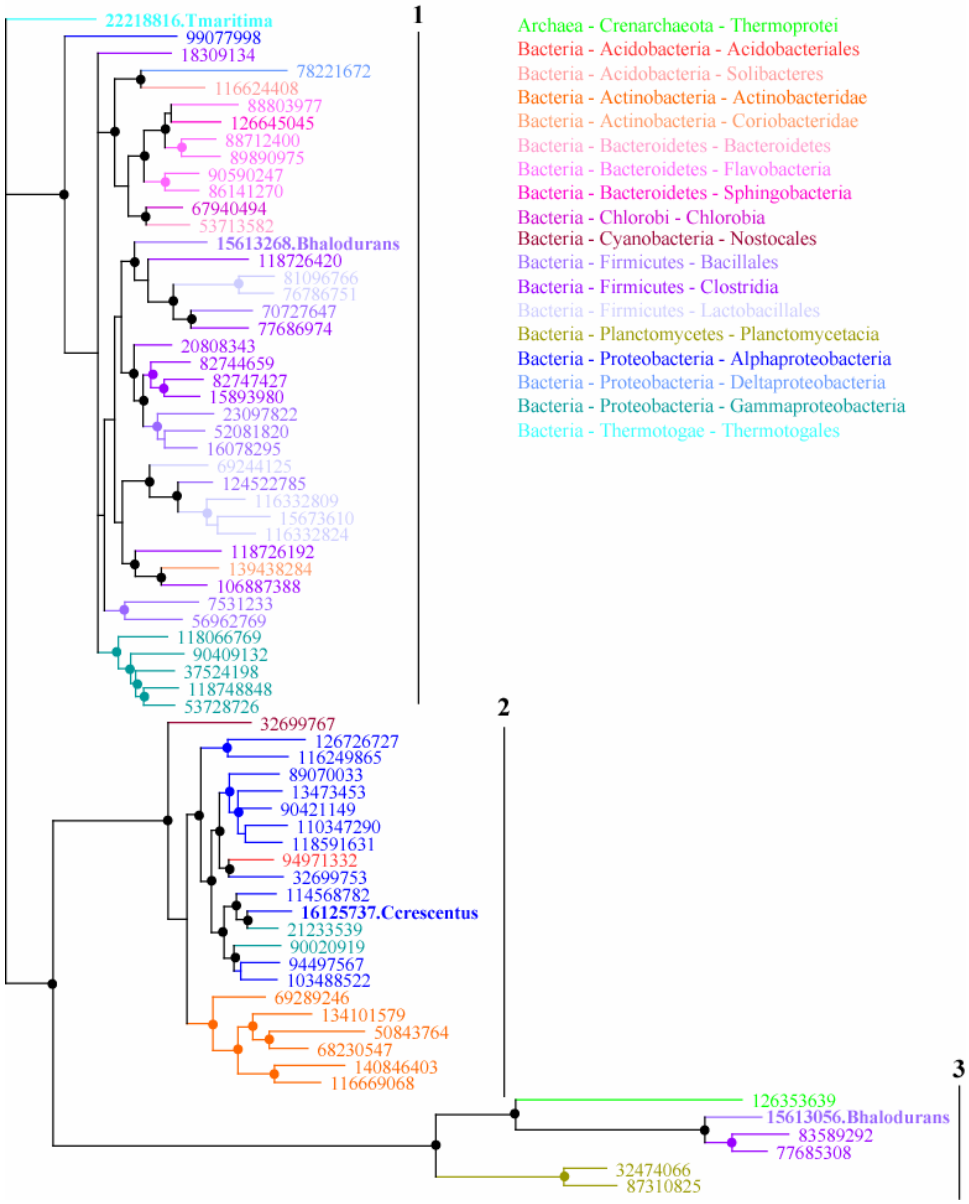


Figure 2.2: Bayesian phylogenetic tree of the proteins in the URI-like sequence set. A representative set of sequences was selected by filtered the complete URI sequence set to 70% identity. Sequences are listed according to their NCBI gi number. The crystallized URI sequences from *T. maritima*, *B. halodurans*, and *C. crescentus* are indicated with species abbreviations. Branch confidence values greater than 0.9 are indicated with circles. The three sequence clusters discussed in the text are indicated with numbers.

Characterization of Bh0705. The catalytic activity for Bh0705 was determined with D-glucuronate and D-galacturonate as substrates. This enzyme was found to isomerize either of these compounds to their respective ketoacid sugar products and the kinetic constants are presented in **Table 2.1**. The protein Bh0705 was found to be more specific for the isomerization of D-glucuronate compare to D-galacturonate since the ratio of k_{cat}/K_m for these two compounds is approximately 100. The two active site directed inhibitors, D-arabinaric acid (**1**) and the hydroxamate derivative (**2**), that were designed to mimic the proposed *cis*-endiol intermediate (**3**) were found to be potent competitive inhibitors for the isomerization of D-glucuronate. The inhibition constants are provided in **Table 2.2**.

Table 2.1: Kinetic parameters for Bh0493 and Bh0705 with D-glucuronate and D-galacturonate

enzyme	substrate	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})
Bh0493	D-glucuronate	5.2 ± 0.3	0.05 ± 0.01	(1.1 ± 0.3) x 10^5
		3.6 ± 0.3	15 ± 5	(2.5 ± 0.8) x 10^2
	D-galacturonate	2.4 ± 0.1	0.05 ± 0.01	(4.7 ± 0.8) x 10^4
		0.9 ± 0.1	12 ± 5	72 ± 30
Bh0705	D-glucuronate	9.7 ± 0.1	0.33 ± 0.01	(2.9 ± 0.10) x 10^4
	D-galacturonate	1.2 ± 0.01	3.3 ± 0.1	(3.6 ± 0.12) x 10^2

Kinetic constants were determined at 30 ºC, pH 8.0 from a fit of the data to equation 2.1 or 2.2.

Table 2.2: Inhibition constants with mimics of *cis*-enediol intermediate[a]

enzyme	inhibitor	K_i (μM)
Bh0493	D-arabinaric acid (**1**)	$(5.5 \pm 0.3) \times 10^{-2}$
	Monohydroxamate derivative (**2**)	2.1 ± 0.1
Bh0705	D-arabinaric acid (**1**)	0.40 ± 0.02
	Monohydroxamate derivative (**2**)	18 ± 2

[a]These constants were obtained at pH 8.0, 30 ºC from a fit of the data to equation 2.3.

Purification and Properties of Bh0493. The plasmid containing the gene encoding protein Bh0493 was transformed into BL21(DE3) cells, and the protein was expressed by induction with IPTG. The protein was soluble and significant amounts of protein were obtained and purified to homogeneity. Using ICP-MS, the protein was found to contain 0.5 equivalents of Zn^{2+} per subunit. SDS-PAGE indicated that the subunit molecular weight is approximately 49 kDa, which corresponds well to the calculated molecular weight of Bh0493 based on the DNA sequence.

Characterization of Bh0493. The enzymatic activity of Bh093 was tested using D-glucuronate and D-galacturonate as substrates. The enzyme catalyzes the isomerization of D-glucuronate to D-fructuronate and D-galacturonate to D-tagaturonate. The double-reciprocal plots for both substrates were biphasic (concave downward), and thus, the initial velocity data were fitted to eq 2.2. The double-reciprocal plots for the isomerization of D-glucuronate and D-galacturonate are shown in **Figure 2.3**. The kinetic constants at high and low concentrations of substrate are provided in **Table 2.1**.

The structural mimics of the proposed cis-enediol intermediate were found to be competitive inhibitors for the isomerization of D-glucuronate when the assays were conducted at low substrate concentrations. The kinetic constants are presented in **Table 2.2**.

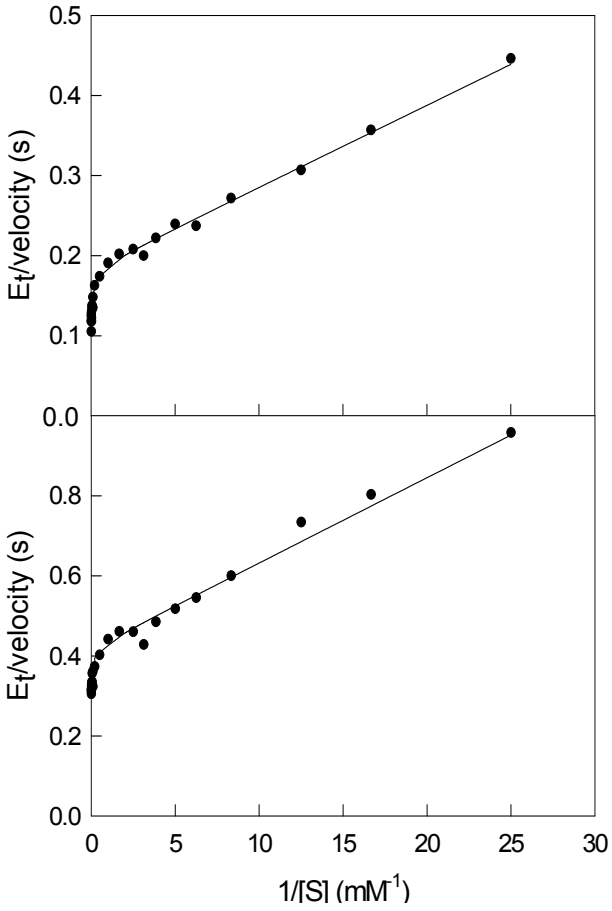


Figure 2.3: Double reciprocal plot for the isomerization of D-glucuronate (top) and D-galacturonate (bottom) by Bh0493. The initial velocities were fit to equation 2.2.

Structure of Bh0493. Two crystal forms of Bh0493 were solved and refined. A tetragonal crystal form was solved to a resolution of 2.4 Å and triclinic crystal form at a resolution of 2.0 Å. Both crystal forms contain similar Bh0493 trimers. There is one protein trimer in the asymmetric unit of the tetragonal form and there are four protein trimers in the asymmetric unit of the triclinic crystal form. Three molecules in every trimer are connected by non-crystallographic three-fold axis. Several ions were observed that sit on the local three-fold axes of every trimer. The Na^+ ion on the local three-fold axis in the tetragonal crystal form is coordinated by six water molecules, connected to protein side chains. This crystal form was crystallized from solutions containing NaCl. The triclinic Bh0493 crystal form had no NaCl in the crystallization solutions and the analogous metal position on the local three fold axis was occupied by Zn^{2+} as shown in **Figure 2.4**. The tetragonal crystal form additionally contains a Cl^- ion that sits on the local three-fold axis of the trimer and is surrounded by the side chains of Lys-278 from three adjacent protein molecules. All of these ions additionally stabilize the internal structure of the trimers.

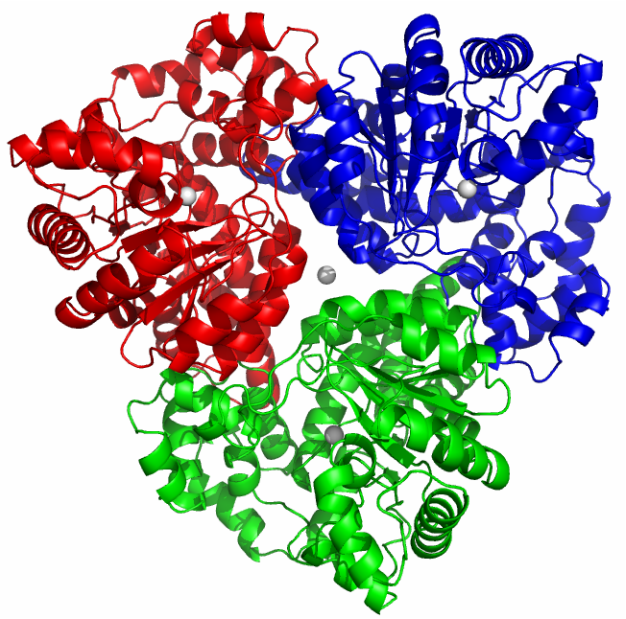


Figure 2.4: Ribbon representation of the trimeric form of Bh0493. Each color represents an individual subunit. The zinc ions are shown as grey spheres.

The structure of the Bh0493 monomer is composed of a $(\beta/\alpha)_7\beta$-barrel with N-terminal and C-terminal extensions of the chain from both sides of the barrel. The following chain segments are included in the eight β-strands of the barrel: β1 (residues 21-28), β2 (137-143), β3 (165-170), β4 (219-224), β5 (253-260), β6 (295-301), β7 (318-323) and β8 (349-356). The N-terminal extension of the chain includes α-helix 6-20. The C-terminal extension includes two long distorted α-helices from residues 360-386 and 393-413. These three long α-helices form a mutual hydrophobic core and stick out of the molecule. The long loop between strands β1 and β2 of the barrel contain 8 helices. The chain segment 40-122 of this loop contains 6 helices, sticks out of the molecule and can be considered as a separate domain B. The long interval between

strands β3 and β4 of the barrel contains 3 helices. The remaining intervals between the β-strands of the barrel contain one helix each.

The active site is located at the C-terminal end of the barrel in large domain A of the Bh0493 monomer and is open to external solvent. The three active sites from three adjacent monomers keep their positions approximately at the vertices of the trimer at a maximum distance from three-fold axis and from each other. The Zn^{2+} ion is bound in the active site of every monomer and is coordinated by His-26 and His-28 from the strand β1 and by Asp-355 from the strand β8 as illustrated in **Figure 2.5**. The fourth ligand for Zn is a water molecule. The inter-atomic distances between the Zn in the active site to these four ligands are 2.3, 2.5, and 2.3 and 2.0 Å, respectively. The other residues whose side chains may be important in substrate binding are: His-49 and Tyr-50 from the domain B, Arg-170 from the strand β3, Trp-325 from the loop after β7 and Arg-357 from the loop after β8.

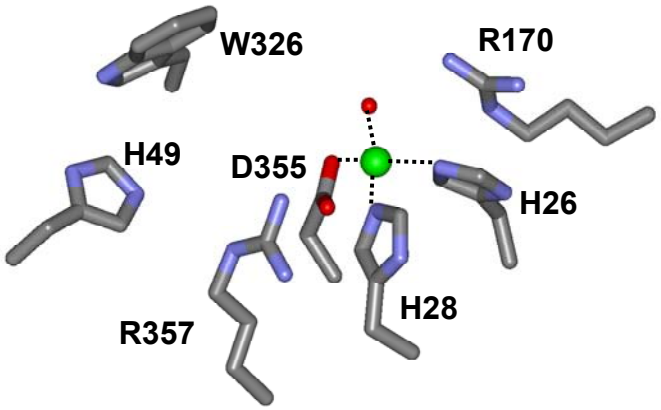


Figure 2.5: Close-up view of the active site region of Bh0493. The relative orientation of the zinc (green sphere), the four metal ligands and other conserved residues that are likely candidates for the binding of substrate and catalytic transformations are shown.

The Bh0493 trimer is mainly stabilized by interactions between the C-terminal helices 360-386 and 393-413 of one monomer and domain B (40-122) of the adjacent monomer. These interactions include main chain – main chain hydrogen bonds and side chain – main chain hydrogen bonds. Additional trimer stabilizing effects are interactions between the loops β5-β6, β6-β7 and β7-β8 of every monomer with the corresponding loops of neighboring monomer within the Bh0493 trimer.

DISCUSSION

Since Holm and Sander first identified a diverse set of amidohydrolases related to adenosine deaminase, phosphotriesterase, and urease in 1997 (*21*), increasingly sophisticated sequence search methods as well as the sequencing of additional genomes have revealed that the AHS contains far more diversity than previously thought, both in terms of function and in terms of sequence (**Figure 2.1A**). For example, α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) and related enzymes, recently identified as part of the AHS (*115*), catalyze nonoxidative decarboxylase reactions. Although these reactions likely share some mechanistic similarity with that of canonical AHS members, and not surprisingly, these sequences are quite different from those of the core superfamily members. As shown by the network representation of the AHS in **Figure 2.1A**, at an *E*-value cutoff where most of the amidohydrolase families identified by Holm and Sander are tightly clustered, ACMSD and related sequences form a distinct, although connected, cluster. The uronate isomerase-like sequences are even more divergent than those of ACMSD and the related decarboxylases. Even at the

permissive *E*-value cutoff used in **Figure 2.1A**, they do not connect to any other AHS sequences. Uronate isomerase is also an outlier of the superfamily in terms of function, being the only characterized superfamily member that catalyzes the interconversion of aldose and ketose functional groups (*98*).

Given currently available tools, enzymes can often be reliably assigned to a superfamily, especially when X-ray structures are available to augment sequence information. However, classifying superfamily sequences into appropriate families each associated with a specific overall reaction can be much more challenging, especially within superfamilies in which many divergent families catalyzing different overall reactions conserve active site residues common to all of the superfamily members. In such cases, even closely related enzymes may catalyze different overall reactions (*63*), and conversely, quite distantly related enzymes may catalyze the same overall reaction (*116*). The problem of functional inference in such superfamilies has been described in some detail for the muconate lactonizing enzyme subgroup of the enolase superfamily (*116*). This is also an issue in the AHS, where, for example, in silico docking approaches were required to identify the function of protein Tm0936 from *T. maritima* (PDB ID 1plm and 1j6p) as an adenosine and thiomethyl adenosine deaminase (*7*). By sequence similarity, this protein most resembles the large chlorohydrolase and cytosine deaminase subgroup within the AHS but shows no sequene similarity with adenosine deaminases.

Even within the canonical members of the URI family, substantial divergence is apparent, with the pairwise sequence identity among the top three sequences shown in the alignment given in **Figure 2.6** ranging from 35 to 52% and the sequence identities between these sequences and that from *C. crescentus* falling to between 22 and 27%. The new outlier uronate isomerase from *B. halodurans*, Bh0493, is even more distant, with a pairwise sequence identity to the other sequences in **Figure 2.6** ranging from 17 to 22%. As shown by the network representation in **Figure 2.1B**, it is only at an *E*-value threshold of 10^{-10} that the outlier uronate isomerases, including Bh0493, connect to the tightly clustered canonical uronate isomerases. At a slightly more stringent *E*-value cutoff of 10-12, the outlier uronate isomerases form a distinct group, losing their connection to the canonical uronate isomerases (not shown). However, because Bh0493 is so distant from even the canonical uronate isomerases, it is useful for identifying minimum sequence requirements for URI activity. The few active site residues conserved across Bh0493 and the canonical uronate isomerases (**Figure 2.6**) are likely candidates for residues required for substrate binding and catalysis. These residues include the conserved HxH motif from β-strand 1, the aspartate from β-strand 8, and a WWF motif from β-strand 7. We have demonstrated in unpublished experiments with

the uronate isomerase from *E. coli* that residues corresponding to His-26, His-28, His-49, Arg-170, Asp-355, and Arg-357 in Bh0493 are critical for the functioning of this enzyme. These residues are conserved in the sequence alignment of Bh0493 with other authentic uronate isomerases (**Figure 2.6**).

The identification of Bh0493 as a uronate isomerase is further supported by the gene context within the genome of *B. halodurans*. Presented in **Figure 2.7** is a schematic representation of the open reading frames in the vicinity of the gene that encodes for Bh0493. The protein Bh0492 is annotated as a tagaturonate oxidoreductase (UxaB) and Bh0490 is an altronate dehydratase (UxaA). The gene for Bh0492 is homologous to the *uxa*B gene in *E. coli* and the two proteins have an amino acid sequence identity of 38%. Likewise, Bh0490 has a sequence identity of 54% to altronate dehydratase from *E. coli* (UxaA). The co-location of these three genes suggests an operon for the metabolism of D-galacturonate to pyruvate and D-glyceraldehyde-3-phosphate (**Scheme 2.1**). This conclusion is further supported by the annotation of Bh0494 as an exopolygalacturonate lyase (PelX).

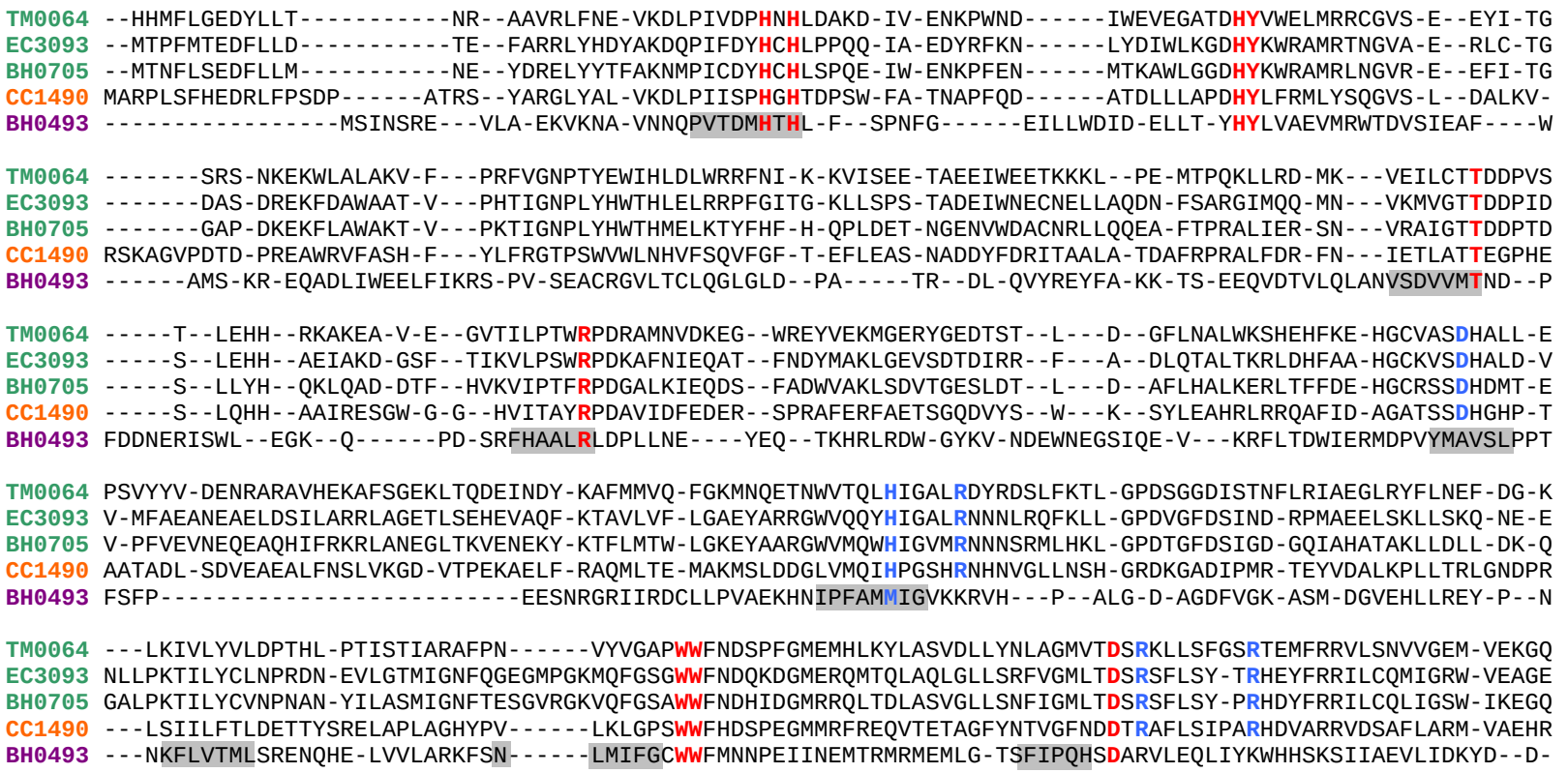


Figure 2.6: Protein sequence alignment of representative URI-like sequences from *T. maritima* (Tm0064), *E. coli* (b3092), *B. halodurans* (Bh0705 and Bh0493), and *C. crescentus* (Cc1490). The three sequence groups shown in **Figure 2.3** are indicated by the colors on the sequence headers (green=1; orange=2; and purple=3). Residues within the active site that may influence substrate binding and/or catalysis are colored. Red indicates residues conserved across the entire URI sequence set, and blue indicates positions that are completely conserved within a group but not across the entire sequence set. β-strands for Bh0493 are indicated with gray shading.

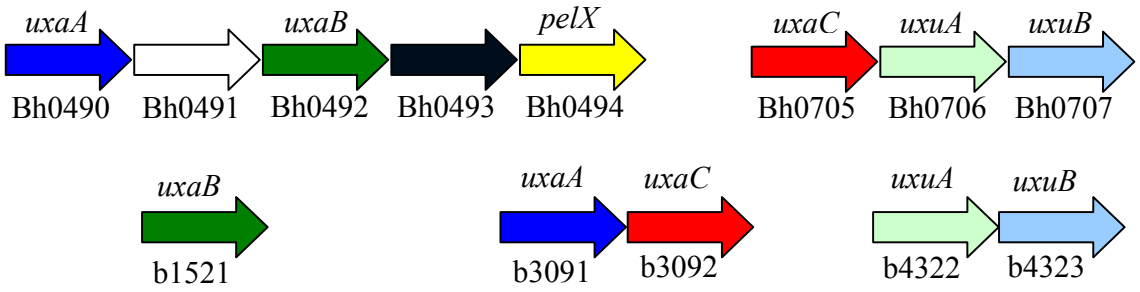


Figure 2.7: Chromosomal arrangement of the genes that code for enzymes involved in the metabolism of D-glucuronate and D-galacturonate. (top) *B. halodurans*. (bottom) *E. coli* K12. The relative lengths of each gene are not drawn to scale.

In nearly all microorganisms the isomerization of the two uronic acids, D-glucuronate and D-galacturonate, are catalyzed by the same enzyme. This is certainly true for *E. coli* and *T. maritima*. However, in the genome of *B. halodurans* there is another gene that is annotated as *uxa*C and apparently this enzyme also catalyzes the isomerization of uronic acids. The gene for this protein (Bh0705) is found adjacent to two other genes currently annotated as *uxu*B (Bh0707) and *uxu*A (Bh0706). The enzymes encoded by these two genes are expected to be D-mannonate oxidoreductase (UxuB) and D-mannonate dehydratase (UxuA). The amino acid sequence of Bh0705 is 53% identical with the prototypical uronate isomerase from *E. coli* (UxaC, b3092) while Bh0706 is 38% identical with the sequence for the authentic mannonate dehydratase from *E. coli* (UxuA, b4322). However, the protein Bh0707 has no apparent homology to the *uxu*B gene product from *E. coli* (b4323). Nevertheless, this enzyme does have a high sequence identity to the gene product of y*jm*F from *B. subtilis* and *uxu*B from *B. stearothermophilus*. Both of these proteins have been annotated as D-mannonate

oxidoreductases (*117*, *118*). Mannonate dehydrogenases generally belong to the long-chain alcohol dehydrogenase superfamily, but the protein sequence of *yjmF* is identified as a member of the short-chain alcohol dehydrogenase superfamily. The protein from *B. subtilis* was first recognized as a mannonate oxidoreductase because it was found in a cluster of genes that are part of the hexuronate catabolic pathway. The identification of Bh0707 as a mannonate dehydrogenase and Bh0706 as a mannonate dehydratase indicates that Bh0705 is in an operon for the metabolism for D-glucuronate. Therefore, it is apparent that in *B. halodurans* there are separate pathways for the metabolism of D-glucuronate and D-galacturonate. Bh0705 is in the operon for D-glucuronate whereas Bh0493 is in the operon for D-galacturonate.

The two putative uronate isomerase enzymes from *B. halodurans*, Bh0493 and Bh0705, were expressed, purified and characterized. It was found that the enzyme Bh0705 will catalyze the isomerization of both D-glucuronate and D-galacturonate. However, the enzyme is significantly more active with D-glucuronate than with D-galacturonate. The value of k_{cat}/K_m for the isomerization of D-glucuronate (2.9×10^4 M^{-1} s^{-1}) is approximately two orders of magnitude higher than it is for the isomerization of D-galacturonate (3.6×10^2 M^{-1} s^{-1}). This indicates that Bh0705, unlike the homologous enzyme from *E. coli*, is relatively specific for the isomerization of D-glucuronate. The corresponding enzyme from *E. coli* (b3092) has been shown to isomerize D-glucuronate with a k_{cat}/K_m of 4×10^5 M^{-1} s^{-1} and D-galacturonate with a k_{cat}/K_m of 2×10^5 M^{-1} s^{-1} (*98*). Therefore, the lone uronate isomerase from *E. coli* cannot discriminate between these two epimeric sugars.

The enzyme Bh0493 can catalyze the isomerization of D-glucuronate and D-galacturonate with a similar catalytic activity for either compound. The observed double-reciprocal plots for Bh0493 are nonlinear for either substrate with an apparent activation at high substrate concentrations. The dependence of the observed catalytic activity as a function of substrate concentration was fit to equation 2.2. At low substrate concentrations of D-glucuronate the value of k_{cat}/K_m is 1.1 x 10^5 M^{-1} s^{-1} whereas at low concentrations of D-galacturonate the value of k_{cat}/K_m is 4.7 x 10^4 M^{-1} s^{-1} and thus there is very little discrimination between these two acid sugars. At saturating concentrations of D-glucuronate the effective turnover number is ~8.8 s^{-1} (5.2 + 3.6 s^{-1}) and somewhat less with D-galacturonate (3.3 s^{-1}). The reason for the nonlinearity in the double reciprocal plots is unknown but it could represent a homotropic cooperativity between subunits. These results validate the hypothesis that Bh0705 catalyzes the isomerization reaction in the D-glucuronate pathway and Bh0493 is primarily responsible for the isomerization reaction that initiates the utilization of D-galacturonate, although this enzyme catalyzes the isomerization of both compounds.

Two compounds that mimic the proposed *cis*-enediol intermediate in the isomerization of D-glucuronate and D-galacturonate were tested as inhibitors of the reactions catalyzed by Bh0493 and Bh0705. D-Arabinaric acid can be considered as a mimic for either compound depending on how the dicarboxylic acid is oriented within the active site. This compound was found to be an excellent competitive inhibitor for the isomerization of D-glucuronate with either enzyme. With Bh0493, the lowest K_i value (at low substrate concentrations) was found to be 55 nM whereas with Bh0705 the

competitive inhibition constant was 400 nM. The competitive inhibition constants for the monohydroxamate derivative of D-arabinaric acid are considerably weaker for both enzymes relative to the effects of D-arabinaric acid. For Bh0493 and Bh705 the observed K_i values are 2 and 18 μM, respectively. The monohydroxamate derivative made for this investigation is an analogue of the *cis*-endiol of D-galacturonate rather than of D-glucuronate. This partially explains the weaker binding of the hydroxamate with Bh705 but it does not explain the weaker binding, relative to D-arabinaric acid, with Bh0493. In any event, the inhibition of these compounds is consistent with the formation of a *cis*-enediol-like reaction intermediate during the isomerization of the aldose substrate to the ketose product.

The first reported crystal structure in the uronate isomerase subfamily of the AHS was from *T. maritima*. The enzyme Bh0493, with an overall sequence identity of 19% to Tm0064, is structurally similar to this enzyme with an r.m.s.d. of 2.3 Å. Like Tm0064, the quaternary structure of Bh0493 is a homotrimer organized as a pinwheel (**Figure 2.4**). Members of the AHS are generally found in oligomerization states with an even number of subunits. For example, phosphotriesterase and dihydroorotase exist as homodimers (*34*, *119*). NagA from *E. coli* is tetrameric (*42*), whereas cytosine deaminase and isoapartyl dipeptidase adopt hexameric and octameric oligomerization states (*35*, *47*). Uronate isomerase is thus the first example for a trimeric quaternary arrangement of subunits within this superfamily of enzymes.

Structural comparisons of Bh0493 with Tm0064 indicate that tertiary fold of Bh0493 is also organized into two distinct domains (**Figure 2.8**). Domain A assumes

the amidohydrolase-like structural fold (TIM-barrel) and is the site for the catalytic transformations. Domain B is composed mainly of α-helices and loops and possesses a unique structural fold. The structural alignment of domain B for these two enzymes indicates a nearly identical fold with the exception of one helix being slightly longer in the *T. maritima* enzyme. Comparison of domain A also reveals a highly conserved structural fold in these two enzymes. However, one major difference between the two structures is an extra helix in Tm0064 of over 20 residues that occurs at the end of β-strand 4. The structural alignment also demonstrates that the histidine contributed from the end of β-strand 5 (highly conserved in nearly all of the enzymes in the AHS) is not conserved in Bh0493, suggesting that this histidine is not important for catalytic activity. Aside from the HxH motif at the end of β-strand 1 and the aspartate at β-strand 8, several other conserved residues are present in this alignment including His-49, Arg-170, Arg-357, and the WWF motif the end of β-strand 7. Although the details of the chemical mechanism for the enzymatic isomerization of D-glucuronate and D-galacturonate are uncertain, it is quite likely that these specific residues will be found to play the critical roles in the catalytic activity of the uronate isomerases. A sequence alignment including all uronate isomerase like sequences shows that the residues identified above are highly (in most cases, completely) conserved, supporting their importance for catalysis and/or substrate recognition (**Figure 2.6**). Additional positions within the active site appear to be conserved within one of the three uronate isomerase groups identified based on the phylogenetic tree, but not across the entire uronate isomerase sequence set (**Figure 2.6**), possibly indicating different strategies for substrate recognition within each group.

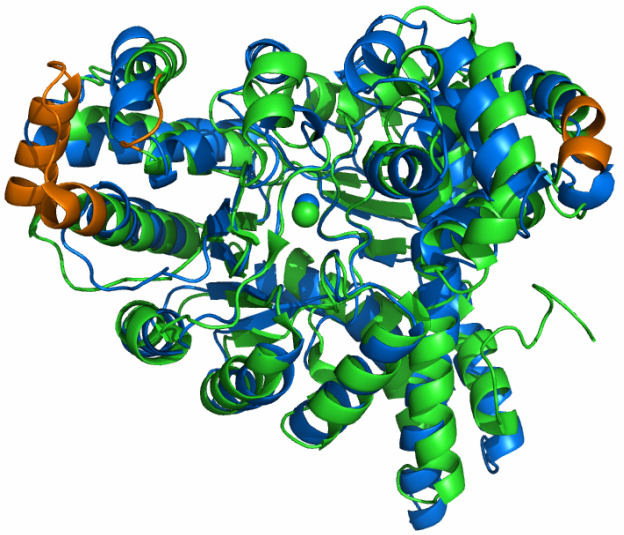


Figure 2.8: Overlay of the ribbon representation of the two structures, Tm0064 (green) and Bh0493 (blue). In orange are the non-overlapping regions between the two structures. The green sphere represents the putative metal ion in Tm0064 and the blue sphere represents the zinc ion in Bh0493.

B. halodurans may not be the only organism with separate enzymes for the isomerization of D-glucuronate and D-galacturonate. *Alkaliphilus metalliredigenes* QYMF, also an alkaliphilic bacterium like *B. halodurans*, has two members in the uronate isomerase sequence set, one of which (gi: 77686974) has 48% sequence identity to the prototypical uronate isomerase of *E. coli* and the other of which (gi: 77685308) has 57% identity to the outlier uronate isomerase Bh0493. The gene encoding protein 77686974 is located near homologues of *E. coli* mannonate dehydrogenase (UxuB, gi: 77686972) and mannonate dehydratase (UxuA, gi: 77686973), enzymes specific for the breakdown of D-glucuronate. Although the *A. metalliredigenes* genome has not been fully assembled, two fragments homologous to *E. coli* UxaB are found near 77685308,

the second putative UxaC (gi: 77685309 and 77685310). Full length homologs of *E. coli* UxaB and UxaA are also found in the genome (gi: 77686989 and 77686988, respectively). The homology of the two putative *A. metalliredigenes* UxaCs to the *B. halodurans* UxaCs, along with genomic context information, suggest that *A. metalliredigenes* has distinct isomerases for the metabolism of D-glucuronate and D-galacturonate.

Clostridium beijerincki NCIMB 8052 and *Saccharophagus degradans* 2-40 also contain multiple uronate isomerase like sequences, with genome context that suggests one may be specific for the isomerization of D-gluconate and the other for D-galacturonate. However, the putative D-galacturonate isomerases from these organisms do not cluster with those from *B. halodurans* and *A. metalliredigenes*, nor with each other. If they are, indeed, specific for D-galacturonate, this specificity appears to have evolved multiple times within the uronate isomerase like group.

Other proteins homologous to the outlier uronate isomerase Bh0493 include protein sequences from *Moorella thermoacetica* ATCC 39073 (gi: 83589292), *Blastopirellula marina DSM3645* (gi: 87310825), *Caldivirga maquilingensis* IC-167 (gi: 126353639) and *Rhodopirellula baltica SH 1* (gi: 32474066). These organisms are unusual because they do not appear to have a protein homologous to the prototypical uronate isomerase in *E. coli*. The *M. thermoacetica* homologue is located near genes homologous to *E. coli* UxuB and UxuA, indicating that it may be a genuine uronate isomerase. The remaining sequences, however, cannot be validated based on genomic

context. The function of the *C. maquilingensis* sequence is particularly open to question, as this is the only nonbacterial organism with a uronate isomerase like sequence.

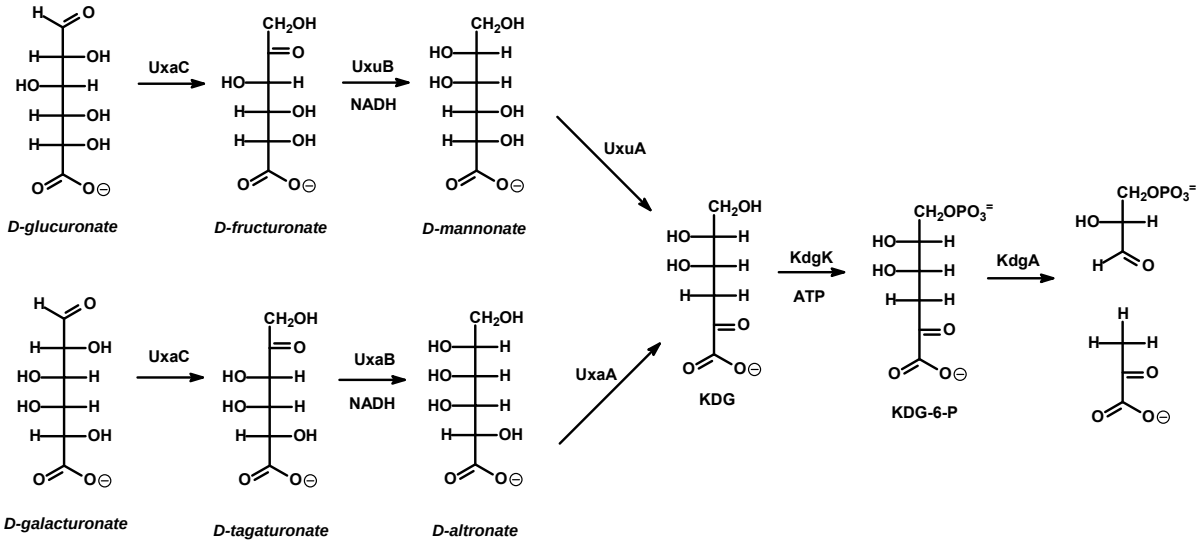
From the kinetic and structural studies, we have determined that the protein Bh0493 from *B. halodurans* is a uronate isomerase and a diverged member of the amidohydrolase superfamily. In *B. halodurans*, the gene encoding uronate isomerase had been identified previously as Bh0705 based on sequence identity. Initial velocity studies of Bh0705 and Bh0493 with the substrates D-glucuronate and D-galacturonate indicate that Bh0705 is highly specific for D-glucuronate while Bh0493 can utilize either substrate with similar efficiency. Bh0493 is located adjacent to genes that are annotated as tagaturonate oxidoreductase, altronate dehydratase, and exopolygalacturonate lyase (an operon for D-galacturonate metabolism), while the gene for Bh0705 is located near mannonate dehydratase and mannonate oxidoreductase (an operon for D-glucuronate metabolism). Sufficient information is not available to hypothesize as to why *B. halodurans* and some other organisms have more than one URI gene or whether this leads to a selective advantage for these organisms. Nor is the evolutionary path that connects this highly divergent family to the AHS clear. However, identification of this outlier Bh0493 sequence as a uronate isomerase extends the boundaries of the superfamily further into the "twilight" of functional and structural divergence than has been previously recognized.

CHAPTER III

THE MECHANISM OF THE REACTION CATALYZED BY URONATE ISOMERASE ILLUSTRATES HOW AN ISOMERASE MAY HAVE EVOLVED FROM A HYDROLASE WITHIN THE AMIDOHYDROLASE SUPERFAMILY

Uronate isomerase (URI[1]) catalyzes the first step in the pathway for the metabolism of D-glucuronate and D-galacturonate. In this transformation, D-glucuronate and D-galacturonate are initially isomerized into their corresponding keto products, D-fructuronate and D-tagaturonate, respectively (*120*). D-Fructuronate and D-tagaturonate are then reduced to D-mannonate and D-altronate, respectively, by mannonate and altronate dehydrogenase in the presence of NADH (*121*). The pathways converge through a dehydration reaction where mannonate dehydrase and altronate dehydrase convert mannonate and altronate to 2-keto-3-deoxy-D-gluconic acid (KDG). This product is then phosphorylated by the enzyme ketodeoxygluconic acid kinase with ATP to form 2-keto-3-deoxy-6-phosphogluconic acid (KDG-6-P). In the final step of this pathway, 2-keto-3-deoxy-6-phosphogluconic acid is cleaved by an aldolase to yield pyruvate and D-glyceraldehyde-3-phosphate, which enter the citric acid cycle and glycolysis. The entire pathway is summarized in **Scheme 3.1** (12*0,121*).

Scheme 3.1: Entire pathway for the metabolism of D-glucuronate and D-galacturonate.

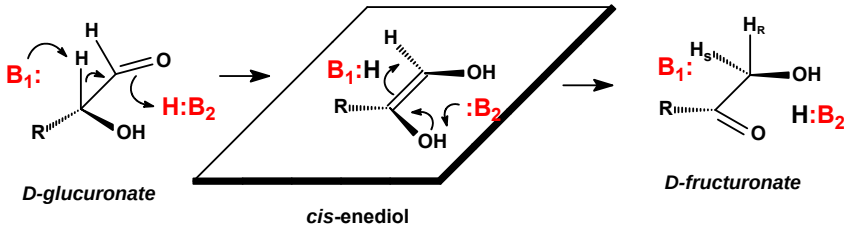


We have demonstrated that uronate isomerase is a member of the amidohydrolase superfamily of enzymes (*44*). The majority of the functionally characterized members of the amidohydrolase superfamily catalyze the hydrolysis of amide or ester bonds to carbon or phosphorus centers (*21,31*). Well characterized examples include dihydroorotase (*60*), urease (*32*) and phosphotriesterase (6*8*). Members of this superfamily also catalyze the deamination of many nucleotides including adenosine (*56*), cytosine (*47*), and guanine (*122*). The active sites of these enzymes generally consist of a mononuclear or binuclear metal center that is perched at the C-terminal end of the β-barrel in proteins within a $(\beta/\alpha)_8$ structural fold. The most highly conserved residues in the AHS include two histidines from β-strand 1, histidines after the ends of β-stands 5 and 6, and an aspartic acid from β-strand 8. Since URI

catalyzes an isomerization of an aldose sugar to the corresponding ketose product, this enzyme is one of the most divergent members of the amidohydrolase superfamily. The mechanistic details of this transformation are therefore of significant interest toward a greater understanding of how an active site that originally evolved to catalyze hydrolytic reactions has been re-forged to undergo an isomerization reaction.

We have previously demonstrated that the proton originally at C-2 of D-glucuronate is ultimately found at the *pro*-R position at C-1 of D-fructuronate and that this proton slowly exchanges with solvent (*98*). These results are consistent with a proton transfer mechanism with a *cis*-enediol intermediate. The general mechanism, shown in **Scheme 3.2**, indicates a requirement for at least two residues that participate in the transformation of D-glucuronate into D-fructuronate. A general base ($:B_1$) abstracts the proton from C-2 of D-glucuronate and a general acid ($H:B_2$) facilitates the transfer of a proton to the carbonyl oxygen at C-1 to produce the *cis*-enediol intermediate. In the subsequent step, the ketose product is generated by a proton transfer from the hydroxyl group at C-2 of the proposed intermediate and protonation of C-1 by $H:B_1$. For compounds such as D-glucuronate the enzymatic transformation is made more complicated by the fact that in solution the substrate exists almost entirely as a mixture of two anomeric cyclic hemiacetals.

Scheme 3.2: Proton transfer mechanism in a 1,2-hydrogen transfer reaction.



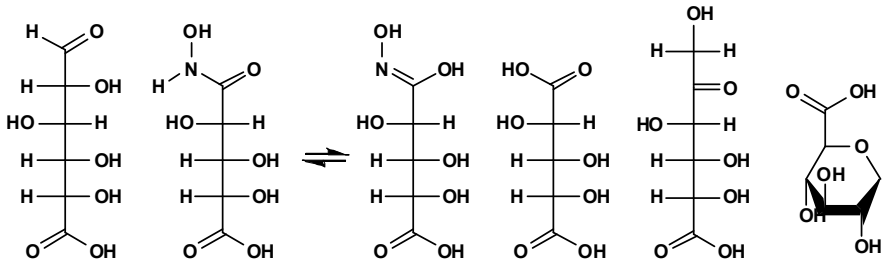
This chapter focuses on a determination of the chemical mechanism for the isomerization reaction catalyzed by URI from *E. coli*. The rate limiting steps have been interrogated by measuring the primary kinetic isotope effects with [2-^{2}H]-D-glucuronate and solvent isotope effects with D_2O for the wild type and mutant enzymes. The rate limitation imposed by product release has been examined using solvent viscosity effects. The identity of the residues involved in the proton transfer events has been probed by pH-rate profiles and characterization of the kinetic constants for mutant enzymes. These approaches have been augmented by the elucidation of the X-ray structure of a uronate isomerase from *Bacillus halodurans* (Bh0493) in the presence of D-glucuronate, D-fructuronate, and two mimics of the *cis*-enediol intermediate.

MATERIALS AND METHODS

Materials. D-Glucuronic acid (**I**), L-gulonic acid γ-lactone (**VI**), NADH, buffers, and all other chemicals were purchased from Sigma-Aldrich or Acros, unless otherwise stated. D-arabinaric acid (**III**) and the monohydroxamate derivative of this compound (**II**) were synthesized by Dr. Yingchun Li. [2-^{2}H]-D-glucuronate was also

synthesized by Dr. Yingchun Li. 2,6-Anhydro-L-gulonic acid (**V**) was synthesized by Dr. Chengfu Xu. The structures of these compounds are presented in **Scheme 3.3**. Oligonucleotide syntheses and DNA sequencing were performed by the Gene Technologies Lab of Texas A&M University. Metal analyses were done using inductively coupled plasma mass spectrometry (ICP-MS).

Scheme 3.3: Structures of D-glucuronate (I), hydroxamate derivative (II), D-arabinaric acid (III), D-fructuronate (IV) and 2,6-anhydro-L-gulonic acid (V).



Site-Directed Mutagenesis. Site-directed mutagenesis of URI was performed using the QuikChange mutagenesis kit from Stratagene. The following mutants were obtained by this method: H33N, H33A, H35N, H35A, H59N, H59A, Y60F, Y60A, R186K, R186M, D238N, H297N, R302K, R302M, H297A, W381F, W381A, D412N, D412A, R414K, and R414M. The mutations were confirmed by DNA sequencing of the modified plasmids.

Protein Expression and Purification of Uronate Isomerase. The *uxa*C gene encoding *E. coli* uronate isomerase (gi: 16130987) was amplified by PCR and inserted into pET28 without a histag. The protein was expressed in the *E. coli* strain BL21-(DE3), and single colonies were used to inoculate 5 mL of LB medium supplemented with 50 *í*g/mL kanamycin. The overnight culture was used to inoculate 1 L of TB medium supplemented with 50 μg/mL kanamycin. The large culture was grown at room temperature until an *A*600 of 0.4-0.6 was reached, and then 1.0 mM $ZnCl_2$ was added, followed by induction with 0.4 mM IPTG. The cells were incubated overnight and then centrifuged at 6000 rpm for 10 min at 4 °C. The cell pellet was resuspended in 50 mM HEPES, pH 8.0, and 500 μM $ZnCl_2$. The cells were lysed by sonication and the nucleic acids precipitated by dropwise addition of 2.0% (w/v) protamine sulfate solution. The solution was centrifuged at 12 000 rpm and the lysate fractioned between 50% and 80% saturation of ammonium sulfate. After centrifugation, the pellet was resuspended in a minimal amount of buffer (50 mM HEPES, pH 8.0), loaded onto a Superdex 200 gel filtration column, and eluted at a flow rate of 2.5 mL/min. Fractions containing uronate isomerase were identified by SDS-PAGE and activity assays. The pooled enzyme was loaded onto a Resource Q anion exchange column preequilibrated with buffer A (20 mM HEPES, pH8.0) and eluted with a 0-30% gradient of buffer B (20 mM HEPES, pH 8.0 and 1.0 M NaCl) at a flow rate of 4.5 mL/min. The fractions were assayed and pooled based on the enzymatic activity. The purified enzyme was stored in the refrigerator at 4 °C for up to 1 year without significant loss of activity. The enzymes contained up to 1 equivalent of zinc (depending on the mutant) as measured by ICP-MS.

Cloning and Purification of D-Mannonate Dehydrogenase. The *uxu*B gene (gi: 26111644) from *E. coli* that encodes D-mannonate dehydrogenase (MDH) was cloned into a pET30 expression vector. The gene was transformed into BL21(DE3) and the cells grown in TB broth at 37 °C. D-Mannonate dehydrogenase was purified using the same methods previously described for the isolation of uronate isomerase, but no metal was added to the purification buffers. The cell extract was made 40-60% of saturation with ammonium sulfate, and the resuspended pellet was subjected to fractionation with the aid of a Superdex 200 gel filtration column. Fractions containing D-mannonate dehydrogenase were pooled, quick-frozen, and stored at -80 °C.

Enzyme Assays. The conversion of D-glucuronate to D-fructuronate by URI was coupled to the reduction of D-fructuronate with NADH by mannonate dehydrogenase (MDH). The assays were monitored spectrophotometrically by following the decrease in absorbance at 340 nm. The standard assay conditions contained 50 mM HEPES (pH 8.0), varying concentrations of D-glucuronate, 0.2 mM NADH, excess MDH, and URI in a final volume of 250 μL. The pH-dependence of the kinetic parameters, k_{cat} and k_{cat}/K_m, were measured over the pH range of 5.3-10.3 at 0.20 pH intervals. The buffers used for the pH-rate profiles were MES, PIPES, HEPES, CHES, and CAPS. The pH values were recorded after the completion of the assays. The effects of solvent viscosity on the kinetic constants were determined at pH 8.0 using sucrose as the micro-viscogen at 25 ºC. The concentrations of sucrose were 0%, 10%, 14%, 20%, 24%, and 32% (w/w), and the corresponding relative viscosities were 1, 1.3, 1.5, 1.9, 2.2, and 3.2 (*15, 16*). The solvent isotope effects on the kinetic parameters for URI and two mutant

enzymes (D412N and R414M) were measured in 99% D_2O at a pD of 8.4. The primary deuterium kinetic isotope effects were obtained by direct comparison of the kinetic constants at pH 8.0 for [2-^{1}H]-D-glucuronate and [2-^{2}H]-D-glucuronate.

Data Analysis. The kinetic parameters, k_{cat} and k_{cat}/K_m, for uronate isomerase with D-glucuronate as the substrate were determined by fitting the initial velocity data to equation 3.1 where ν is the initial velocity, E_t is the total enzyme concentration, k_{cat} is the turnover number, [A] is the substrate concentration, and K_m is the Michaelis constant. The profiles for the variation of k_{cat} or k_{cat}/K_m with pH were fit to equation 3.3, where c is the pH-independent value of y, K_a and K_b are the dissociation constants of the ionizable groups and H is the proton concentration. The competitive inhibition patterns were fit to equation 3.2, where K_{is} is the slope inhibition constant and I is the concentration of the inhibitor.

$$v / E_t = (k_{cat}\,[A])/(K_a + [A]) \tag{3.1}$$

$$\nu / E_t = (k_{cat}[A])/(K_a(1 + (I/K_{is})) + [A]) \tag{3.2}$$

$$\log y = \log (c / (1 + (H / K_a) + (K_b / H))) \tag{3.3}$$

RESULTS

Requirement for Divalent Cation. The importance of a metal ion for the catalytic activity of uronate isomerase was investigated. The apo-enzyme was prepared and subsequently tested for enzymatic activity using D-glucuronate as the substrate. The wild-type URI from *E. coli* was found to contain 0.9 equivalents of zinc after purification. This protein (3 mL) at a concentration of 3.0 mg/mL was dialyzed against

1 L of dialysis buffer containing 20 mM dipicolinate in 50 mM MES, pH 6.0. The buffer was changed three times over the course of 48 h, after which the catalytic activity and metal content of the enzyme were determined. The chelator effectively removed more than 98% of the bound zinc as indicated by ICP-MS. The activity of the enzyme was assayed with mannonate dehydrogenase and NADH to detect the formation of D-fructuronate in the presence of 10 μM dipicolinate. The apo-enzyme exhibited less than 1% activity of the native enzyme with bound zinc.

*Inhibition by 2,6-Anhydro-L-Gulonic Acid (**V**).* Compound **V** was synthesized as a cyclic analogue mimic of the pyranose form of D-glucuronate. The inhibitory properties of **V** were determined with the wild-type uronate isomerase from *E. coli* and *B. halodurans* (Bh0493). This compound was found to be a competitive inhibitor for both enzymes. The data were fit to equation 3.2 and the values of K_{is} were determined to be 45 ± 4 and 24 ± 2 μM for the URI from *E. coli* and Bh0493, respectively.

pH-Rate Profiles. The kinetic constants for the conversion of D-glucuronate to D-fructuronate were obtained as a function of pH. The pH-rate profiles for the effects of pH on k_{cat} and k_{cat}/K_m are presented in **Figures 3.1A** and **3.1B**, respectively. The pH profiles are bell-shaped and are consistent with a single functional group that must be unprotonated for activity and another functional group that must be protonated for catalytic activity. From a fit of the data to equation 3.3 the kinetic pK_a values from the k_{cat}/K_m plot are 5.5 ± 0.1 and 9.5 ± 0.1, respectively. From the plot of k_{cat} vs. pH, kinetic pK_a values of 5.8 ± 0.1 and 10.2 ± 0.1 were obtained.

Site-Directed Mutants. Site-directed mutagenesis was utilized to identify the involvement of specific amino acids in metal binding, substrate recognition, and the catalytic mechanism of uronate isomerase. Conserved residues were chosen based on the location within the active site of Bh0493, a uronate isomerase found in *B. halodurans*. His-33 and His-35 were mutated to investigate the importance of metal binding and the potential role of the divalent cation on catalytic activity. Mutations at either of these two residues resulted in the dramatic loss of affinity for the divalent cation and a significant reduction in the catalytic activity. The diminution of catalytic activity for the mutants is more severe when these histidine residues are changed to alanine than to asparagine. The highly conserved histidine at the end of β-strand 5 (His-297) and the invariant aspartate at the end of β-strand 8 (Asp-412) were mutated to asparagine and alanine. For the mutations at His-297 there were substantial increases in the Michaelis constant. In contrast, with the mutation of Asp-412 the reduction in catalytic activity was more pronounced on k_{cat}. These four residues are broadly conserved in all members of the amidohydrolase superfamily.

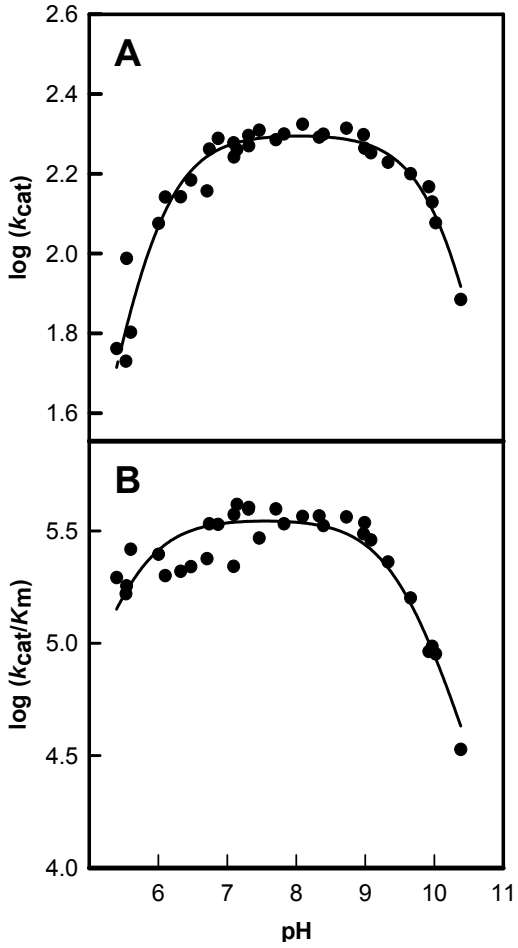


Figure 3.1: pH-rate profile for the wild type uronate isomerase containing 1 equivalent of zinc. The data were fit to equation 3.3. (**A**) Plot of log k_{cat} vs. pH. (**B**) Plot of log k_{cat}/K_m vs. pH.

Additional residues that are conserved among all of the known uronate isomerases were mutated as a probe of functional participation in binding and catalysis. These residues include Trp-381 from the conserved WWF motif after β-strand 7, His-59 and Tyr-60 after β-strand 1, and three conserved arginines (Arg-186, Arg-302, and Arg-414). Mutation of residues Trp-381, Arg-186, and Arg-302 resulted in increases in the value of K_m and small changes in k_{cat}, indicating that these residues most likely take part in substrate recognition and binding. For His-59 and Arg-414 there were relatively small changes in K_m but drastic reductions in the value of k_{cat}. The kinetic constants for the mutants constructed for this investigation are presented in **Table 3.1**.

Kinetic Isotope Effects. Primary deuterium kinetic isotope effects on the isomerization of D-glucuronate were obtained as a probe of the rate limiting steps in the overall reaction mechanism. For the solvent deuterium isotope effects, the kinetic parameters were obtained for the wild type enzyme in H_2O and D_2O with D-glucuronate as the substrate. The double reciprocal plots are presented in **Figure 3.2**. The measured solvent isotope effects for $^{D_2O}k_{cat}$ and $^{D_2O}(k_{cat}/K_m)$ are 1.22 ± 0.02 and 1.10 ± 0.09, respectively. Solvent isotope effects were also determined for four of the site directed mutants, H59N, Y60F, D412N and R414M. The solvent isotope effects for the H59N mutant were 1.7 ± 0.1 for $^{D_2O}k_{cat}$ and 1.4 ± 0.1 for $^{D_2O}(k_{cat}/K_m)$, while the effects for the Y60F mutant were 2.1 ± 0.1 and 1.2 ± 0.1 for $^{D_2O}k_{cat}$ and $^{D_2O}(k_{cat}/K_m)$, respectively. For the D412N mutant, the solvent isotope effects were determined to be 1.3 ± 0.1 and 1.5 ± 0.2 for $^{D_2O}k_{cat}$ and $^{D_2O}(k_{cat}/K_m)$, respectively. For the R414M mutant, the solvent isotope effects were 1.8 ± 0.1 for $^{D_2O}k_{cat}$ and 2.0 ± 0.3 for $^{D_2O}(k_{cat}/K_m)$.

Table 3.1: Kinetic parameters and metal content of mutants

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (M^{-1} s^{-1})	URI/Zn
WT	196 ± 6	0.50 ± 0.05	(4.0 ± 0.4) x 10^5	0.90
H33N	2.1 ± 0.1	(5.0 ± 0.4) x 10^{-2}	(4.7 ± 0.4) x 10^4	0.07
H33A	0.60 ± 0.01	0.20 ± 0.01	(3.0 ± 0.2) x 10^3	0.20
H35N	4.0 ± 0.2	9.4 ± 1.1	(4.3 ± 0.5) x 10^2	<0.05
H35A	0.70 ± 0.04	39 ± 5	18 ± 2	<0.05
H59N	15 ± 1	0.70 ± 0.04	(2.1 ± 0.1) x 10^4	0.96
H59A	0.60 ± 0.01	0.70 ± 0.07	(8.3 ± 0.1 x 10^2	0.95
Y60F	21.7 ± 0.1	0.16 ± 0.01	(1.4 ± 0.1) x 10^5	0.8
Y60A	13.9 ± 0.1	0.21 ± 0.01	(6.6 ± 0.3) x 10^4	0.8
R186K	54 ± 2	2.6 ± 0.2	(21 ± 2) x 10^3	0.94
R186M	4.7 ± 0.1	38 ± 3	(1.3 ± 0.1) x 10^2	0.91
D238N	60 ± 1	1.3 ± 0.1	(4.6 ±0.1) x 10^4	0.70
H297N	30 ± 2	56 ± 5	(5.0 ± 0.5) x 10^2	1.00
H297A	10 ± 1	(2.2 ± 0.3) x 10^2	43 ± 7	0.41
R302K	160 ± 4	2.5 ± 0.2	(6.3 ± 0.5) x 10^4	0.90
R302M	180 ± 9	(2.0 ± 0.3) x 10^2	(8.8 ± 1.3) x 10^2	0.99
W381F	16 ± 1	1.7 ± 0.1	(9.5 ± 0.4) x 10^3	0.90
W381A	250 ± 6	21± 2	(1.2 ± 0.1) x 10^4	0.69
D412N	0.60 ± 0.01	1.00 ± 0.04	(6.0 ± 0.3) x 10^2	0.36
D412A	(9.0 ± 0.3) x 10^{-3}	0.40 ± 0.05	21 ± 3	0.12
R414K	5.8 ± 0.1	0.82 ± 0.02	(7.1 ± 0.2) x 10^3	0.92
R414M	0.70 ± 0.01	1.4 ± 0.1	(5.4 ± 0.2) x 10^2	0.91
These data were obtained at 30 ºC, pH 8.0, with D-glucuronate as the substrate.				

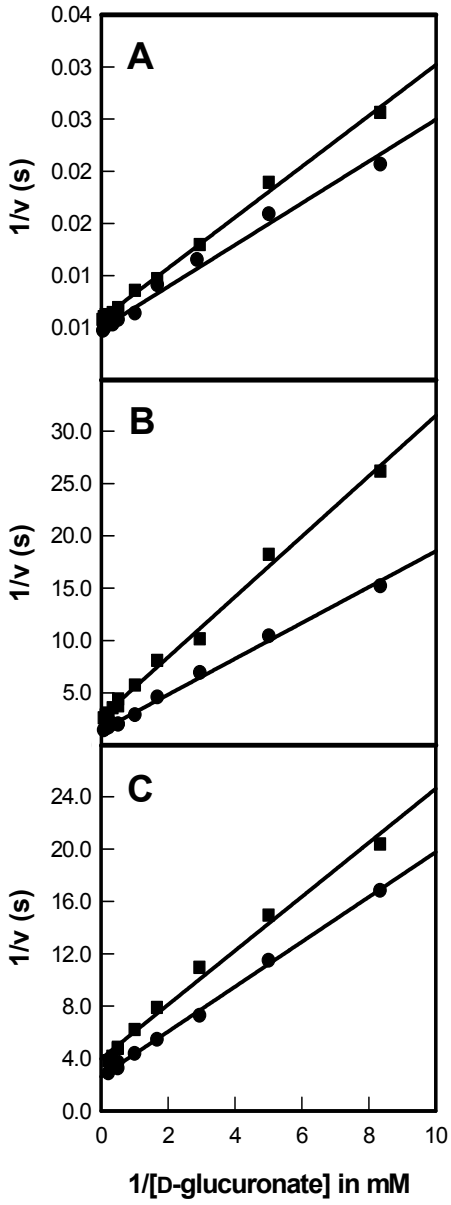


Figure 3.2: Double-reciprocal plots for the solvent isotope effect (H_2O *vs* D_2O) where $1/v$ (s) is plotted against 1/[D-glucuronate] (in mM^{-1}). The data sets in H_2O are represented by circles, and the squares correspond to the values obtained in D_2O. (**A**) represents the plot for the wild type enzyme, (**B**) for the R414M mutant, and (**C**) is the plot of the D412N mutant.

The primary deuterium isotope effects for abstraction of the proton from C-2 of D-glucuronate were determined for the wild type enzyme and two mutants, D412N and R414M. The double-reciprocal plots are shown in **Figure 3.3**. For the wild-type enzyme the primary deuterium isotope effects on k_{cat} and k_{cat}/K_m were determined to be 1.4 ± 0.04 and 1.2 ± 0.01, respectively. For the D412N mutant, the primary deuterium isotope effects were determined to be 2.0 ± 0.2 for $^{D}k_{cat}$ and 1.9 ± 0.3 for $^{D}(k_{cat}/K_m)$. For the R414M mutant, the isotope effect on k_{cat} was 3.2 ± 0.1 and the effect on k_{cat}/K_m was 3.5 ± 0.4.

Solvent Viscosity Effects. Alterations in solvent viscosity were utilized to probe the degree of rate limitation by the binding and dissociation of products and substrates on the kinetic constants of uronate isomerase (*123*). The effects of changes in solvent viscosity on k_{cat} and k_{cat}/K_m were made by the addition of sucrose (*124*). A plot of $^{o}k_{cat}/{}^{\eta}k_{cat}$ versus the relative solvent viscosity for the wild-type enzyme exhibits a slope of 0.72 ± 0.08. The slope for the effect of solvent viscosity on k_{cat}/K_m is 0.64 ± 0.05 for the wild-type enzyme. For the two mutant enzymes, D412N and R414M, the slopes for the effect of solvent viscosity on k_{cat} were found to be 0.07 ± 0.05 and 0.05 ± 0.01, respectively. For the effect on k_{cat}/K_m the slopes were 0.03 ± 0.05 and -0.17 ± 0.02, respectively, with the D412N and R414M mutant enzymes. The kinetic data are presented in **Figures 3.4A** and **3.4B**.

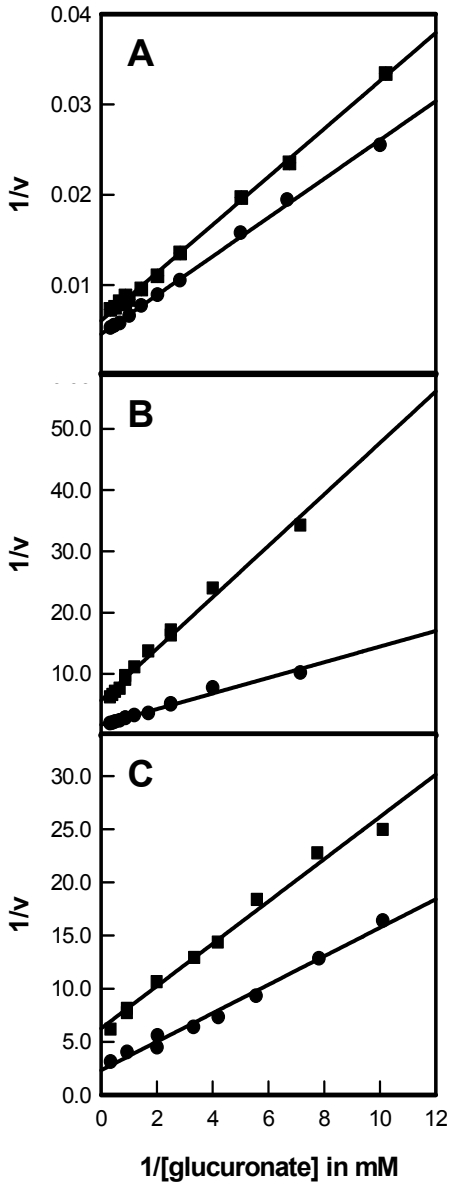


Figure 3.3: Primary isotope effects using protonated and deuterated D-glucuronate at the C-2 position are presented as double-reciprocal plots where (**A**) is the plot for the wild-type enzyme, (**B**) R414M mutant, and (**C**) D412N mutant. The data for the protonated substrate are represented as circles, and the values for the deuterated substrate are denoted as squares.

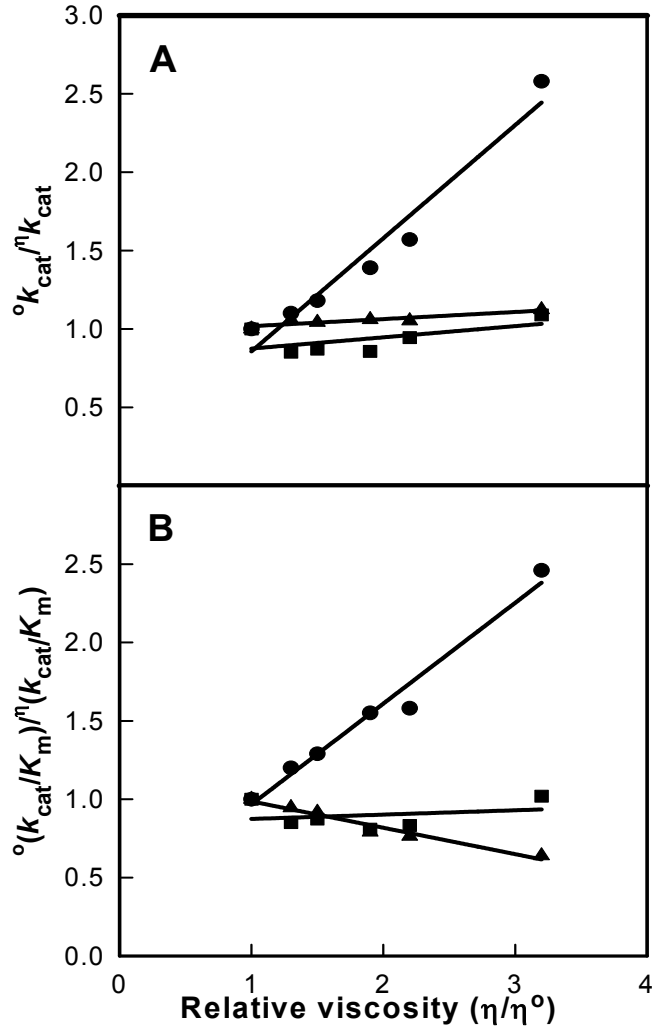


Figure 3.4: The effect of viscosity on the relative values of k_{cat} (**A**) and k_{cat}/K_m (**B**) using sucrose as the microviscogen. The circles indicate wild type enzyme, squares represent the data set for the D412N mutant, and triangles signify the values for the R414M mutant.

Structures of Inhibitor Complexes. The crystal structure of Bh0493 from *B. halodurans* was previously solved in the absence of bound substrates or inhibitors (PDB code: 2QO8, *44*). In that structure zinc is bound in the active site and coordinated to two histidines from β-strand 1 (H26 and H28) and the invariant aspartate found at the end of β-strand 8 (D355). The structure of Bh0493 was solved in the presence of the substrates D-glucuronate (**I**), D-fructuronate (**IV**), and two inhibitors that mimic the proposed *cis*-enediol(ate) intermediate, D-arabinarate (**III**) and its hydroxamate derivative (**II**). Portions of the electron density maps that show the D-glucuronate and D-fructuronate in the active site of Bh0493 are presented in **Figures 3.5A** and **3.5B**. These compounds are bound in the active site in a similar manner and are interacting with the same set of amino acid residues. Stereoscopic images of the active site complexes for compounds **I**, **II**, **III,** and **IV** are presented in **Figures 3.6A-D** and the distances between specific amino acid residues and the bound ligands are provided in **Figures 3.7A-D**.

In the complex with D-glucuronate, the substrate is bound in the open chain configuration. The terminal carboxylate is ion-paired with the guanidino group of Arg-170 and there is a monodentate coordination with the bound zinc. The zinc is also ligated with the hydroxyl group attached to C-5 of the substrate. The hydroxyl group at C-4 does not make any specific interactions with the protein, which is consistent with the observation that both D-glucuronate and D-galacturonate are substrates for this enzyme (*120*). The hydroxyl at C-3 interacts with both Arg-357 and His-49. The nearest residue to the hydroxyl at C-2 of D-glucuronate is Arg-357. The hydroxyl group from Tyr-50 hydrogen bonds with the carbonyl group at C-1 of the substrate. The closest residue to

the proton that is abstracted from C-2 of D-glucuronate is the side chain carboxylate of Asp-355. Similar interactions are found in the complexes with D-fructuronate, D-arabinarate and the hydroxamate derivative. The D-fructuronate complex was identified in a crystal that was grown after a long incubation and subsequent co-crystallization of Bh0493 with D-glucuronate.

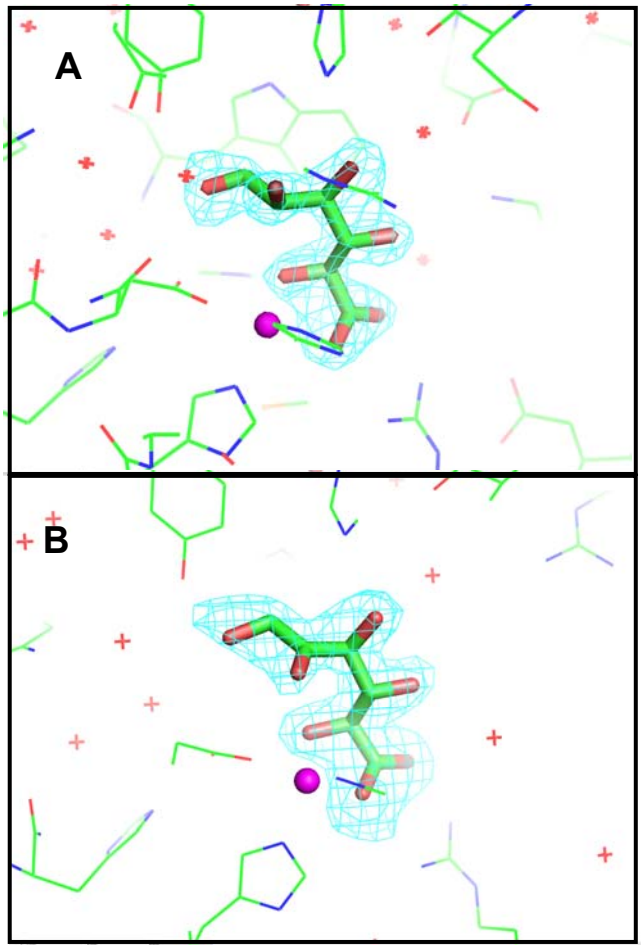


Figure 3.5: Electron density maps of (**A**) D-glucuronate and (**B**) D-fructuronate in the active site of Bh0493.

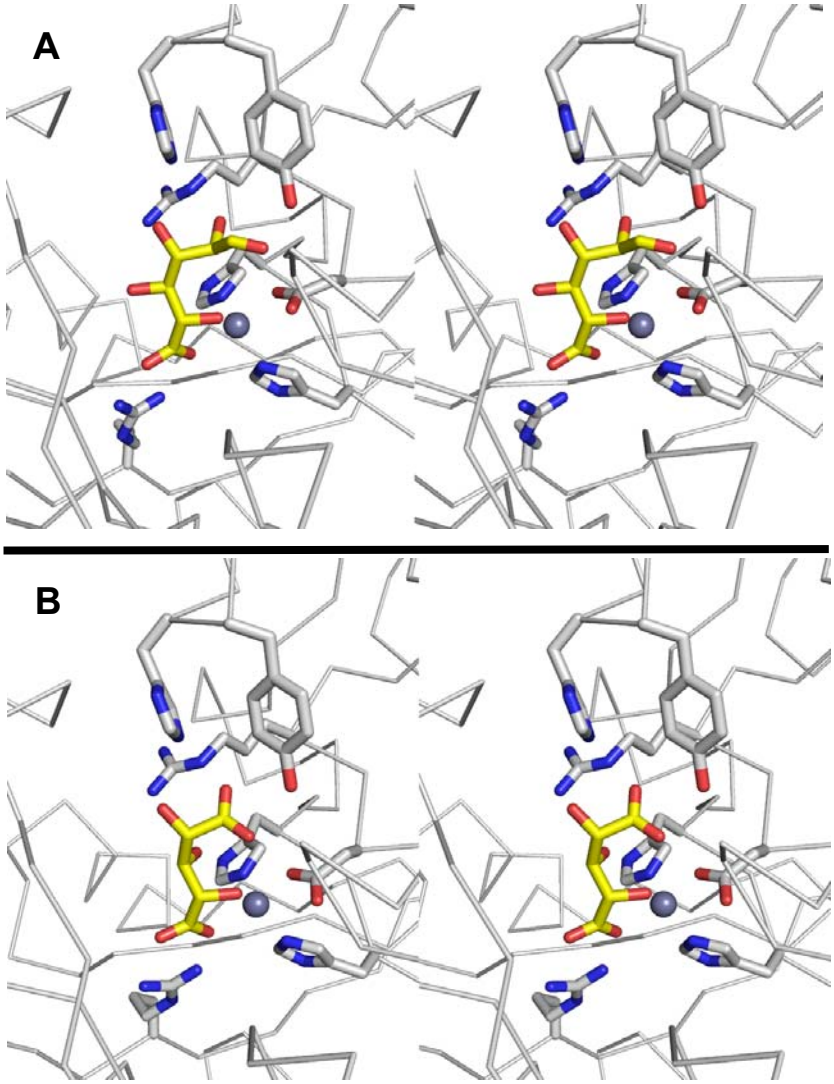


Figure 3.6: Stereoview images of Bh0493 in the presence of bound (**A**) D-glucuronate (**I**), (**B**) hydroxamate of arabinarate (**II**), (**C**) D-arabinarate (**III**), and (**D**) D-fructuronate (**IV**). This figure was created with PyMOL v0.99.

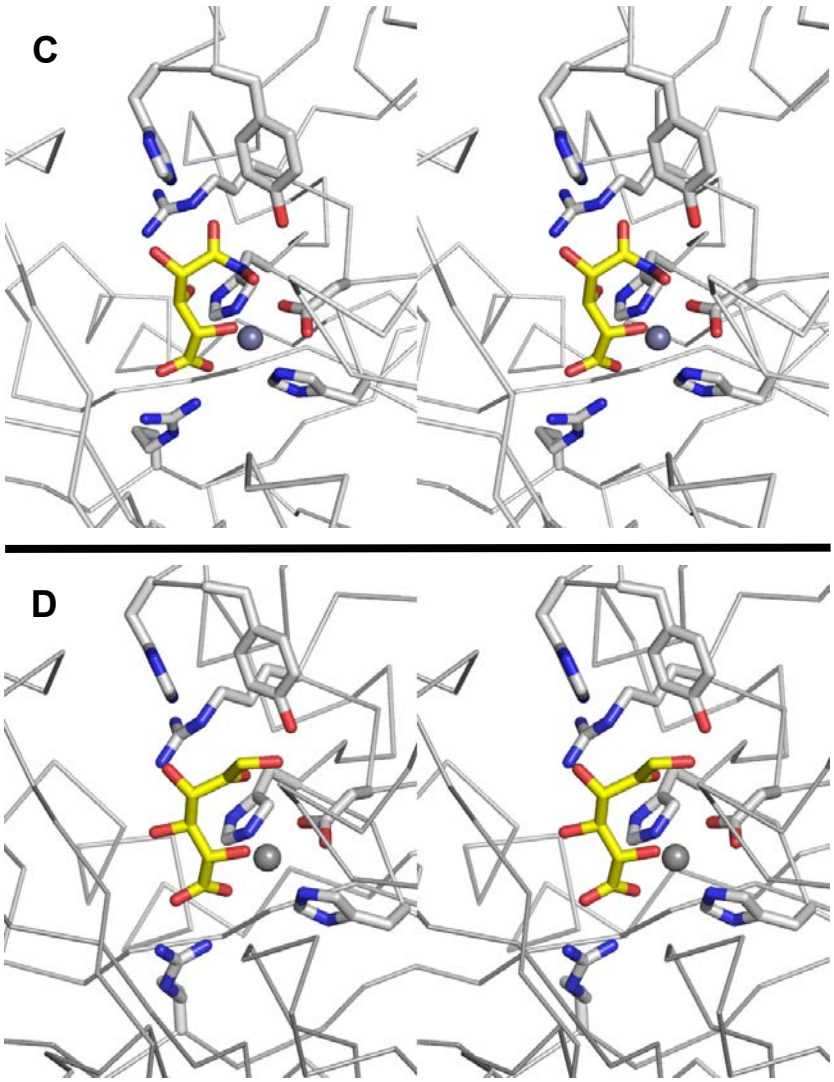


Figure 3.6: Continued.

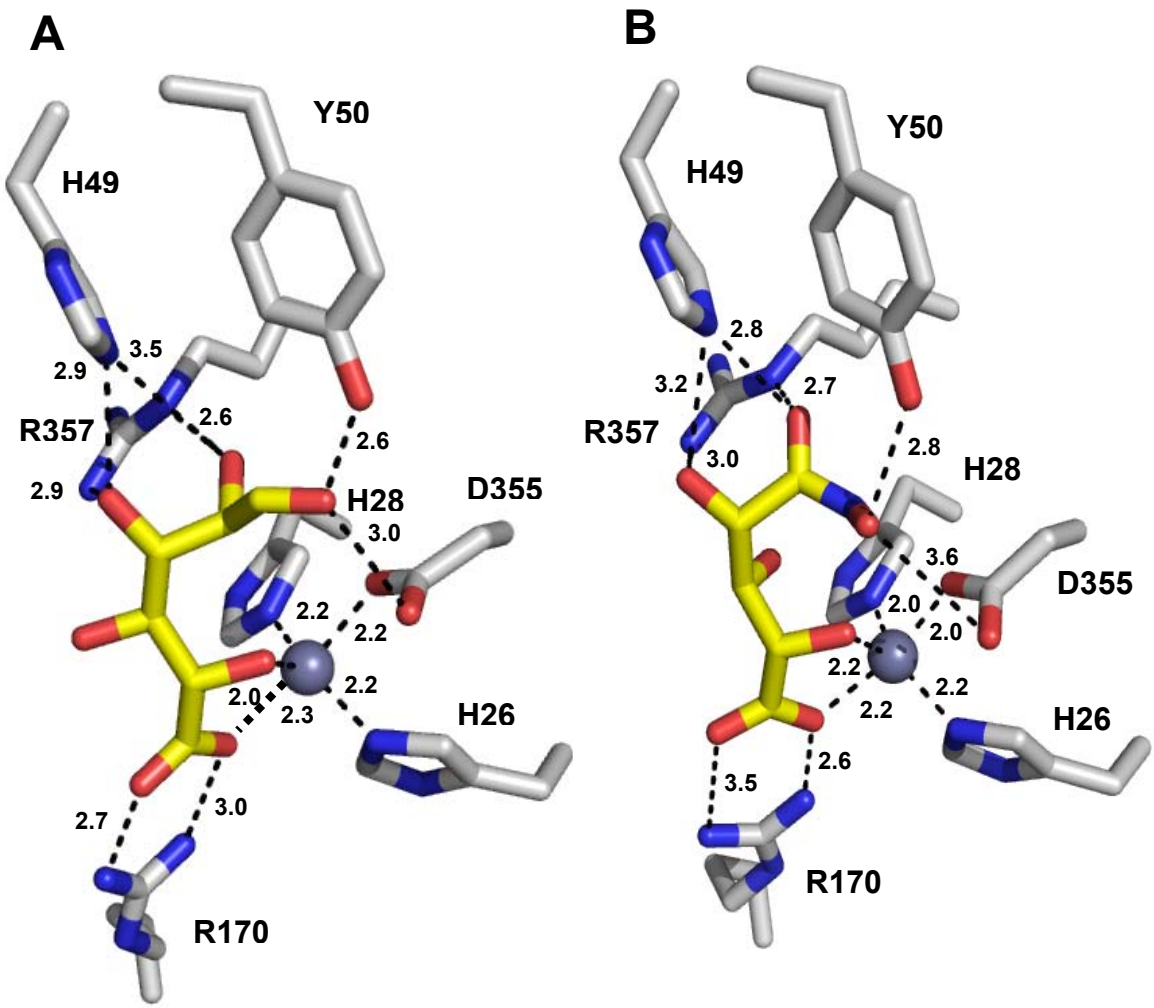


Figure 3.7: Active site of Bh0493 with (**A**) D-glucuronate, (**B**) hydroxamate of D-arabinarate, (**C**) D-arabinarate, and (**D**) D-fructuronate. Interactions between the enzyme and ligand with the distances are shown.

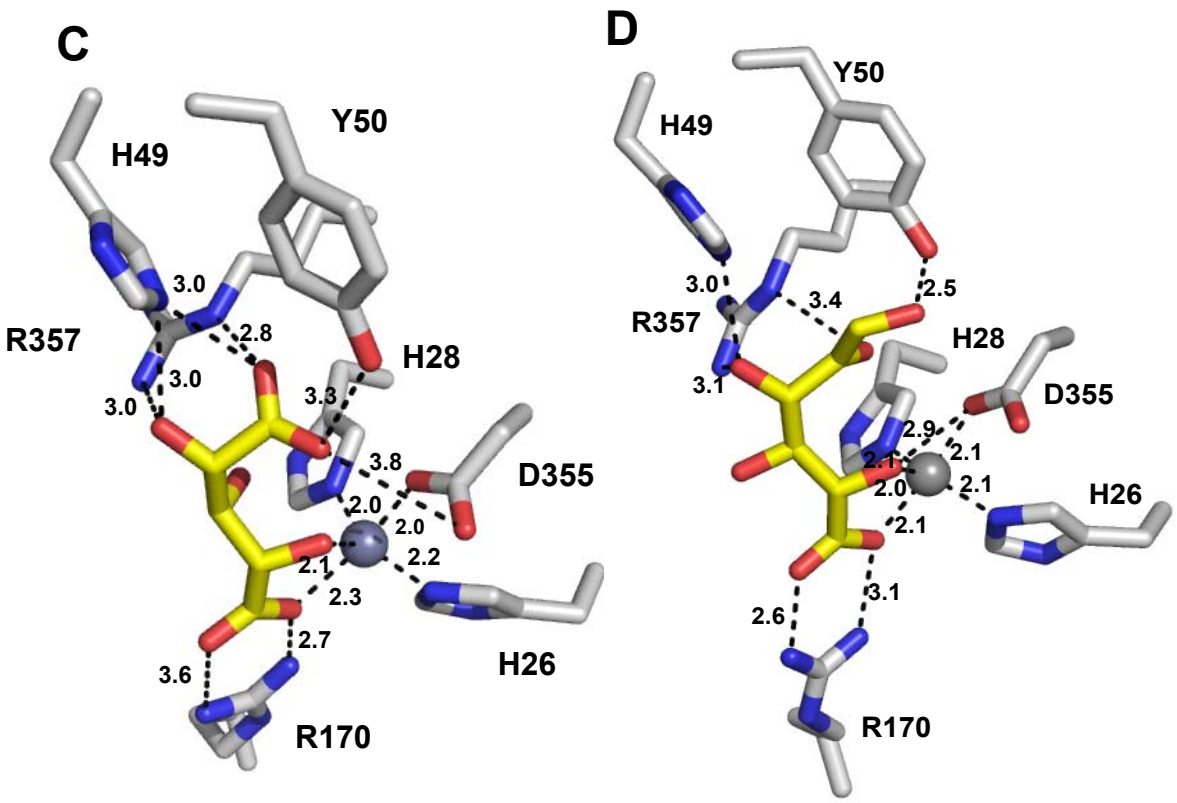


Figure 3.7: Continued.

DISCUSSION

The uronate isomerase from *E. c*oli has previously been shown to catalyze the isomerization of D-glucuronate and D-galacturonate to D-fructuronate and D-tagaturonate, respectively (*120*, *121*). With D-glucuronate the hydrogen at C-2 was shown to be transferred to the *pro*-R position at C-1 of the product (*98*). This hydrogen exchanges with solvent at a rate that is 4 orders of magnitude slower than the net interconversion of substrate and product. These results were interpreted to be consistent with a reaction mechanism that was initiated by proton abstraction at C-2, formation of a

cis-enediol(ate) intermediate and subsequent reprotonation at C-1 (*98*). This minimal reaction mechanism thus requires at least two different residues within the active site for these proton transfers. A general base (B_1:) is required for the removal of the proton at C-2 and delivery to C-1 and a general acid (B_2:H) is needed to shuttle a proton between the oxygens attached to C-1 and C-2 of the substrate/product pair. This transformation has been summarized in **Scheme 3.2**. A requirement for a minimum of two amino acids that must be in a specific state of protonation is experimentally supported by the measurement of the effects of pH on the magnitude of k_{cat} and k_{cat}/K_m for the conversion of D-glucuronate to D-fructuronate. The pH-rate profiles for both k_{cat} and k_{cat}/K_m are bell-shaped and indicate that one residue must be unprotonated and another protonated for catalytic activity. In these profiles the general base has a kinetic pK_a between 5.5 and 5.8 whereas the general acid has a kinetic pK_a between 9.5 and 10.2. Candidates for these residues were identified through the elucidation of the three-dimensional crystal structure of a uronate isomerase from *B. halodurans* (Bh0493) in the presence of D-glucuronate, D-fructuronate, and two mimics of the *cis*-enediol(ate) intermediate.

The crystal structure of Bh0493 determined with D-glucuronate has identified those residues in the active site that interact directly with the substrate. In this structure the zinc is ligated by three amino acids from the protein: the two conserved histidine residues from β-strand 1 and the aspartate from β-strand 8. The C-6 carboxylate group of the substrate is ligated to the zinc and also ion-paired with Arg-170. The hydroxyl group from C-5 is ligated to the zinc in the active site and hydrogen bonded to the aspartate from β-strand 8. At the other end of the substrate, Arg-357 hydrogen bonds

with the two hydroxyls from C-3 and C-2. The phenolic group of Tyr-50 forms a hydrogen bond with the carbonyl group at C1. The closest residue from the protein to the hydrogen at C-2 that must be abstracted during the chemical transformation is Asp-355 at 3.15 Å. In addition, His-49 is within hydrogen bonding distance to the hydroxyl at C-2 in the hydroxamate inhibitor (**II**) but not in the complex with the bound D-glucuronate (**I**) Thus, the most likely residues that are required for the isomerization reaction (in Bh0493) are His-49, Tyr-50, Asp-355 and Arg-357. In the *E. coli* enzyme these residues are equivalent to His-59, Tyr-60, Asp-412, and Arg-414, respectively. These residues, in addition to His-33 and His-35 (ligands to the zinc) Arg-186 (equivalent to R170 in Bh0493), His-297 (a conserved histidine at the end of β-stand 5 in most members of the amidohydrolase superfamily), Arg-302 and Trp-381, were mutated as probes of functional significance.

The mutation of specific residues within the active site of URI results in significant perturbations to the magnitude of the kinetic constants for substrate turnover. Changes to either of the two histidine residues that originate from the end of β-strand 1 weakens the binding of zinc to the active site and this results in a diminution of catalytic activity. This observation is consistent with the proposed role of zinc in the direct ligation of the substrate through the C-6 carboxylate and the hydroxyl from C-5. A drastic reduction in the affinity of the substrate occurs with the mutation of Arg-186 to methionine. In this case the K_m for the substrate increases by nearly two orders of magnitude and the value of k_{cat}/K_m is reduced by more than three orders of magnitude. This result is consistent with an ion-pair interaction between the C-6 carboxylate and the

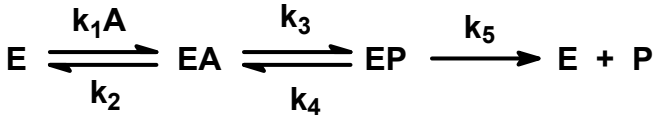
guanidino group of Arg-186 (R170 in Bh0493) that is observed in the X-ray structure of Bh0493. There is also a substantial increase in the K_m for D-glucuronate when His-297 is mutated to alanine or asparagine. This residue originates from the end of β-strand 5 and is highly conserved in nearly all members of the amidohydrolase superfamily but this residue is not conserved in Bh0493 and thus it is not easy to discern the effect on the structure of URI from *E. coli*.

The most dramatic reductions in k_{cat} occur with the mutation of Asp-412 to alanine or Arg-414 to methionine. Arg-414 is equivalent to Arg-357 from Bh0493, which is hydrogen bonded to the C-2 and C-3 hydroxyls of the bound D-glucuronate. Therefore, this residue is a suitable candidate for assisting in the movement of the proton from the hydroxyl at C-2 during the transformation to D-fructuronate. The other candidate for this process is the phenolic side chain from Tyr-60. This group is hydrogen bonded to the carbonyl oxygen at C-1 of D-glucuronate in the X-ray crystal structure. Mutation of Tyr-60 to phenylalanine results in reductions in k_{cat} and k_{cat}/K_m of about an order of magnitude. The only residue from the protein that is suitably positioned to function as the general base for the abstraction of the proton from C-2 and delivery to C-1 is Asp-412. Many members of the amidohydrolase superfamily that catalyze hydrolytic reactions have been shown to use this residue from the end of β-strand 8 to abstract a proton from the hydrolytic hydroxide or water (*60*, *68*, *70*, *71*).

Changes in solvent viscosity and isotopic substitution were used to address the degree of rate limitation on the bond breaking steps and product release (*123*, 1*24*, *126*, *127*). A simplified kinetic mechanism for the conversion of substrate to product for the

uronate isomerization reaction is presented in **Scheme 3.4**, where EA represents the enzyme-glucuronate complex and EP represents the enzyme-fructuronate complex. In this mechanism the expression for k_{cat} is given by $(k_3k_5)/(k_3 + k_4 + k_5)$. If one assumes that k_5 is inversely proportional to the relative solvent viscosity then the value of k_5, relative to the sum of k_3 and k_4 can be determined from a plot of the ratio of $^{o}(k_{cat})/^{\eta}(k_{cat})$ as a function of the relative solvent viscosity, η. The slope of this plot is equal to $(k_3 + k_4)/(k_3 + k_4 + k_5)$. For the wild-type enzyme the slope was found to be ~0.7 and thus the sum of k_3 and k_4 is greater than the product release step, k_5. This result is consistent with the release of the product as the rate limiting step for the wild type enzyme and the relatively small primary isotope effect for [2-^{2}H]-D-glucuronate. With the D412N and R412M mutants the slope of this plot, for changes in the relative value of k_{cat} as a function of solvent viscosity, is reduced substantially. This result is consistent with a significant reduction in the rate constants for the interconversion of the substrate/product pair to the point where k_5 is now greater than the sum of k_3 and k_4. This result is also consistent with the significant increase in the value of $^{D}k_{cat}$. With the D412N and R414M mutants, the primary isotope effects are 2.0 and 3.2, respectively, and thus with these two mutants the interconversion of the substrate/product pair is substantially rate limiting.

Scheme 3.4: A simplified kinetic mechanism for the conversion of substrate to product for the uronate isomerization reaction.



Mechanism of Action. Based upon the X-ray crystal structure of Bh0493 in the presence of the bound substrate, the catalytic properties of selected site-directed mutants and the stereochemical constraints for the conversion of D-glucuronate to D-fructuronate, a minimal chemical mechanism can be written for uronate isomerase. In the proposed mechanism for URI from *E. coli*, D-glucuronate is bound in the active site through electrostatic interactions to five highly conserved amino acid residues and the divalent cation. The carboxylate group at C-6 is coordinated to the divalent cation and Arg-186. The hydroxyl group at C-5 is also coordinated to the zinc. The hydroxyl groups at C-3 and C-2 interact with the side chain guanidino group of Arg-414 and the hydroxyl at C-2 is also apparently able to hydrogen bond to His-59. The carbonyl group at C-1 is hydrogen bonded to the phenolic oxygen of Tyr-60. The pH-rate profiles for URI are consistent with two amino acid residues that must be in a specific state of protonation for catalytic activity. The general base, with a kinetic pK_a of approximately 5.8, is consistent with Asp-412. This residue is conserved in all members of the amidohydrolase superfamily and for those enzymes that catalyze hydrolytic reactions it has been shown to initiate proton transfers from water/hydroxide to the leaving group

(*60*, *68*, *70*, *71*). The general acid, with a kinetic pK_a of approximately 10.2 from the pH-rate profiles, may be due to Tyr-60 but it is difficult to exclude a role for Arg-414.

In the simplest mechanism D-glucuronate binds in the open chain conformation in the active site and then Asp-412 abstracts the proton from C-2 as the carbonyl group of C-1 is protonated by Tyr-60 to form a *cis*-enediol intermediate. In the subsequent step, Tyr-60 abstracts the proton from the hydroxyl at C-2 as Asp-414 delivers a proton to C-1 with *pro*R stereochemistry. It should be noted that the mutation of Tyr-60 to alanine or phenylalanine diminishes k_{cat} by only a factor of 10. This reduction in rate is perhaps smaller than what may be expected for this role in catalysis. However, the lack of a primary deuterium isotope effect for the wild-type enzyme indicates that the chemical step is not rate-limiting. In addition, a water molecule may substitute for the phenolic group in the mutant enzymes. Additional mechanisms can be written that utilize a combination of Tyr-60, Arg-414, and His-59 to facilitate the proton movements between the oxygens at C-2 and C-1 within the *cis*-enediol intermediate and in the opening of the hemiacetal.

An alternative mechanism can be proposed in which Asp-412 abstracts the proton from the hydroxyl at C-5, which then abstracts a proton from C-2 to initiate the formation of the *cis*-enediol intermediate. The hydroxyl at C-5 is additionally activated through direct ligation to the bound zinc. These two variations are presented graphically in **Figure 3.8.** This latter mechanism is particularly attractive since it retains elements that are common to nearly all members of the amidohydrolase superfamily that have been interrogated mechanistically (*60*, *68*, *70*, *71*). The structural similarities in the

active site of uronate isomerase with those members of the amidohydrolase superfamily nicely illustrates the evolutionary link between those members of the AHS that catalyze hydrolytic reactions and those that catalyze 1,2-proton transfers. A structural alignment of the active sites of Bh0493/D-glucuronate and dihydroorotase (DHO)/dihydroorotate (PDB entry code 1j79) supports this proposition as illustrated in **Figure 3.9**. In this structural alignment one of the C-6 carboxylate oxygens from D-glucuronate is positioned in nearly the same place as the carboxylate oxygen of the bridging carbamate functional group in DHO. These oxygen atoms interact directly with the alpha-metal (M_α) in their respective structures. Moreover, the C-5 oxygen of D-glucuronate is positioned in the same way as the nucleophilic hydroxide in DHO and is oriented to favor proton abstraction by Asp-412 at a distance of 2.9 Å. The C-5 hydroxyl is 2.0 Å away from the hydrogen at C-2 of the bound substrate.

In addition to the crystal structures of Bh0493, the structures of uronate isomerase from *T. maritima* (Tm0064) and *C. crescentus* (Cc1490) are also available (PDB codes: 1j5s and 2q01, respectively). The three-dimensional structures of Bh0493, Tm0064 and Cc1490 are similar to each other despite the low sequence identities between Bh0493 and the other two proteins. All three proteins exist as a homotrimers. One significant difference in the three structures is that Bh0493 has a zinc ion in the active site, while the structure of Tm0064 contains electron density that corresponds to a water molecule that is in the same position as the zinc ion in the Bh0493 structure (*43*). The Cc1490 structure contains no metal ion or water molecules in the active site. Moreover, the histidine from β-strand 5 is conserved in Tm0064 and Cc1490 but not for

the Bh0493 enzyme. Despite the divergence in the amino acid sequences, the mechanism utilized by uronate isomerase family members is expected to be the same. Although the Tm0064 and Cc1490 structures do not contain a divalent cation, the metal ligand residues are present and therefore it is expected that they are capable of binding/utilizing a divalent cation for catalysis.

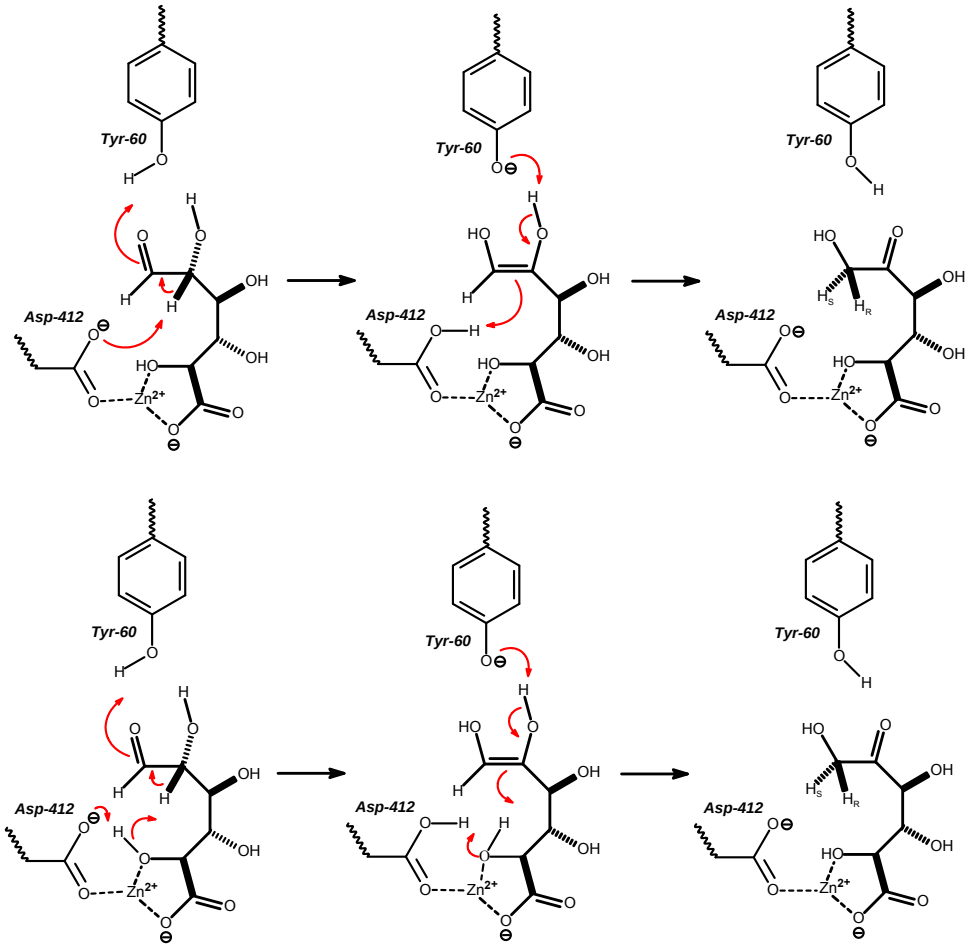


Figure 3.8: Proposed mechanisms for the isomerization of D-glucuronate by URI.

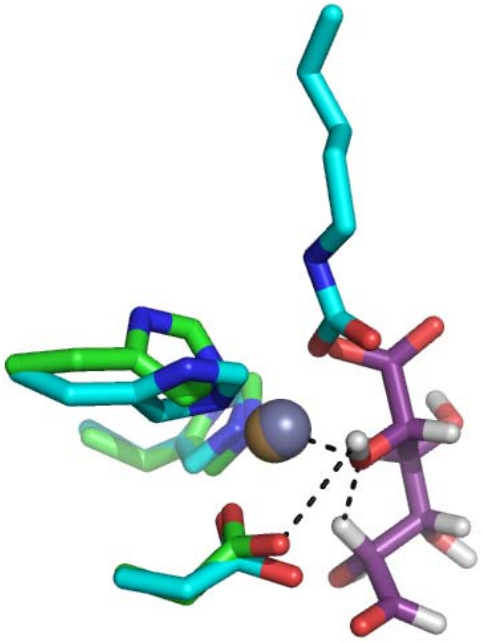


Figure 3.9: A structural alignment for portions of the active sites of Bh0493 (green) and DHO from *E. coli* (blue). The alpha metal (M_α) is shown as a yellow sphere for Bh0493 and in grey for DHO. The bound D-glucuronate for Bh0493, is shown in purple. The interactions between the C-5 oxygen of D-glucuronate and M_α, C-2 hydrogen and Asp-412 are indicated by dashed lines. The other ligands to the metal ion are shown as stick representations, including the HxH from β-strand 1 and aspartate from β-strand 8. The C-6 oxygen of D-glucuronate, aligns closely the the position occupied by one of the oxygens from the carboxyl group of the carboxylated lysine from strand 4 of DHO. The C-5 hydroxyl of D-glucuronate is orientated in the same position as the bridging hydroxide in the DHO structure.

CHAPTER IV

MECHANISTIC STUDIES OF SCO3058: A BACTERIAL RENAL DIPEPTIDASE FROM *STREPTOMYCES COELICOLOR*

The amidohydrolase superfamily (AHS) contains enzymes that catalyze primarily hydrolysis reactions of amide and ester bonds of sugars, amino acids, nucleic acids, and organophosphate ester substrates. A few rare families of the anidohydrolase superfamily catalyze nonhydrolytic reactions including isomerization and decarboxylation reaction (*21*,*31*). Structurally, all members of the AHS comprise of an $(\beta/\alpha)_8$-TIM barrel fold with a variation of a mononuclear or a binuclear metal center active site located at the C-terminal end of the beta strands. The metal center functions to activate the hydrolytic water as well as the scissile bond of the substrate for nucleophilic attack and to stabilize the resulting tetrahedral or trigonal bipyramidal transition state (*31*). Holm and Sander first discovered the amidohydrolase superfamily in 1997 based on the structural similarities among urease, adenosine deaminase, and phosphotriesterase (PTE) (*21*). Since then, more than 10,000 members of the AHS have been identified in prokaryotes and eukaryotes, and these enzymes have been shown to be involved in the catalysis of at least 40 different reactions.

One of the most common divalent cation centers within the AHS is the binuclear metal center found in enzymes such as urease, PTE, dihydroorotase, and isoaspartyl dipeptidase (*32-35*). In these proteins, the more buried metal, designated as M_α, is coordinated to the two histidines from the HxH motif at the end of β-strand 1 and an

aspartate from strand 8. M_β, the more solvent exposed metal, is ligated to the two histidines from strands 5 and 6. The two metals are bridged by a solvent hydroxide as well as a carboxylated lysine or a glutamate that originates from strand 4.

Another distantly related binuclear member of the AHS is the renal dipeptidase from mammals. This family of enzymes was first identified and characterized in the 1970s based on its physicochemical properties (*100*). In addition to a wide range of dipeptides, renal dipeptidase is also involved in the renal metabolism of glutathione and its conjugates and is the sole enzyme that is responsible for the hydrolysis of penem and carbapenem β-lactam antibiotics (*101-106*). The crystal structure of the human renal dipeptidase was solved in the presence of the inhibitor cilastatin (PDB entry code: 1itu) (*40*). The quarternary structure of the hRDP is a homodimer with each polypeptide subunit consisting of 369 amino acid residues (42 kDa). The sequence of the hRDP shows a 75% identity to its homologs from rat and pig. From the crystal structure, the protein utilizes a HxD motif at β-strand 1 to bind to the M_α zinc ion instead of the more conventional HxH. Moreover, the residue that bridges the two metals is now a glutamate from β-strand 3 and not strand 4. The aspartate at the end of β-strand 8 is conserved but at a distance of 3.8 Å away, it is no longer coordinating to the M_α. The metal center of renal dipeptidases is illustrated in **Figure 4.1**. According to the X-ray structure, the aspartate residue is interacting with the hydrolytic water via a hydrogen bond. It was also suggested by Nitanai et al that the hydrolytic water is activated by the conserved aspartate from β-strand 8. The deprotonated water then nucleophilicly attacks the carbonyl carbon of the peptide bond to form a tetrahedral intermediate. It was

proposed that the positively charged His-152 acts as an oxyanion hole to stabilize the negatively charged intermediate (*40*). This mechanism of hydrolysis is distinctly different from the mechanism proposed for the enzymes in the AHS in which the tetrahedral intermediate is stabilized by the solvent exposed M_{β}.

The first renal dipeptidase-like gene *acdp* from a prokaryote was cloned from *Acinebacter calcoaceticus* (*107*). This enzyme has a low sequence identity of 23% to the human renal dipeptidase, but the essential residues identified in the mammalian dipeptidases are all conserved. Unlike its mammalian homologs, the gene product of *acdp* could not hydrolyze the unsaturated dipeptide glycyldehydrophenylalanine; instead the protein showed substrate specificity towards various dipeptides with a preference for dipeptides contaiting a D-amino acid at the C-terminus (*107*). In this chapter, the kinetic and structural properties of another bacterial renal dipeptidase-like enzyme from *Streptomyces coelicolor* are described. According to the sequence alignment shown in **Figure 4.2**, this protein is 45% and 22% identical in sequences to the hRDP and *acdp*, respectively. Substrate specificity analysis of Sco3058 against 55 dipeptide libraries indicated promiscuous dipeptidase activity with a preference for the hydrolysis of dipeptides with D-Leu, D-Met, D-Arg, and D-Lys at the N-terminus and either L- or D-Asp at the C-terminus. Overall, this enzyme is promiscuous towards a wide range of dipeptides containing L- and D-amino acids but the libraries with L-Xaa D-Xaa dipeptides were hydrolyzed the fastest. The best substrate tested was the dipeptide L-Arg-D-Asp, with kinetic constants of 68 s^{-1} for k_{cat} and (4.1 ± 0.5) x 10^5 for k_{cat}/K_m. The hydrolysis reaction catalyzed by Sco3058 is shown in **Scheme 4.1**.