Summer 8-12-2014

Structural and Mechanistic Studies on α-Amino β-Carboxymuconate ε-Semialdehyde Decarboxylase and α-Aminomuconate ε-Semialdehyde Dehydrogenase

Lu Huo
Georgia State University

STRUCTURAL AND MECHANISTIC STUDIES ON α-AMINO β-CARBOXYMUCONATE ε-SEMIALDEHYDE DECARBOXYLASE AND α-AMINOMUCONATE ε-SEMIALDEHYDE DEHYDROGENASE

by

LU HUO

Under the Direction of Dr. Aimin Liu

ABSTRACT

α-Amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) and α-aminomuconate-ε-semialdehyde dehydrogenase (AMSDH) are two neighboring enzymes in the L-tryptophan and 2-nitrobenzoic acid degradation pathways. The substrates of the two enzymes, α-amino-β-carboxymuconate-ε-semialdehyde (ACMS) and α-aminomuconate-ε-semialdehyde (2-AMS), are unstable and spontaneously decay to quinolinic acid and picolinic acid, respectively.

ACMSD utilizes a divalent zinc metal as cofactor and is a member of the amidohydrolase superfamily. In this dissertation work, we have identified an important histidine residue in the active site that plays dual roles in tuning metal selectivity and activating a metal bound water ligand using mutagenesis, resonance Raman, EPR, crystallography, and ICP metal analysis techniques. The crystal structures of ACMSD from *Pseudomonas fluorescens* (*Pf*ACMSD) have been solved as homodimers in our laboratory while human ACMSD (*h*ACMSD) was annotated

as a monomer by another group. To resolve this structural difference, we used two conserved active site arginine residues as probes to study the oligomeriztion state of ACMSD and demonstrated that these two arginine residues are involved in substrate binding and that both *Pf*- and *h*- ACMSD are catalytically active only in the dimeric state. Subsequently, we solved the crystal structure of *h*ACMSD and found it to be a homodimer in both catalytically active and inhibitor-bound forms.

AMSDH is an NAD^+ dependent enzyme and belongs to the aldehyde dehydrogenase superfamily. Due to the high instability of its substrate, AMSDH has not been studied at the molecular level prior to our work. We have cloned and expressed *Pf*AMSDH in *E. coli*. The purified protein has high activity towards both 2-AMS and 2-hydroxymuconate semialdehyde (2-HMS), a stable substrate analog. We have successfully crystallized AMSDH with/without NAD^+ and solved the crystal structure at up to 1.95 Å resolution. Substrate bound ternary complex structures were obtained by soaking the NAD^+ containing crystals with 2-AMS or 2-HMS. Notably, two covalently bound catalytic intermediates were captured and characterized using a combination of crystallography, stopped-flow, single crystal spectroscopy, and mass spectrometry. The first catalytic working model of AMSDH has been proposed based on our success in structural and spectroscopic characterization of the enzyme in five catalytically relevant states in this dissertation work.

INDEX WORDS: α-Amino β-carboxymuconate ε-semialdehyde decarboxylase, α-Aminomuconate ε-semialdehyde dehydrogenase, Quinolinic acid, Amidohydrolase superfamily, Crystallography, Oligomeriztion, Reaction intermediate.

STRUCTURAL AND MECHANISTIC STUDIES ON α-AMINO β-CARBOXYMUCONATE ε-SEMIALDEHYDE DECARBOXYLASE AND α-AMINOMUCONATE ε-SEMIALDEHYDE DEHYDROGENASE

by

LU HUO

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2014

STRUCTURAL AND MECHANISTIC STUDIES ON α-AMINO β-CARBOXYMUCONATE ε-SEMIALDEHYDE DECARBOXYLASE AND α-AMINOMUCONATE ε-SEMIALDEHYDE DEHYDROGENASE

by

LU HUO

Committee Chair: Aimin Liu

Committee: Binghe Wang

Jenny Yang

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

August 2014

DEDICATION

I dedicate my dissertation work to my dear parents, Ms. Qiuju Dong and Mr. Huimin Huo, who supported me each step of the way.

ACKNOWLEDGEMENTS

This dissertation would not have been completed without the help and support from a number of people over the past four and half years. I wish to take this chance to show my gratitude to all of them.

I would like to express my deepest appreciation to my advisor, Professor Aimin Liu, for his guidance and constant support. I thank Dr. Liu for introducing me to the field of biochemistry with his enthusiasm for science, providing me with all the best research and learning opportunities with his care for students, preparing me for future challenges with his profound knowledge and life experience.

I wish to thank Dr. Jenny Yang and Dr. Binghe Wang for serving on my dissertation committee. I thank Dr. Yang for setting a good example to me as a great female scientist. I sincerely thank both Dr. Wang and Dr. Yang for their insight and guidance during my annual meeting and thesis preparation. I want to express my thanks to Dr. Gabor Patonay, Dr. Giovanni Gadda, and Dr. Siming Wang, who advised me inside and outside of the Bioanalytical Chemistry, Enzymology and Mass Spectrometry courses.

I want to thank all my dear friends in the lab. I thank Mr. Jiafeng Geng for training me when I first joined the lab, for being a good friend and a knowledgeable senior co-worker who provides me with valuable suggestions whenever I ask him for opinion. I thank Mr. Ian Davis, my neighbor and teammate in the lab, who cannot be more dependable as a lab manager and more insightful as an internal collaborator. I thank Mr. Kednerlin Dornevil, who is so warm hearted and always ready to help. I want to thank Mr. Shingo Esaki and Mr. Ryan Gumpper for their cooperation in experiments. I thank Dr. Vesna de Serrano and Dr. Lirong Chen for teaching me crystallography from setting up trays, mounting crystals, collecting and processing data, to

data analysis, step by step. I wish to thank all other lab members who have helped me in my Ph.D. project.

I would like to thank all my collaborators for helping me with the work in this dissertation. I thank Dr. Yoshie Hasegawa and Dr. Hiroaki Iwaki at Kansai University, Japan for providing the expression systems of *Pf*ACMSD, *h*ACMSD and *Pf*AMSDH. I thank Dr. Lawrence Que and Dr. Andrew Fielding at University of Minnesota, USA for performing resonance Raman experiment on H228Y ACMSD. I want to thank Dr. Allen Orville and Dr. Babak Andi at Brookhaven National Laboratory, USA for performing single crystal spectroscopic experiments on AMSDH. I also want to thank the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID and 22-BM beamlines staff scientists for assistant with remote data collections.

I acknowledge the Molecular Basis of Disease Area of Focus program for the generous supports, including the fellowship, travel award, and the great honor. I also want to thank Ms. Rita Bennett and Mr. Will Lovett, the graduate coordinators in Chemistry Department, for their help in the past few years.

Finally, to my beloved family, their love is the motive force that keeps me going forward. I must thank my parents, Ms. Qiuju Dong and Mr. Huimin Huo, for their dedication, unconditional love and supports. Although we are 12200 kilometers apart, I can always feel the deepest and warmest love from them. I also want thank my husband and best friend, Mr. Xiaochen Lin, for his encouragement, inspiration, help, and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

Part of chapter 1 has been published in our two review papers: Amidohydrolase Superfamily. Aimin Liu, and Lu Huo. (2014) Els. John Wiley & Sons, Ltd: Chichester. DOI: 10.1002/9780470015902.a0020546.pub2. Decarboxylation mechanisms in biological system. Tingfeng Li, Lu Huo, Christopher Pulley, and Aimin Liu. (2012) *Bioorg. Chem*., 43, 2-14.

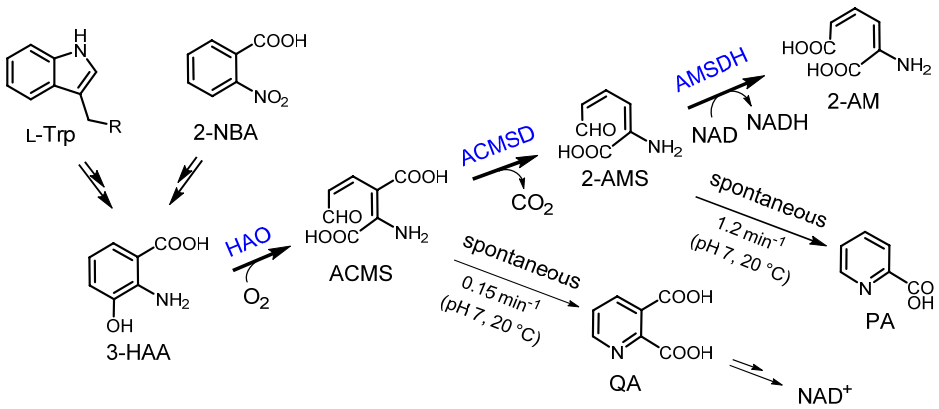
1.1 The Kynurenine Pathway

As the major route for L-tryptophan (TRP) catabolism, which takes place in the liver, kidney, and brain of mammals, the kynurenine pathway (KP) has been found to be highly related with numerous disease states, and thus has gained significant attention from the scientific field. Studies in the past few decades have shown that the KP is a powerful regulator of neurodegenerative disorders: Alzheimer's disease, Huntington's disease, anxiety, depression, and epilepsy (1-7). This is because several neurologically active compounds including kynurenic acid, 2-hydroxykynurenine, and quinolinic acid (QUIN) are metabolites generated by the KP. QUIN is an agonist of *N*-methyl-D-aspartate receptors and could cause overexcitement of neurons and cell death in the central nervous system (2,5,8,9). The precursor of QUIN is α-amino-β-carboxymuconate-ε-semialdehyde (ACMS), which is an unstable metabolic intermediate and spontaneously decays to QUIN with a $t_{1/2}$ of 33 min at 20 °C, pH 7.0 (10). The formation of QUIN *in vivo* is controlled by the enzyme ACMS decarboxylase (ACMSD) because the decarboxylation reaction it catalyzes competes with the decay of ACMS. The product of ACMSD is α-aminomuconate-ε-semialdehyde (2-AMS), which is even more unstable than ACMS and decays to picolinic acid (PIC) with a $t_{1/2}$ of 9 s at pH 7.4, 37 °C. The biological role of PIC remains controversial. Several studies have shown various physiological effects when using millimolar concentrations of PIC; however endogenous concentrations have been

measured to be in the sub-micromolar range. The formation of PIC is controlled by the downstream enzyme 2-AMS dehydrogenase (AMSDH) which oxidizes the unstable aldehyde to a stable acid, 2-aminomuconic acid (Scheme 1.1). This dissertation focuses on structural and mechanistic studies of two consecutive enzymes of the KP, ACMSD and AMSDH.

1.2 The 2-Nitrobenzoic Acid Degradation Pathway

The toxic organic compound 2-nitrobenzoic acid (2-NBA) is biodegradated by a bacterial strain KU-7, identified as *Pseudomonas fluorescens* (11), which can utilize 2-NBA as the sole source of carbon, nitrogen, and energy (11). The initial two steps catalyze the transformation of 2-NBA to 3-HAA, at which point the 2-NBA degradation pathway "borrows" part of the kynurenine pathway as shown in Scheme 1.1. These two pathways share three enzymes, non-heme Fe^{2+} dependent 3-hydroxyanthranilinic acid 3,4-dioxygenase (HAO), ACMSD, and AMSDH, and also the mechanism to regulate nonenzymatic production of QA and PA. This discovery suggests that the genes for these three enzymes of the kynurenine pathway are evolutionarily conserved genome segments which are important for at least two independent catabolic pathways. Before our work, human ACMSD can only be expressed as insoluble inclusion body in *E. coli*, and human AMSDH had not been identified. ACMSD and AMSDH



Scheme 1.1 The L-tryptophan and 2-NBA degradation pathways share an enzyme trio system.

from the 2-NBA pathway had been used as model enzymes for structural and mechanistic studies before our recent progresses on the human analog enzymes.

1.3 ACMSD is a Member of the Amidohydrolase Superfamily

ACMSD was initially thought to be a cofactor-free enzyme (12,13). A comprehensive *in vitro* study on ACMSD from *Pseudomonas fluorescens* (*Pf*ACMSD) demonstrated that this enzyme is actually critically dependent on transition metal ions (14). Further sequence analysis and mutagenesis studies suggest that ACMSD belongs to the amidohydrolase superfamily (15), which is a structure-based cluster of metalloenzymes that contain a sturdy and versatile triosephosphate isomerase (TIM)-like $(\beta/\alpha)_8$-barrel fold embracing the catalytic active site. Later on, the crystal structure of *Pf*ACMSD was solved with a zinc metal center and a TIM-barrel overall fold, supporting that ACMSD is a new member of the amidohydrolase superfamily.

1.3.1 The concept of amidohydrolase superfamily

Up to date, the amidohydrolase superfamily has grown into one of the largest families of enzymes, with tens of thousands of members catalyzing a wide range of hydrolytic and nonhydrolytic metabolic reactions which are important in amino acid and nucleotide metabolism as well as biodegradation of agricultural and industrial compounds. Previously, presence of a mono- or di-nuclear *d*-block metal cofactor in the active site was thought to be one of the main characters for the members in this superfamily. Recently, new members that contain trinuclear metal cofactor or no cofactor were also discovered. It has become evident that activating a well-ordered water molecule by an active site residue for nucleophilic attack on the organic substrate is a common mechanistic feature for all members of the superfamily.

Amidohydrolase superfamily is one of the largest enzyme families. The concept of the ‘amidohydrolase superfamily’ was introduced by Holm and Sander. The sbtriking similarities of

the three-dimensional structures of adenosine deaminase (ADA), phosphotriesterase (PTE) and urease (URE) inspired the unification of a broad set of *d*-block metal-dependent hydrolase enzymes into a unique enzyme superfamily (16). The amidohydrolase superfamily is more than just a group of hydrolase enzymes. It is a functionally diverse enzyme group. Members of this enzyme superfamily catalyze the cleavage of not only C–N but also C–C, C–O, C–Cl, C–S and O–P bonds of organic compounds. Most of the characterized members of the amidohydrolase superfamily are important enzymes for histidine, tryptophan and lignin derived compounds metabolism, *de novo* biosynthesis of purine and pyrimidine nucleotides, and biodegradation of agricultural and industrial materials including rubber chemicals, herbicides, leather, paper and others. Some members of this enzyme superfamily are medically relevant. For instance, ADA catalyzes ammonia elimination from the heterocyclic nitrogenous base of the substrate. The deficiency of this enzyme is linked to a common form of severe combined immunodeficiency (SCID). Since the superfamily is not function-designated, an enzyme with amidohydrolase in its name is not necessarily a member of the amidohydrolase superfamily. For instance, penicillin amidohydrolase (also known as penicillin acylase, EC 3.5.1.11) contains neither a catalytic metal ion nor a TIM-barrel domain; thus, it is not a member of the amidohydrolase superfamily.

The amidohydrolase superfamily contains more than five related subfamilies in the Pfam database annotation (17). The Amidohydrolase_1 subset (accession number: PF01979) contains a large group of 18,418 protein sequences (http://pfam.janelia.org). Its members catalyze the hydrolysis of a wide range of substrates bearing amide, ester, halogen or other functional groups at carbon and phosphorus centers. This family includes well-characterized enzymes such as ADA, PTE, URE, cytosine deaminase (CDA), D-amino acid deacetylase (AAD), dihydroorotase (DHO), N-acetylglucosamine-6-phosphate deacetylase (AGD) and renal dipeptidase (RDP).

Among them, DHO and URE also belong to MEROPS peptidase family M38 (β-aspartyl dipeptidase, clan MJ), where they are classified as nonpeptidase homologues.

The Amidohydrolase_2 subset (accession number: PF04909) contains 6,287 aligned protein sequences. These proteins are related to the metal-dependent hydrolases but they appear to be either structurally or functionally divergent from the Amidohydrolase_1 group. Members of this branch are known to catalyze nonhydrolytic reactions including decarboxylation and hydration (18). The prototypic member of the Amidohydrolase_2 is ACMSD (EC 4.1.1.45), which converts ACMS to 2-AMS in a nonhydrolytic C-C bond cleavage process (14). Despite the similar tertiary structures, it is not clear if members of the Amidohydrolase_2 subset have arisen from a single evolutionary origin with those in the Amidohydrolase_1 because of their very low overall sequence similarities.

The Amidohydrolase_3 subset (accession number: PF07969) consists of 5,270 protein sequences, most of which are hypothetical proteins. Some, including D-aminoacylase, formyltransferase/hydrolase complex Fhc subunit A, are amidohydrolase-like enzymes. This branch of the amidohydrolase superfamily contains the most diverse set of sequences, including a high proportion of outlier sequences that have only low levels of sequence identity to their closest superfamily relatives.

The Amindohydolase_4 subset (accession number: PF13147) consists of 10,639 protein sequences. The representative member of this subset includes allantoinase, D-hydantoinase (DYH), dihydropyrimidinase (DHPase), imidazolonepropionase (IPase), and a latent form of dihydroorotase.

The Amidohydrolase_5 subset (accession number: PF13594) is a relative small group of enzymes with 3,712 sequences to date. Isoaspartyl dipeptidase is its representative member.

More than thirty additional new subsets of amidohydrolases are being added, most of which are based on the functional distinctions. These include Creatininase (accession number: PF02633, 1,362 sequences), which is a group of urease-related amidohydrolases hydrolysing creatinine to creatine, FGase (accession number: PF05013, 1,531 sequences) which consists of a group of N-formylglutamate amidohydrolases, and TatD_DNase (accession number: PF01026, 8,430 sequences) which is a rapidly growing subgroup of DNAse.

1.3.2 Structural features of members in the amidohydrolase superfamily

Most members of the amidohydrolase superfamily consist of a central $(\beta/\alpha)_8$ barrel in which eight parallel β strands flanked on the outer face by eight α helices (Figure 1.1). Members

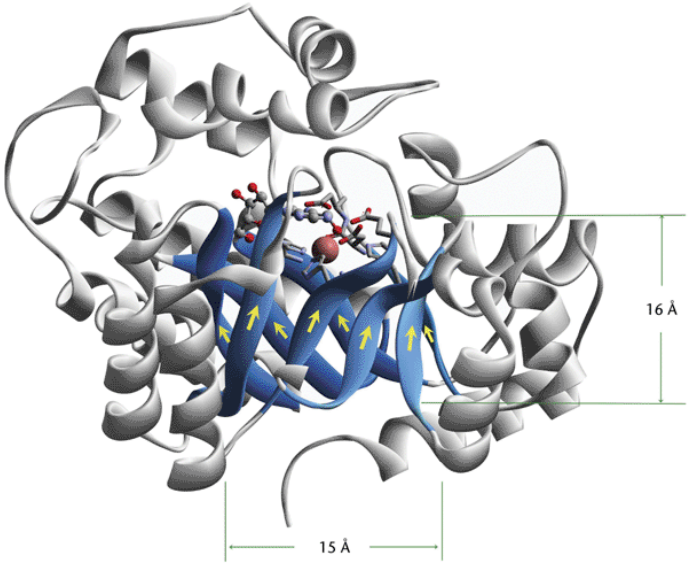


Figure 1.1. Metal cofactor and the TIM-like parallel barrel core in ADA (PDB entry 1A4M). The metal ion is depicted as CPK sphere, metal ligands are shown in sticks and the propeller structural fold is highlighted in blue colour in the representation. Substrate nalogue 6-hydroxy-1,6-dihydro purine nucleoside is represented in scaled ball and stick.

contain a trinuclear metal center instead possesses a distorted $(\beta/\alpha)_7$ barrel fold. This gives rise to a prominent architectural feature: a sturdy pocket with an internal cavity adjacent to the active site. The pocket has a propeller-like shape made by the seven or eight β strands with a depth of around 15-18 Å and a diameter of about 14-16 Å that forms a compact thermo-stable core. The

$(\beta/\alpha)_8$-barrel fold is a common structural platform found in about 10% of all proteins with known three-dimensional structures and in about 33 superfamilies in the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop). It is also known as the TIM-barrel fold because it was first defined from the three-dimensional structure of TIM. Containing a *d*-block metal cofactor with one or two metal ions conjunction with the TIM-barrel used to be another structural character for the amidohydrolase superfamily. Recently, several new members have been identified either with a trinuclear metal center or cofactor free. However, four active site residues, three histidine and one aspartate/glutamate, serve as metal ligands in the metal containing amidohydrolase superfamily proteins are still conserved in the metal independent members with new catalytic roles. Figure 1.2 shows examples of the TIM-barrel along with active site that contains no-, mono-, di- or tri- nuclear metal cofactors. There are currently 25 functionally annotated unique members of the amidohydrolase superfamily for which high-resolution X-ray crystal structures are available (Table 1.1).

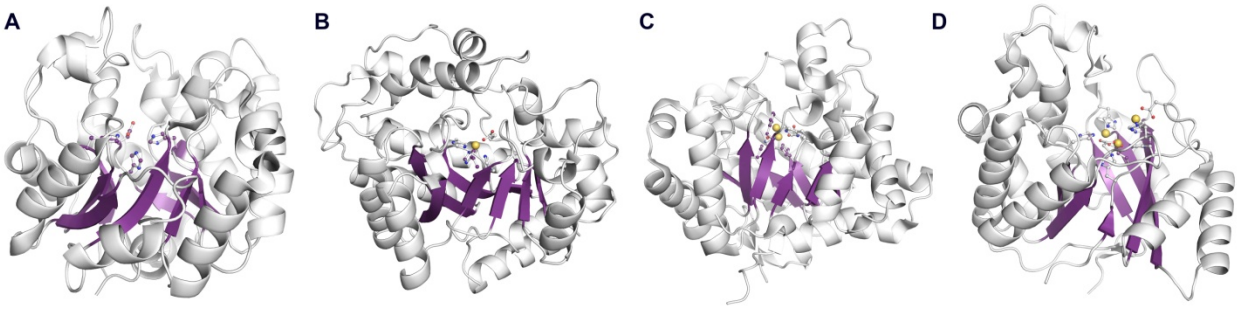


Figure 1.2. The crystal structures of TIM-barrel along with different metal center. (A) BmulJ_04915 (PDB entry 4DNM, cofactor free), (B) ADA (PDB entry 1A4M, mononuclear metal cofactor), (C) PTE (PDB entry 1HZY, dinuclear metal cofactor) and (D) HPP (PDB entry 4GC3, trinuclear metal cofactor). Metal ions are shown as yellow spheres. Metal ligandresidues are shown in sticks and the propeller structural fold is highlighted in purple.

Table 1.1. Functionally annotated and structurally characterized members of the amidohydrolase superfamily

Enzyme	Abbreviation	E.C. number	PDB code
α-Amino-β-carboxymuconate-ε-semialdehyde decarboxylase	ACMSD	4.1.1.45	2HBV
Adenosine deaminase	ADA	3.5.4.4	1A4M
Adenine deaminase	ADase	3.5.4.2	2ICS
N-acetylglucosamine-6-phosphate deacetylase	AGD	3.5.1.25	1O12
Allantoate amidohydrolase	ATase	3.5.2.5	1Z2L
Adenosine 5′-monophosphate deaminase	AMPD	3.5.4.6	2A3L
Cytosine deaminase	CDA	3.5.4.1	1K6W
D-Aminoacylase	DAA	3.5.1.81	1M7J
Dihydroorotase	DHO	3.5.2.3	1J79
Dihydropyrimidinase (L-hydantoinase)[a]	DHPase	3.5.2.2	1GKR
Guanine deaminase	GAH	3.5.4.3	1WKQ
D-Hydantoinase	HYD	3.5.2.2	1NFG
Isoaspartyl dipeptidase	IAD	3.4.19.5	1ONW
Imidazolonepropionase	IPase	3.5.2.7	2BB0
Phosphotriesterase	PTE	3.1.8.1	1HZY
Renal dipeptidase	RDP	3.4.13.11	1ITQ
γ-Resorcylate decarboxylase	RSD	4.1.1.44	2DVT
Urease	URE	3.5.1.5	2UBP
5-Methylthioadenosine/S-adenosylhomocysteine deaminase	MTAD	3.5.4.28	1J6P
4-Oxalomesaconate hydratase	OMAH	4.2.1.83	2GWG
2-Pyrone-4,6-dicarboxylic acid hydrolase	PHD	3.1.1.57	4D8L
N-Acyl-D-glutamate deacylase	AGD	3.5.1.82	3GIP
N-Isopropylammelide isopropyl amidohydrolase	IIA	3.5.99.4	2QT3
Enamidase	ENA	3.5.2.18	2UVN
Isoorotate decarboxylase	IDCase	4.1.1.66	4HK7
L-Histidinol phosphate phosphatase	HPP	3.1.3.15	4GC3
Atrazine chlorohydrolase	ATZA	3.8.1.8	3LS9

[a] Although DHPase and HYD have the same E.C. number, they are different enzymes and have different substrate specificities. DHPase enzymes (including L-hydantoinase) catalyze the reversible hydrolytic ring opening of six- or five-membered cyclic diamides such as dihydropyrimidines and 5′-monosubstituted hydantoins to the corresponding 3-ureido acids and carbamoyl amino acids, respectively. HYD is an industrial enzyme that is widely used in the production of D-amino acids which are precursors for synthesis of antibiotics, peptides and pesticides.

Biochemical studies on the amidohydrolase superfamily have revealed that the metal cofactor is catalytically essential for the majority members (19). The metal cofactor is located at the opening of the barrel near the C-terminal ends of several β strands, coordinated by several histidine, aspartic acid and in some occasions cysteine residues of the β strands and βα loops. The metal-containing site is referred to as the catalytic face of the barrel, in contrast to the stabilizing face of the opposite end of the barrel (20). To accommodate the metal center and substrates, the catalytic face often has a wider width than the stabilizing face, consequently making the other side of the barrel appear more compressed. Thus, the barrel sometimes looks like a conoid.

The most common metal in this enzyme superfamily is a zinc ion or a dinuclear Zn_2 pair. However, transition metals, such as divalent iron, nickel and manganese ions, are also observed in amidohydrolase enzymes. Cobalt ion can often be substituted for zinc *in vitro*, producing a comparable or even higher level of enzyme activity (14). Non *d*-block metal ions, such as Mg^{2+} and Ca^{2+}, are usually unable to perform the biochemical functions at the enzyme active site, suggesting that the role of the *d*-block metal cofactor is beyond simply providing a cationic charge for substrate binding.

The metal cofactor is tethered to the protein through a few protein residues. There are several types of active site structural architecture known in this enzyme superfamily, and the origin of the metal ligands is a valuable tool to divide the enzyme superfamily into seven subgroups (19). A mononuclear cofactor requires four or five ligands, while a dinuclear cofactor demands five or six protein ligands. One notable structural feature is that at least one water molecule or water-derived hydroxide is coordinated to the metal ion(s) in the enzyme structures. The solvent-derived ligand generally remains bound to the metal in the enzyme–substrate complex, but it is

replaced by substrate on some occasions such as 2,6-dihydroxybenzoate (γ resorcylate) decarboxylase (RSD) (21).

For a mononuclear metal cofactor of this enzyme superfamily, the fifth β strand invariably provides a histidine ligand for the metal coordination. The sixth β strand contributes a histidine residue either as a metal ligand or a nonligand but an important catalytic component at the enzyme active site. These two histidine residues are the signature amino acids conserved across the mononuclear and dinuclear metal cofactor containing amidohydrolase superfamily members. It is common for the first and/or eighth β strands to add more protein ligands for metal coordination. An aspartic acid residue from the eighth β strand in many amidohydrolase enzymes plays a dual role, i.e. acid/base catalyst and metal ligand.

The two divalent metals are separated by ~3.6 Å in the dinuclear metal cofactor. The more buried metal cation is coordinated to two histidine residues from the end of β strand 1 and an aspartate from β strand 8. The more solvent-exposed metal ion is ligated to protein through two imidazole side chains of histidine from β strands 5 and 6 (19). The two divalent metal ions are bridged by a hydroxide in addition to a carbamate functional group originating from the posttranslational modification of a lysine residue from β strand 4 in the presence of bicarbonate. In some cases, an unmodified glutamic acid from β strand 4 or a cysteine from β strand 2 is employed as a bridge instead of the modified lysine. Until now, the characterized dinuclear amidohydrolase enzymes have a homogeneous metal preference. However, some mixed metal cofactors may exist. Enamidase, for instance, is a bifunctional enzyme belonging to the amidohydrolase family that mediates hydrolysis of 1,4,5,6-tetrahydro-6-oxonicotinate to ammonia and (S)-2-formylglutarate; it reportedly contains a Fe-Zn cofactor (22).

A subfamily under the amidohydrolase superfamily, polymerase and histidinol phosphatase, catalyzes the reaction of phosphoester hydrolysis. Members in this subfamily have a trinuclear metal center buried in a distorted $(\beta/\alpha)_7$ barrel. The β-strand 3 is much longer than other strands composed the barrel and overlays with the β-strand 3 and 4 of the $(\beta/\alpha)_8$ structures. All the metal ligands for the mono- and di- nuclear metal center are conserved and coordinate the α and β metals. In addition, the third metal is coordinated by an aspartate or histidine from the end of β-strand 1 and two histidine residues from the end of β-strand 2 and 8.

The identification of several metal independent members further expanded the structural diversity of the amidohydroloase superfamily. The overall structure of metal independent members still adopts a $(\beta/\alpha)_8$ TIM-barrel fold and the four conserved residues, three histidine (two from the end of β-strand 1 and one from β-strand 6) and one aspartate (from the end of β-strand 8) ligand residues in the mononuclear metal center, are still present in the active site. These conserved residues are now proposed to conduct new catalytic roles since they no longer serve as metal ligand. Other than the histidine residues from the sixth β strands, which are conserved throughout the entire family, the remaining metal ligands exhibit some variations. For example, an HxH metal-binding motif is commonly seen in strand 1 of the β-barrel, but HxD and ExH have been observed from the ACMSD protein subfamily of Amidohydrolase_2 (18). The histidine from the fifth β-strand is absent in the cofactor free members. Moreover, the metal-binding motif from strand 1 does not always serve as metal ligands. For instance, in the structures of AGD, the HxH motif is present in the active site but it is not ligated directly to the divalent metal.

The overall sequence conservation among the amidohydrolase superfamily is rather low, indicating that the TIM-like barrel fold is not dictated by details of sequences but rather by

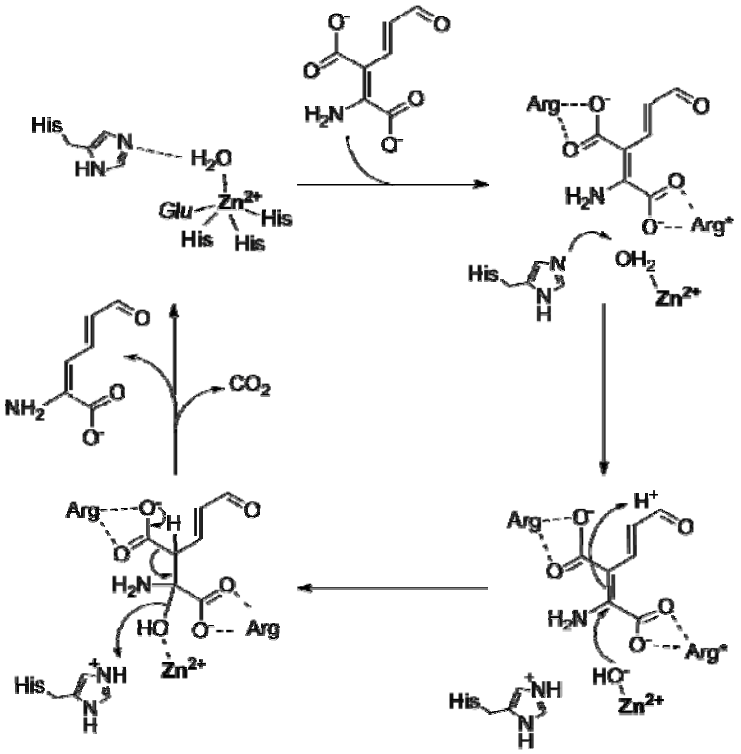
overall distribution of polar or charged and nonpolar or noncharged residues. Also, it is not unusual that enzymes in this superfamily contain noncatalytic domains in addition to the catalytic domain. Thus, conventional sequence alignments occasionally fail to reveal the general characteristics of a potential new member of the amidohydrolase superfamily. When either the metal-binding motif or the TIM-like barrel is not obvious, an advanced sequence alignment may be required. If this is the case, a secondary structure pattern can be calculated by the PSIPRED protein structure prediction server at http://bioinf.cs.ucl.ac.uk/psipred (23). Then a 'secondary structure-based' sequence alignment with those members with known three-dimensional structures may reveal clusters of similar residues at topologically equivalent positions. An example using this strategy was described in a recent study of ACMSD, which successfully predicted that this is a new member of the amidohydrolase superfamily. A site-directed mutagenesis analysis was followed and the results revealed important information concerning the enzyme such as metal ligand identities and active site residues as well as their possible roles in the catalytic process (15).

However, one should be cautious in drawing conclusions solely based upon a sequence study. Sequence, and even structural, similarities do not always translate into functional similarities. One extreme example in the functionally diverse amidohydrolase superfamily is the difference between melamine deaminase and atrazine chlorohydrolase. These two enzymes share 98% sequence identity, but catalyze completely different reactions (24).

1.3.3 Catalytic mechanism of members in the amidohydrolase superfamily

The best-characterized members of the amidohydrolase superfamily share a common catalytic mechanism. Scheme 1.2 illustrates the proposed hydrolytic mechanism for the mononuclear Zn-dependent enzyme ACMSD. The common feature of the mechanism is that a

metal-bound water becomes a hydroxide with the assistance of an active site catalyst. It then attacks the substrate-bearing amide or other functional groups at carbon and phosphorus centers and forms a tetrahedral carbon intermediate on the substrate. Subsequent collapse of the substrate-based intermediate leads to the hydrolytic products (25,26).



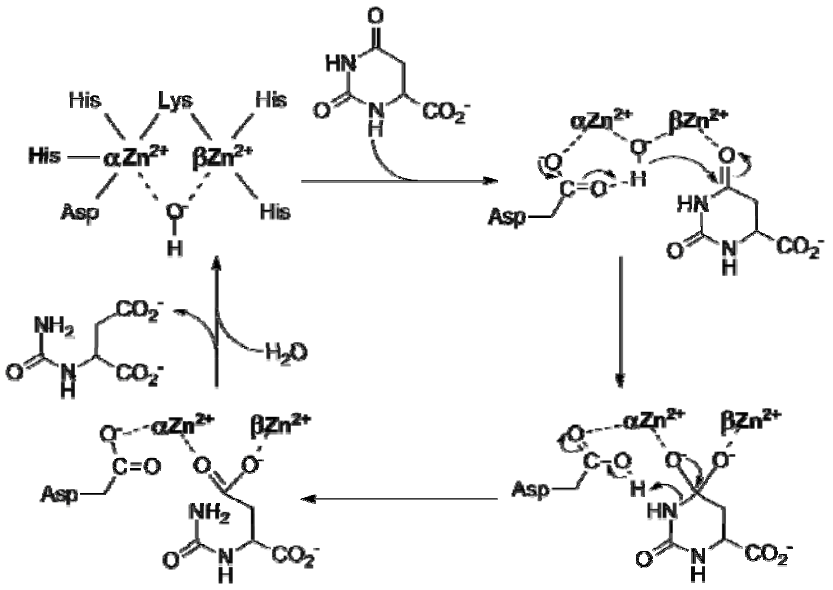
Scheme 1.2. A working model of ACMSD catalytic cycle for a nonhydrolytic C–C bond cleavage. Arg* represents a conserved arginine residue intruded to catalytic center from a neighbouring subunit of the homodimeric protein.

Mechanistic understanding of the nonhydrolytic C–C bond and C=C bond cleavage reactions is still in its infancy. However, recent biochemical and structural studies have yielded evidence supporting the idea that these reactions have mechanisms similar to those proposed for the hydrolytic reactions. The core elements for the required chemistry, a proper Lewis acid metal center, a water ligand and a nearby histidine (the conserved histidine from the end of β-strand 6), are all in place at the active site of the characterized members of the Amidohydrolase_2 branch, suggesting a common mechanistic paradigm for substrate activation at the early steps of the

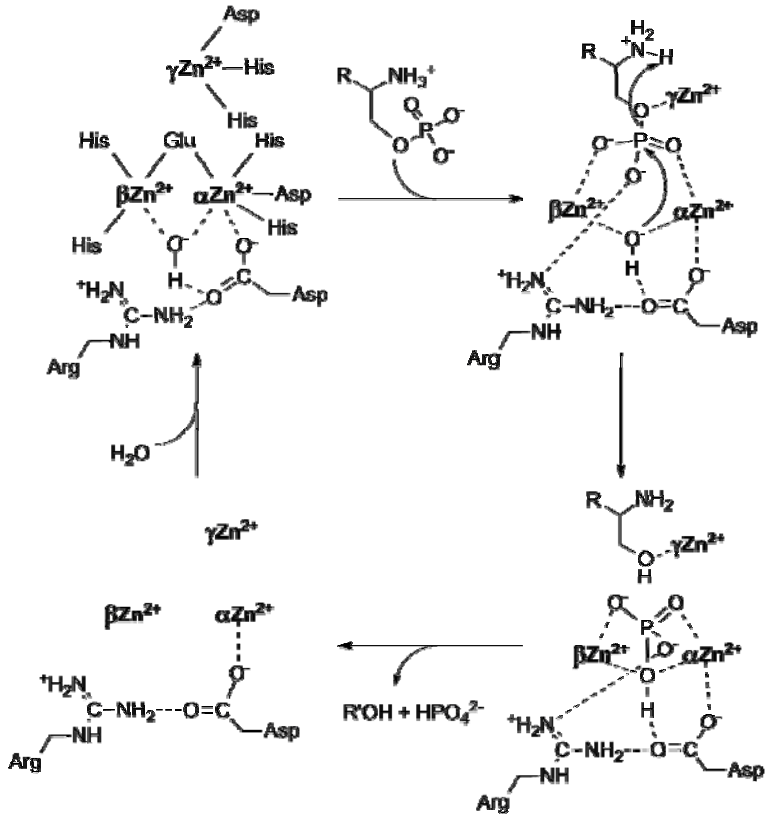
reaction catalyzed by the enzyme of this superfamily. From a chemical perspective, the enzyme OMAH of the ACMSD protein family catalyzes a mechanistically insightful reaction; it adds a water molecule to the substrate, as if it were the half-reaction in a decarboxylation mechanism proposed for ACMSD (Scheme 1.2). Although there is no characterized member of the Amidohydrolase_3 subfamily at the present stage, it is almost certain that some structural variations and new functionalities will soon be described for members of this branch.

An arginine residue, Arg239 in ACMSD, is conserved across the characterized members catalyze the decarboxylation reaction and proved to be catalytically essential. Arg239 has been shown to play important role in substrate binding in solution for ACMSD. Crystal structure of IDCase in complex with a substrate analogue further demonstrated that this arginine specifically binds to the leaving carboxyl group. Unlike other active site residues, Arg239 is intruded from a neighbouring subunit, indicating this subfamily of enzymes needs to function as dimers. This is also verified by the gel filtration studies done for OMAH and RDC.

The dinuclear enzymes proceed similarly, with a second divalent metal ion providing additional activation power towards the substrate. DHO is one of the best characterized dinuclear amidohydrolase enzymes. It is believed that the original keto oxygen of the substrate interacts with the more solvent exposure metal ion in DHO, while the more buried metal ion activates the nucleophilic attack during catalysis. A proposed mechanism of DHO catalyzed reaction is shown in Scheme 1.3.



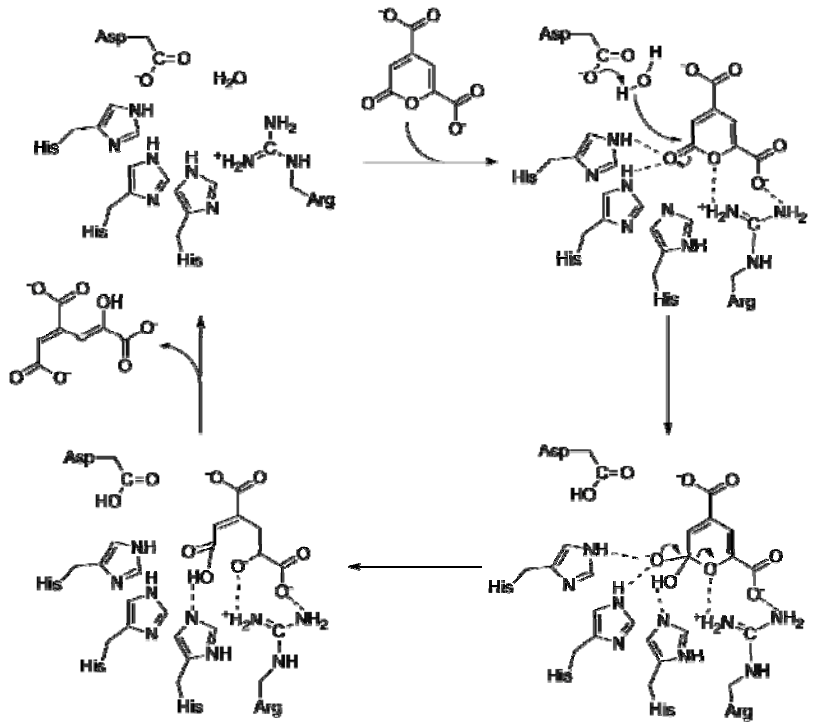
Scheme 1.3. The proposed catalytic mechanism of dihydroorotase (DHO). DHO contains a dinuclear metal center and catalyzes a ring opening reaction using dihydroorotate as substrate.



Scheme 1.4. The mechanistic model for a trinucler zinc center-mediated L-histidinol phosphate phosphatase (HPP) reaction. L-Histidinol and inorganic phosphate are the two reaction products.

The subgroup of polymerase and histidinol phosphatases that contain a trinuclear metal center is relatively new. Tyrosine phosphatase and L-histidinol phosphate phosphatase (HPP) are two representative members (27-29). For HPP catalyzed reaction, three metal ions and one arginine residue are involved in binding with the phosphate oxygens of the substrate. The γ-metal binds to the leaving oxygen and is proposed to serve as a Lewis acid to facilitate the O-P bond cleavage. A hydroxide group, which initially serves as a bridge between the α-metal and β-metal, nucleophilic attacks the phosphorus of the substrate. The amino group of the substrate may function as a general acid to protonate the leaving oxygen. A proposed mechanism is shown in Scheme 1.4.

LigI, an enzyme catalyzes the reversible hydrolysis of 2-pyrone-4,6-dicarboxylate from Sphingomonas paucimobilis is the first member of amidohydrolase superfamily that found does



Scheme 1.5. The proposed catalytic cycle of LigI. LigI catalyzes the reversible hydrolysis of 2-pyrone-4,6-dicarboxylate to 4-oxalomesaconate and 4-carboxy-2-hydroxymuconate in the degradation of lignin. The enzymatic reaction is independent to the presence of a metal ion.

not require a divalent metal (30). Later on, BmulJ_04915 and GLI are found as the second and third examples (31). Although this group of enzymes does not contain a metal cofactor, the original metal binding residues, including three histidine and one aspartate, are still conserved in the active site with new catalytic functions. For LigI and BmuIJ_04915, the three histidine residues are proposed to bind and polarize the substrate. The Aspartate residue is proposed to activate a water molecule for nucleophilic attack (Scheme 1.5).

The amidohydrolase superfamily has attracted substantial attention as a large structure-based cluster of enzymes with thousands of members and divergent catalytic functions. In the past few years there have been huge advances in the studies of the amidohydrolase superfamily. Large volumes of sequence are now becoming functionally annotated, and many members that exhibit distinct biological significances are now structurally defined. It is becoming clear that the metal-cofactor elaborated by the sturdy and versatile TIM-like β barrel is enormously powerful. The substrate specificity appears to be mostly dictated by the loops, insertions and conformational restrictions of the catalytic face of the TIM-like β barrel. This insight has come from site-directed mutagenesis, kinetics, structural and spectroscopic studies of many members of the superfamily.

Major advances have been made, but a structure-based functional annotation is still challenging. Many principals and strategies related to the design of loops, insertions, subunit interactions and the catalytic impact of protein dynamics remain to be elucidated. Nonetheless, the rapid expansion of knowledge concerning the metal-dependent TIM-barrel enzymes is making the amidohydrolase superfamily a well characterized large enzyme group that may enable a reliable evolutionary analysis for the origin of the divergent members. The amidohydrolase superfamily as a whole is emerging as a popular and valuable asset in the enzymology field for studying structure–function relationships and evolution.

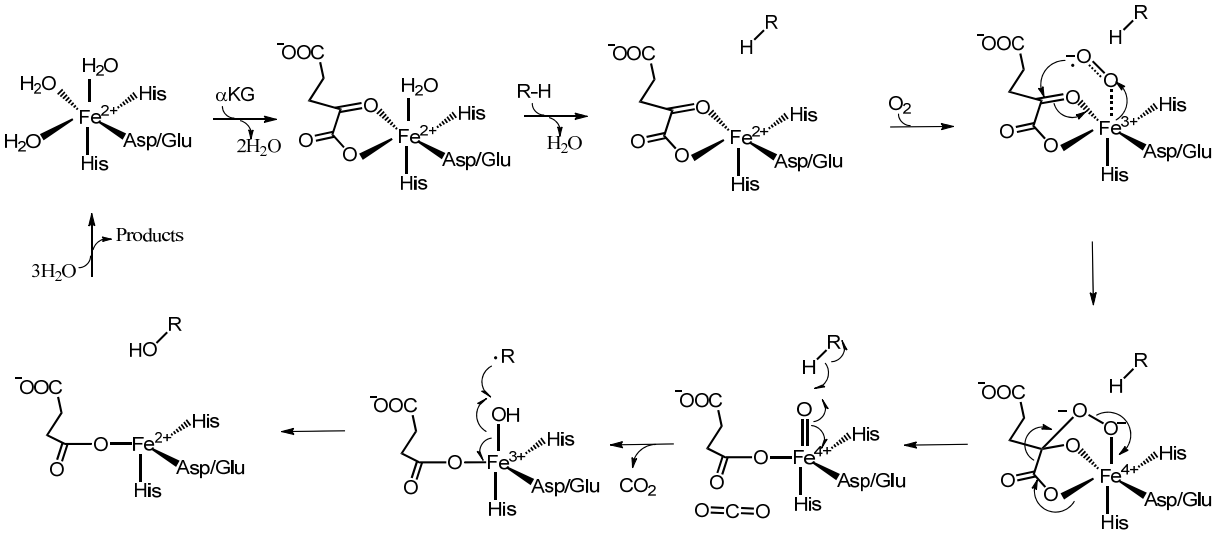
1.4 ACMSD is a Transition Metal Dependent Decarboxylase

Decarboxylation is one of the most common processes in nature and one of the most fundamentally important reactions in biological systems. Essentially all of the carbon dioxide evolved in fermentation and respiration is generated by the decarboxylation of organic acids (32). Decarboxylases are known for their roles in a wide variety of catabolic and anabolic pathways including decarboxylation of α- and β-keto acids, amino acid conversions, and carbohydrate synthesis (33). Within the IUPAC classification scheme, these are currently divided into at least 90 different classes. Substantial effort has been applied to the study of the 60 origin and the mechanisms of production of metabolic carbon dioxide, and considerable knowledge has been accumulated regarding the decarboxylation mechanisms in biological systems. Enzymatic decarboxylation usually utilizes either an organic cofactor such as pyridoxal 50-phosphate and biotin, or an inorganic cofactor, such as an iron or zinc complex, in the catalytic reaction. ACMSD is one of the decarboxylases whose activity is dependent of inorganic metals. Metal ions are employed as cofactors in many decarboxylases, some of which do not function as catalytic centers. For instance, the Zn^{2+} ion in the biotin-dependent decarboxylases helps to properly orient the substrate while the Mg^{2+} ion in the ThDP-dependent decarboxylases assists in binding the thiamine cofactor. In $NAD(P)^{+}$-dependent decarboxylases, such as malic enzymes, a divalent metal ion (Mn^{2+} or Mg^{2+}) plays an important role in both catalysis and structure stability. In some cases, metal cofactors play critical catalytic roles.

1.4.1 Mechanism of ferrous-dependent oxidative decarboxylases

α-Ketoglutarate (αKG)- and nonheme ferrous-dependent dioxygenases are classic ferrous-dependent decarboxylases (34,35). The αKG-dependent dioxygenase reaction proceeds in two half reactions. The first half of the reaction is a ferrous and O_2-dependent oxidative

decarboxylation reaction of αKG. The cleavage of the O-O bond results in one oxygen atom inserting into αKG while the other oxygen remaining at the Fe center (Scheme 1.6). The αKG- and Fe-dependent enzymes are widespread and versatile, catalyzing an exceptionally wide range of energetically demanding biosynthetic and degradative reactions such as the stereo-selective desaturation of inactivated C-C single bonds, ring expansion, oxidative ring closure, and hydroxylation processes (35), which are important in the biosynthesis of collagen (36), antibiotics biosynthesis (37,38), cellular response of hypoxia (39), repair of alkylatedDNA(40,41), regulation of gene expression by demethylation of histones (42), and many other biological processes.



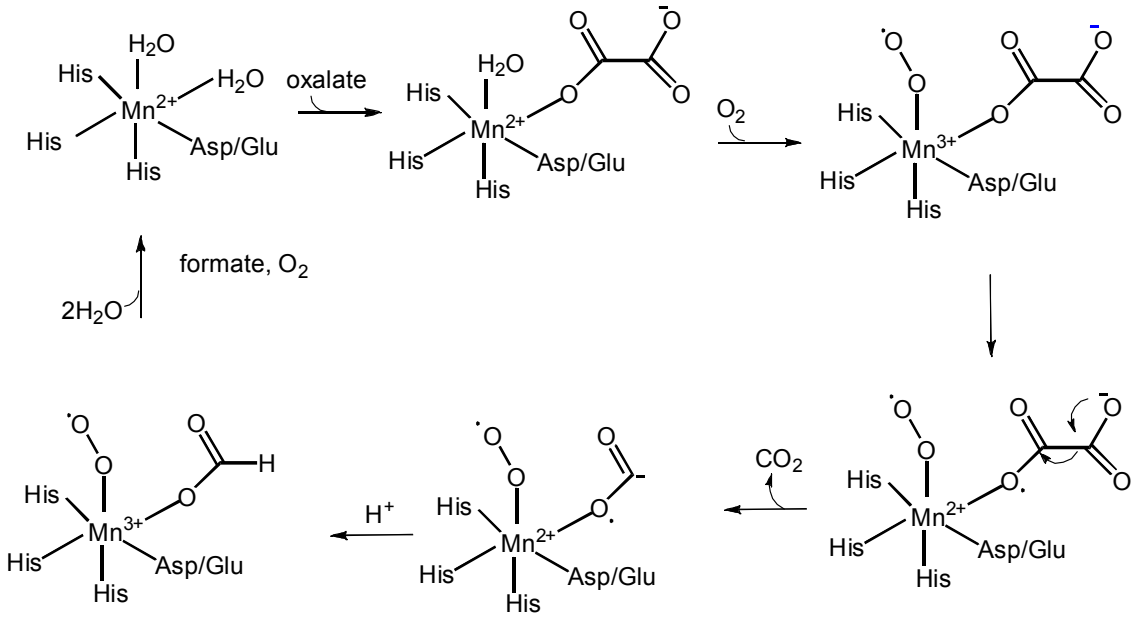
Scheme 1.6. A proposed mechanism for ferrous-dependent oxidative decarboxylation of αKG.

The catalytic mechanism of αKG-dependent dioxygenases has been extensively studied. This family of enzymes all require an iron(II) center, which is relatively weakly coordinated by three amino acid side chains, typically in a 2-his-1-carboxylate facial triad (43). The mechanism for this group of decarboxylases proceeds via a Fe(IV)=O intermediate (Scheme 1.6), which was first proposed three decades ago and recently confirmed by Bollinger and Krebs using EPR and

Mössbauer spectroscopy (44-49) and by Hausinger using continuous-flow resonance Raman spectroscopy (50,51). In this model, the binding of the enzyme to αKG, theprimary substrate, and the binding of molecular oxygen follows an ordered binding mechanism (49,52). The iron center of the resting enzyme is coordinated with two histidines and one carboxylate (Asp or Glu) residues together with three water molecules. The addition of the co-substrate, αKG, will replace the two water ligands with its 1-carboxyl and 2-keto groups. The primary substrate will then enter and expel the third water molecule, which leads to an empty coordination site ready for a dioxygen molecule to come in and bind to the iron center (49,53). Addition of dioxygen to the substrate/co-substrate/iron-bound enzyme complex leads to the oxidation and decarboxylation of αKG to produce succinate and form a substrate-hydroxylating Fe(IV)=O species through two non-accumulating complex species in which the Fe(IV)=O intermediate has been trapped and characterized (44,48,50,54-56). Although the details of the decarboxylating mechanism remain elusive, it has been demonstrated that decarboxylation of αKG occurs prior to or concomitantly with the formation of this substrate-hydroxylating Fe(IV)=O intermediate (55). Carbon dioxide may remain bound to the metal center in its hydrated or nonhydrated form as observed in the self-hydroxylation reaction study in taurine/aKG dioxygenase (57) and the crystallographic study in deacetoxycephalosporin C synthase (58). In the next step, the highly reactive Fe(IV)=O intermediate abstracts a hydrogen atom from the primary substrate and generates a radical species on the substrate. It also causes a reduction of the metal center. The radical abstracts the hydroxyl group from the iron center, resulting in the hydroxylation of the substrate (47,48). The rate-determining step in the catalytic cycle is the release of products (55). For some enzymes in this family, the Fe(IV)=O intermediate is believed to perform an electrophilic attack on the aromatic ring of the primary substrate rather than hydrogen abstraction (47,59,60).

1.4.2 Mechanism of manganese-dependent decarboxylases

Enzymes that require Mn^{2+} and O_2, but not any organic cofactors, to decarboxylate substrates include oxalate decarboxylase (EC 4.1.1.2) and the closely related oxalate oxidase (EC 1.2.3.4). Moreover, 4-oxalocrotonate decarboxylase (EC 4.1.1.77) forms a complex with vinylpyruvate hydratase and utilizes either Mn^{2+} or Mg^{2+} as a cofactor to convert 2-oxo-3-hexenedioate to 2-oxo-4-hydroxypentanoate (61,62). Oxalate decarboxylase catalyzes the conversion of oxalate to formate and carbon dioxide. Oxalate decarboxylase utilizes both Mn^{2+} and dioxygen as its cofactors, but the reaction does not involve a net redox change (63). Oxalate oxidase also requires Mn^{2+} and dioxygen for its catalytic activity. Furthermore, it degrades oxalate to produce carbon dioxide. However, the reaction uses dioxygen as a substrate involves a net redox reaction (64,65). Both oxalate decarboxylase and oxalate oxidase catalyzing reactions utilize free radicals to cleave a relatively stable carbon–carbon bond. Current proposed mechanisms for oxalate degradation by oxalate decarboxylase and oxalate oxidase suggest several essentially identical steps up to formation of the Mn-bound formyl radical (64,66). In fact, it has been demonstrated that oxalate decarboxylase can be converted to an efficient oxalate oxidase by mutating amino acids on a flexible active site lid (67). Scheme 1.7 illustrates a proposed mechanism for oxalate decar boxylase based on a number of hypotheses (64-69).



Scheme 1.7. Proposed mechanism for oxalate decarboxylase.

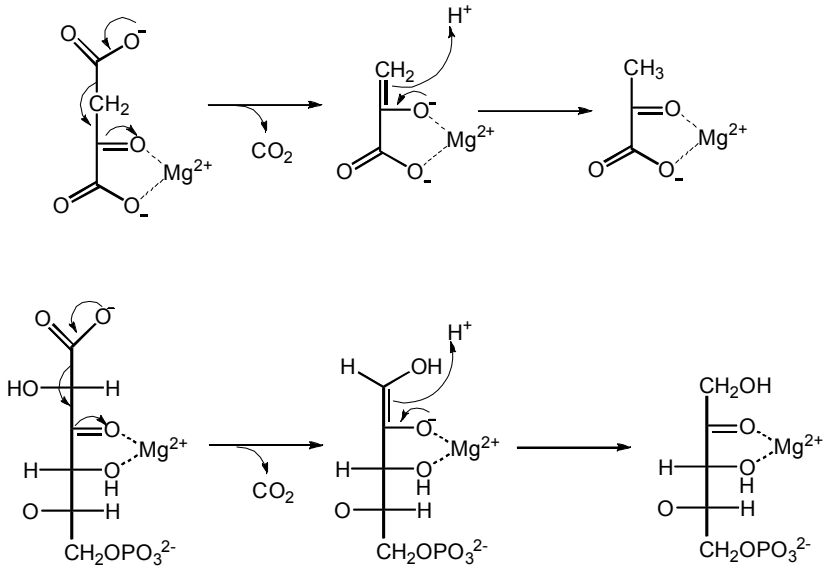
In 2007, the two metal centers in oxalate decarboxylase were individually characterized using high field, multifrequency EPR spectroscopy (70). The carbon dioxide anion radical shown in step five in Scheme 1.7 was trapped in oxalate decarboxylase by Angerhofer and co-workers in 2011 using an EPR spin trapping technique. Through mutagenesis, the authors were able to disrupt a hydrogen bond responsible for closing the flexible loop across the substrate binding channel. This led to a buildup of a carbon dioxide anion radical. A closed conformation is required for decarboxylation, which ensures that the harmful anion radical intermediate remains bound at the active site during catalysis (71). At the resting state of the enzyme, Mn^{2+} coordinates to three conserved histidine residues, a conserved glutamate or aspartate residue, and two water molecules (65,69). When malate comes into the active site, it binds to the metal ion monodentately and replaces one of the water molecule ligands. Then dioxygen binds to Mn^{2+} and replaces the other water molecule. One electron transferred from the metal to the binding dioxygen results in a Mn^{3+}-superoxo species, which functions as an electron sink to facilitate the

decarboxylation reaction. An electron is drawn from the bound substrate to the metal to generate a radical on the bound oxalate. Cleavage of the C-C bond of this radical intermediate releases CO_2 and gives a manganese ion-bound formyl radical anion intermediate. At this point the mechanisms of oxalate oxidase and oxalate decarboxylase purportedly divert. In the hypothetical mechanism of oxalate decarboxylase, protonation occurs at the carbanionic carbon of the formyl radical intermediate, which is followed by an electron transfer from the manganese ion to the formyl radical forming a product bound Mn^{3+}-superoxo species. Product dissociation from the active site completes the reaction cycle and regenerates the enzyme.

1.4.3 Mechanism of magnesium-dependent decarboxylases

Divalent metal ions can catalyze a decarboxylation reaction at the β-carbon of α-keto carboxylic acid, while decarboxylases can increase the rate of the reaction dramatically. One of the most extensively studied cases is the decarboxylation of oxaloacetate. The reaction catalyzed by oxaloacetate decarboxylase is dependent on divalent metal ions. This enzyme is active with either Mn^{2+} or $Mg^{2+.}$ Besides the divalent ion, the decarboxylation catalyzed by this enzyme from some organisms also requires biotin, while $NAD(P)^+$ is essential for the enzyme from other species. A proposed mechanism for oxaloacetate decarboxylase from *Pseudomonas putida*, which requires only divalent metal cations, is shown in Scheme 1.8 upper (72). The decarboxylation catalyzed by 4-oxalocrotonate decarboxylase is believed to proceed in a mechanism similar to that of oxaloacetate decarboxylase (61). The metal divalent cation is thought to form a complex with the α-keto acid moiety and one of the oxygen atoms from thecarboxyl group of the substrate, functioning as an electron sink to stabilize the enolate intermediate (73,74). The enzyme serves to provide substrate specificity and to lower the energy of the ground state to a much more favorable value. Protonation at the Cb position yields the

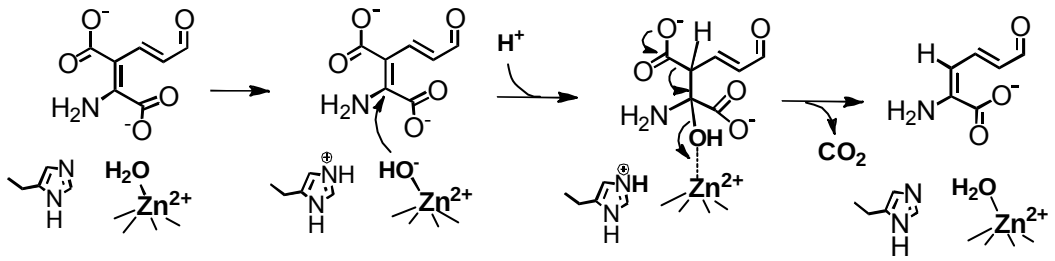
product pyruvate. The proton incorporated in the product can be in a retained or inverted stereochemical configuration compared to the leaving carboxyl group of the substrate. A recent study reported that the decarboxylation catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase is also Mg^{2+} ion-dependent. The reaction mechanism of this enzyme has been extensively studied (75-79). These studies have provided a hypothetical mechanism shown in Scheme 1.8 lower). In the proposed mechanism, the departure of CO_2 is facilitated by forming an enediolate intermediate stabilized by the Mg^{2+} ion. However, the required Mg^{2+} ion does not ligate to the oxygen of the carboxyl group. Instead it coordinates to the oxygen of the keto-group at the C3 position and the oxygen of the hydroxyl-group at the C4 position of the substrate, forming a cis-1,2-enediolate intermediate. Protonation of C1 of the intermediate is completed through a proton relay system involving two alternative water molecules in the active site and the side chains of a conserved histidine or a conserved arginine as the ultimate proton sources. Protonation can lead to either R or S configurations depending on which of the two conserved residues is used as the ultimate source of the proton.



Scheme 1.8. Hypothetical mechanisms for magnesium-dependent decarboxylases.

1.4.4 Transient metal-dependent, O_2-independent decarboxylases

ACMSD was initially thought to be a cofactor-free enzyme (12,13). It has recently been found to be a metalloenzyme belonging to the amidohydralase superfamily (14,15). This enzyme is an illustrative example of the ACMSD protein family (also known as the Amidohydrolase_2 subfamily) (18,80), whose members are decarboxylases or hydrotases rather than hydrolases or deaminases. Members of the ACMSD protein family catalyze a CAC bond cleavage reaction instead of hydrolytic reaction (80,81). Understanding the mechanism of ACMSD will shed light on the mechanisms of the continually emerging members of the ACMSD protein family. However, it is unknown whether ACMSD adopts the mechanism from the amidohydrolase_1 subfamily or follows a model evolved by other decarboxylases (82). Scheme 1.9 is a mechanistic model adopted from the well-characterized amidohydrolase_1 subfamily. In the first step, the substrate binds to the enzyme with its C2 carboxylate close but not to themetal center at a rate of 2.4 * 106 $s^{-1}M^{-1}$ (10). An active site histidine, His228 as numbered in the *Pseudomonas fluorescens* enzyme (80), assists in the deprotonation of a water ligand to form hydroxide, which then nucleophilically attacks the double bond between C2 and C3 of the substrate, resulting in a hydrated tetrahedral intermediate. The subsequent collapse of the tetrahedral intermediate is assisted by the concerted deprotonation of His228 and the rebinding of the hydroxyl group to themetal ion. The carboxylate attached to the C3 position becomes the leaving group. The decarboxylation step is a rapid reaction (>500 s^{-1}) but the release of the product aminomuconate-e-semialdehyde is slow (9 s^{-1}) (10).



Scheme 1.9. A working hypothesis for mechanism of the ACMSD reaction.

The amidohydrolase-derived mechanism shown in Scheme 1.9 predicts that the metal-bound hydroxide is the catalytic driving force of the decarboxylation reaction. The pK_a value of the water ligand is lowered from 15.7 of the unbound state to about 9.6 because of its coordination to zinc (83). This value can be further reduced to nearly the physiological pH with the assistance of an active site base catalyst, presumably His228 in *Pf*ACMSD, and therefore generates a hydroxide anion, which performs a nucleophilic attack on the C2 and C3 double bond of the substrate. As expected, mutation at His228 significantly reduces the enzyme activity and removal of the metal bound water molecule completely deprives the enzyme activity. This hydroxide attack mechanism is believed to be present in the enzyme 4-oxalomesaconate hydratase, which was identified as a part of this family in a recent bioinformatics study (80). This enzyme catalyzes the addition of a hydroxide ion and a proton across a C=C bond of its substrate. The added hydroxide is preserved in the product so that the proposed hydroxide attack for the ACMSD enzyme family is visualized in this specific member, as if it were an interruption of an ACMS decarboxylation reaction. The metal-bound hydroxide attack of the substrate is also proposed to occur in another zinc-dependent structurally characterized decarboxylase member of the ACMSD family, c-resorcylate decarboxylase (21). Therefore, the metal-bound hydroxide assisted C-C/C=C cleavage is likely to be a common mechanism for the ACMSD subfamily of the amidohydrolase superfamily.

1.5 AMSDH is a Member of the Aldehyde Dehydrogenase Superfamily

AMSDH is an NAD^+ dependent aldehyde dehydrogenase that oxidizes the aldehyde group of 2-AMS to its corresponding acid, competing with the spontaneous decay of 2-AMS to picolinic acid (Scheme 1.1). Sequence alignment indicates that AMSDH is a member of the hydroxymuconate semialdehyde dehydrogenase subfamily under the aldehyde dehydrogenase (ALDH) superfamily. AMSDH has not been studied at the molecular level, and no structure was available before our work.

1.5.1 Aldehyde dehydrogenase superfamily

Due to the innate chemical reactivity of aldehyde molecules, many of them are toxic even at low concentrations. Aldehydes are generated from both endogenous (formed during metabolism of alcohols, amino acids, biogenic amines, vitamins, steroids, and lipids) and exogenous (formed from the metabolism of a number of drugs and environmental agents) precursors. Members of the ALDH superfamily catalyze the detoxification of endogenous and exogenous aldehydes and are widely spread in both prokaryotic and eukaryotic organisms.

The first aldehyde dehydrogenase was identified and purified from beef liver in 1949 (84). The first primary structure of an ALDH was obtained from human liver in 1984 (85). Since then, 43,773 ALDH sequences (listed as aldedh in the Pfam database) are identified and 607 structures have been solved. Although ALDH are found in 5325 species with very diverse substrate preferences, they all utilize $NAD(P)^+$ as a cofactor to provide oxidation power during catalysis and their overall tertiary structure is highly conserved. The primary sequence, however, has relatively low conservation. Based on a primary sequence alignment of 145 ALDH proteins, only four residues are strictly conserved across the superfamily, including two glycines, one phenylalanine, and one glutamate (86). The catalytic thiol, a cysteine residue, is also conserved

in all sequences except one. In cephalopod omega-crystallins, the otherwise conserved cysteine is replaced by an arginine, and it has been reported to lack ALDH activity (87).

To date, nineteen human ALDH genes have been identified according to their divergent evolution based on the ALDH nomenclature system established in 1998. Human ALDH genes are closely related to human health and many diseases caused by defects in aldehyde metabolism are caused by mutations in the genes which encode the respective enzymes. For example, ALDH1a1 helps maintain the corneal transparency (88). A single nucleotide polymorphism in encoded ALDH1b1 was found to be related with alcohol induced hypersensitivity symptoms including asthma and flushing (89). Inactivation of ALDH2 will cause both heart disease (90) and alcohol sensitivity problems (91). It has been shown that a much higher percentage of Asian exhibit acute alcohol intoxication compared to Europeans. The underlying reason is a mutation of a glutamate residue to lysine in the ALDH2 gene which leads to inactivation of the expressed enzyme (92). Mutations in the ALDH7a1 gene were found to be related to pyridoxine-dependent seizures in children. This is because the mutation abolishes the activity of antiquitin and leads to the accumulation of δ 1-piperideine-6-carboxylate-α-aminoadipic semialdehyde (93). Recent studies also indicate that increased activity of ALDH1a1 could serve as a hallmark of cancer stem cells (94).

1.5.2 Structural features of members in the ALDH superfamily

The first X-ray crystal structure of an ALDH was solved in 1997 as a homo-dimer (95). Quaternary structures of most ALDH solved later are either homodimers or homotetramers, which consists of a dimer of dimers. All members of the ALDH superfamily have very similar three-dimensional structures, even those with low sequence identities. These similarities arise from several conserved glycine and proline residues which are responsible for constructing the

critical turns and loops, therefore maintaining the conserved tertiary structure (96). Similar to alcohol dehydrogenases, each single chain of an ALDH is composed of three domains: a subunit interaction domain, an $NAD(P)^+$- binding domain, and a catalytic domain (Figure 1.3).

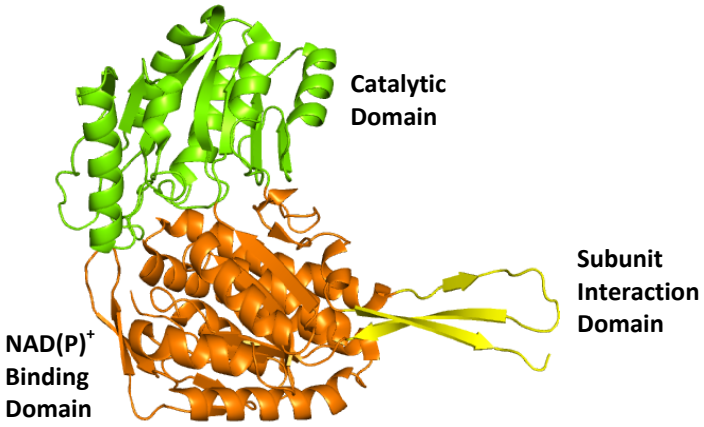


Figure 1.3. A single ALDH chain is composed of three domains. The subunit interaction domain is shown as yellow, co-substrate binding domain is shown in orange, and the catalytic domain is shown in green.

The subunit interaction domain is conserved throughout the ALDH superfamily, indicating that oligomerization is important for the function and stability of all ALDH enzymes. This domain consists of three β-strands, a β-hairpin motif and the final β-strand of the polypeptide chain, which extend to the catalytic domain of another protomer of ALDH in both dimer and tetramer quaternary structures. Studies have shown that ALDH with a dimeric quaternary structure has an extra C-terminal tail and misses some N-terminal residues compared to those tetramers. It is possible that the presence of the C-terminal tail prevents tetramer formation because it is located in the dimer-dimer interface for ALDH with tetramer structures (97).

The co-substrate, NAD^+ or $NADP^+$, is embedded in a Rossmann fold composed of five β-strands and four α-helices (95). The Rossmann fold is a protein tertiary structure motif that binds nucleotides. Binding of dinucleotides like NAD^+ or $NADP^+$ normally requires a pair of

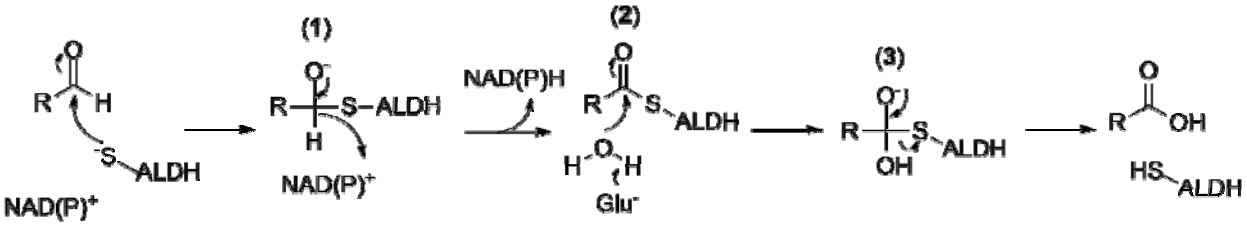
Rossmann folds with six β-strands and four α-helices. The majority of ALDH enzymes have a strong preference for either NAD^+ or $NADP^+$. Studies have shown that residues close to the adenosine ribose are responsible for co-substrate selectivity in $NAD(P)^+$ dependent dehydrogenases. Presence of an acidic residue near the adenosine 2'-position is a major characteristic for NAD^+ binding proteins because the anion side chain can interact with both 2'- and 3'-hydroxyl group of the NAD^+ adenosine ribose. Also, the carboxyl group of the acidic residue could prevent $NADP^+$ from binding by repelling its 2'-phosphate (98-101). On the other hand, placing a positively charged lysine or arginine at the same position will tune the co-substrate selectivity towards $NADP^+$ (102-104).

Multiple hydrogen bonds and salt bridges are involved in stabilizing the co-substrate. However, the nicotinamide head of NAD(P) remains flexible and stays in different conformations dependent upon the oxidation state (105,106). In the oxidized state, $NAD(P)^+$ lies in the Rossmann fold in an extended conformation, allowing the hydride to transfer from the substrate to reduce the co-substrate during the first half of the reaction. Reduced NAD(P)H adopts a bent conformation, in which the nicotinamide head moves back towards protein surface and leaves space for the second half of the reaction, acyl-enzyme hydrolysis, to take place. The catalytic domain has a topologically related βαβ-type polypeptide fold. The catalytic driving force, a cysteine residue, sits in the active site in close proximity to the $NAD(P)^+$ nicotinamide group and is connected to the protein surface by a channel, presumed to be the substrate binding channel. As an active-site catalytic thiol, this cysteine is easier to be oxidized than regular cysteine residues. Hence, purified ALDH enzymes easily lose their catalytic activity if they are not protected by reducing agents. It was suggested that the cysteine residue can be protected by forming a covalent bond with the $NAD(P)^+$ co-substrate (107,108). In some ALDH enzymes, a

specific mechanism has been developed to protect the catalytic cysteine. An extra cysteine is kept on a nearby loop which will reversibly form a disulfide bond with the active-site cysteine upon redox change (109).

1.5.3 Catalytic mechanism of members in the ALDH superfamily

Members in the aldehyde dehydrogenase superfamily share a similar catalytic mechanism. The dehydrogenase reaction is initiated by a nucleophile attack on the aldehyde carbon to form a covalently bound thiohemiacetal tetrahedral intermediate (110) (intermediate **1** in Scheme 1.10). A strictly conserved cysteine residue was assigned as the nucleophilic thiol. This idea was later proved by a series of inactivating chemical modification and site specific mutagenesis experiments (111-115). Stronger evidence was obtained when this intermediate was trapped in crystal structure of L-aspartate-β-semialdehyde dehydrogenase in the absence of co-substrate (116).



Scheme 1.10. ALDH enzymes catalytic mechanism

The next important step is hydride transfer from the tetrahedral carbon to reduce $NAD(P)^+$, leading to the formation of a thioacyl intermediate as shown in Scheme 1.10, intermediate **2**. Once the co-substrate has been reduced, it will adopt a bent conformation and move away from the active site. At the same time, a glutamate residue will rotate into the active site and take the position once occupied by the co-substrate. This intermediate has a better stability and has been trapped *in crystallo* from several ALDH enzymes (108,117,118).

After the glutamate residue rotates into an active position, it will work as a general base to deprotonate an ordered water molecule. The activated water will then attack the thioester carbon (previously the aldehydic carbon), leading to the formation of the second tetrahedral intermediate (Scheme 1.10, intermediate **3**). This glutamate is another strictly conserved residue in the ALDH superfamily and its function as a general base has been well studied in the past decades (108,117,119,120). Finally, collapse of the second tetrahedral intermediate and scission of the C-S bond leads to product release.

CHAPTER 2 EVIDENCE FOR A DUAL ROLE OF AN ACTIVE SITE HISTIDINE IN α-AMINO-β-CARBOXYMUCONATE-ε-SEMIALDEHYDE DECARBOXYLASE

The section of chapter 2 is our published work on ACMSD: Evidence for a Dual Role of an Active Site Histidine in α-Amino-β-Carboxymuconate-ε-Semialdehyde Decarboxylase. Lu Huo, Andrew J. Fielding, Yan Chen, Tingfeng Li, Hiroaki Iwaki, Jonathan P. Hosler, Lirong Chen, Yoshie Hasegawa, Lawrence Que, Jr., and Aimin Liu. (2012) *Biochemistry*, 51(29), 5811-5821.

This work is supported in part by grants from NSF Grant MCB-0843537 (AL), NIH grants GM 038767 (LQ) and GM 056824 (JPH), the Japanese strategic projects to support the formation of research bases at private universities: matching fund subsidy from MEXT, 2008-2012 and the University of Minnesota Chemical Biology Initiative in conjunction with the NIH-supported Chemistry Biology Interface Training Grant GM 08700 (AJF).

2.1 Abstract

The previously reported crystal structures of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) show a five-coordinate Zn(II)$(His)_3(Asp)(OH_2)$ active site. The water ligand is H-bonded to a conserved His228 residue adjacent to the metal center in ACMSD from *Pseudomonas fluorescences* (*Pf*ACMSD). Site directed mutagenesis of His228 to tyrosine and glycine in the present study results in complete or significant loss of activity. Metal analysis shows that H228Y and H228G contain iron rather than zinc, indicating that this residue plays a role in metal selectivity of the protein. As-isolated H228Y displays a blue color, which is not seen in wild-type ACMSD. Quinone staining and resonance Raman analyses indicate that the blue color originates from Fe(III)-tyrosinate ligand-to-metal-charge-transfer (LMCT). Co(II)-substituted H228Y ACMSD is brown in color and exhibits an EPR spectrum showing a high-spin Co(II) center with a well-resolved ^{59}Co ($I = 7/2$) eight-line hyperfine splitting pattern. The

X-ray crystal structures of the as-isolated Fe-H228Y (2.8 Å), Co- (2.4 Å) and Zn-substituted H228Y (2.0 Å resolution) support the spectroscopic assignment of metal ligation of the Tyr228 residue. The crystal structure of Zn-H228G (2.6 Å) was also solved. These four structures show that the water ligand present in WT Zn-ACMSD is either missing (Fe-H228Y, Co-H228Y, and Zn-H228G) or disrupted (Zn-H228Y) in response to His228 mutation. Together, these results highlight the importance of His228 for *Pf*ACMSD's metal specificity as well as maintaining a water molecule as ligand of the metal center. His228 is thus proposed to play a role in activating the metal-bound water ligand for subsequent nucleophilic attack on the substrate.

2.2 Introduction

Histidine, an essential amino acid, is found at the active site of a myriad of enzymes. The imidazole side chain of histidine can serve as a coordinating ligand in metalloproteins and in many cases is a catalytically important component in the active sites of enzymes. The pK_a of the imidazolium ion (~ 7) permits significant concentrations of both acidic and basic forms near neutral pH, making it a commonly found general acid-base catalyst in enzymes (121). The participation of histidine in catalysis has also been illustrated in the well-studied mechanisms proposed for members of the amidohydrolase superfamily (19,81). The best-characterized members of the amidohydrolase superfamily share a common catalytic mechanism by which a metal-bound water is proposed to be activated by an active site base to yield a metal-hydroxo species. Subsequently, the hydroxide attacks the substrate bearing the amide or other functional groups at the carbon or phosphorus center to form a tetrahedral carbon or pentacoordinate phosphorus intermediate on the substrate (19,81). The collapse of the substrate-based intermediate leads to the hydrolytic products.

The enzyme α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) catalyzes the decarboxylation of its substrate ACMS to form 2-aminomuconate semialdehyde product (AMS). It is the first enzyme known to perform an O_2-independent nonoxidative decarboxylation with *d*-block divalent metal cofactors such as zinc, cobalt, or iron (14). ACMSD is proposed to be a prototypical member of a new protein subfamily in the amidohydrolase superfamily representing a novel nonhydrolytic C-C bond breaking activity (15,18). Scheme 2.1 outlines a working mechanism of ACMSD action.

ACMSD plays an important role in two distinct metabolic pathways, the mammalian kynurenine pathway for tryptophan degradation (122) and the bacterial 2-nitrobenzoic acid pathway for biodegradation of nitroaromatic compounds (12,123,124). In both pathways, ACMSD controls the final fate of the metabolites by competing with a slow spontaneous reaction that produces the excitotoxin quinolinic acid (QA) and directs the metabolic flux away from QA to energy production (125), QA is an endogenous selective agonist of *N*-methyl-D-aspartate receptors (2). It modulates neurotransmission and mediates immune tolerance. Reduced activity of ACMSD can lead to abnormal QA concentration in body fluids, which has been linked to numerous diseases, including diabetes (125), neuropsychiatric diseases (126), neurodegenerative disorders such as Alzheimer's and Huntington's diseases (127,128), stroke, and epilepsy (2,129). Therefore, the results of mechanistic study on ACMSD have significant medical implications.

Although ACMSD was discovered more than 55 years ago (130), little was known about the catalytic mechanism until our recent collaborative studies of *Pseudomonas fluorescences* ACMSD (*Pf*ACMSD) (14) unmasked its metal cofactor and determined its three dimensional structure (80). The as-isolated enzyme contains a zinc ion (80) and it works equally well when substituted with a cobalt(II) ion (14,15). The decarboxylation is rapid ($> 500\ s^{-1}$); the product

release from the active site is the rate-limiting step (10). Recently, the crystal structure of human ACMSD was reported with a bound inhibitor 1,3-dihydroxyacetonephosphate (131). The human ACMSD and *Pf*ACMSD present a nearly identical overall structure consisting of an $(\alpha/\beta)_8$-barrel core and a five-coordinate protein-bound zinc ion (80,131). The metal ion is bound to the enzyme by three histidines (His9, His11, and His177), one aspartic acid residue (Asp294), and one water ligand. A nearby histidine from β-strand 6, *i.e*. His228 in *Pf*ACMSD residue numbering, is located in the secondary sphere and is part of a hydrogen-bonded system that appears to facilitate deprotonation of the water ligated to the catalytic zinc (Figure 2.1). A comprehensive biochemical, spectroscopic and structural study presented in this work suggests that His228 is a major determinant of metal preference in *Pf*ACMSD and that it plays an important role in maintaining the hydrolytic water ligand necessary for the decarboxylase chemistry in this enzyme.

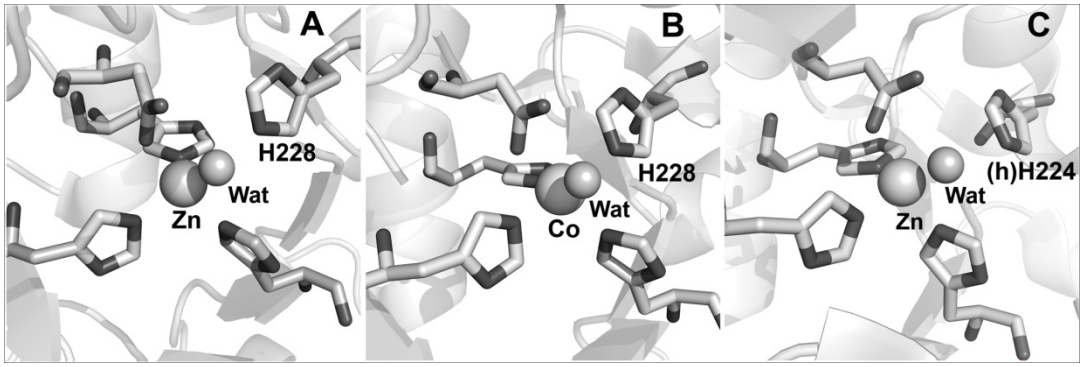


Figure 2.1. The active site histidine (His228 or *h*His224) is in H-bond distance with the metal-bound water ligand (Wat) in crystal structures of ACMSD. (A) as-isolated Zn-*Pf*ACMSD (PDB: 2HBV), (B) Co(II)-substituted *Pf*ACMSD (2HBX), and (C) an inhibitor-bound structure of *h*ACMSD (2WM1).

2.3 Experimental Procedures

Site-Directed Mutagenesis. The plasmid containing His-tagged ACMSD from *Pseudomonas fluorescens* was used as a template for construction of all of the mutants (15). The forward

primers used in the site-directed mutagenesis are 5'-CAAGATCTGTTTCGGGggTGGTGGGGGAAGTTTCG-3' for H228G and 5'-CAAGATCTGTTTCGGGtATGGTGGGGGAAGTTTCG-3' for H228Y. ACMSD mutants were constructed by the PCR overlap extension mutagenesis technique (132). The insertion of each construct was verified by DNA sequencing to ensure that base alterations were introduced correctly as expected and with no undesired changes. After sequencing, the positive clone was used for overexpression of these mutants in *E. coli* BL21(DE3).

Bacterial Growth and Protein Preparation. The expression and purification of WT *Pf*ACMSD and its mutants were the same as described previously (15). Apo H228G was prepared by EDTA treatment; Zn(II) and Co(II) containing H228G proteins were prepared by metal reconstitution reaction as previously described (14). Because iron cannot be removed from as-isolated H228Y by high concentration EDTA, Zn(II)- and Co(II)-H228Y proteins were obtained by substituting 0.1 mM $CoCl_2$ or $ZnCl_2$ to M9 medium prior to the addition of IPTG for induction during cell culture. The protein was purified in the presence of 0.1 mM $CoCl_2$ or $ZnCl_2$ during affinity chromatographic step by using Co(II)- and Zn(II)-charged IMAC columns, respectively.

Metal Analysis. The metal content of the as-isolated ACMSD from LB was determined by inductively coupled plasma optical emission spectroscopy (ICP-OES) using a Spectro Genesis spectrometer (Spectro Analytical Instruments). Assays were performed in triplicate for the mutants the same day as the assay samples of the wild-type (WT) enzyme. Standard curves were prepared using a transition element standard mixture (CCS-6) from Inorganic Ventures. Error (standard deviation) within three measurements of each sample was less than 1%, while error between different ACMSD preparations was within 4%. Metal content is reported per ACMSD

monomer. Protein concentration was measured using Coomassie Plus protein assay reagent (Pierce) according to the manufacturer's instruction. Molar concentration of *Pf*ACMSD monomer was determined by using 39 kDa as the molecular weight (14).

Circular Dichroism Spectroscopy. The CD spectra (190 - 250 nm) of ACMSD and mutant proteins were acquired on a JASCO J-810 spectropolarimeter (JASCO, Easton, MD, USA) at ambient temperature. In each measurement, a protein sample (10 μM) was placed in a 1-mm path length quartz cell in 0.05 M potassium phosphate, pH 7.5. All spectra were the average of ten scans with a scan rate of 50 nm min^{-1}.

Preparation of ACMS and Enzyme Activity Assay. ACMS was generated from 3-hydroxy-anthranilic acid using 3-hydroxyanthranilate 3,4-dioxygenase containing no free transient metal ion as reported previously (15). Specific activities of ACMSD proteins were measured in triplicate at room temperature on an Agilent 8453 diode-array spectrophotometer by monitoring the absorbance of ACMS at 360 nm as described previously (15).

Electronic Spectroscopy. UV-Vis absorption spectra were obtained in 25 mM HEPES-NaOH buffer, pH 7.0, and 5% glycerol at room temperature on a Cary 50 spectrophotometer. Reduced Fe(II)-H228Y for UV-Vis and resonance Raman experiments was prepared by adding 13 μL 30 mM sodium dithionite or L-ascorbate (~ 2 equivalents per Fe) to 150 μL 1.4 mM Fe(III)-H228Y in an 1-mm path length cuvette.

Quinone-based NBT Staining. Purified proteins were first electrophoresized on a 12% SDS–PAGE gel and then transferred onto a nitrocellulose membrane using the Mini Trans-Blot Cell Assembly (Bio-Rad). The transblotting was conducted in the transfer buffer (25 mM Tris-HCl, 192 mM glycine, and 20% methanol) at 100 V for 1 h. The proteins were temporarily visualized by 0.1% ponceau S in 5% acetic acid (133). After the temporary stains were removed by washing

with water, the protein-containing nitrocellulose membrane was immersed in a solution of 0.24 mM NBT and 2 M potassium glycinate, pH 10, in the dark for 45 min to visualize the quinone-containing protein band (134). This method was designed to identify quinonoid compounds, such as DOPA, and is a convenient assay for quinone-containing proteins (135). Purified methylamine dehydrogenase (MADH) from *Paracoccus denitricans* (a generous gift from Dr. Victor L. Davidson), a protein consisting of two large α-subunits and two small β-subunits with a tryptophan tryptophyquinone prosthetic group in each of its small subunit (136), was used as both of a positive and a negative control in Ponceau S staining and NBT staining experiments.

Electronic Paramagnetic Resonance (EPR) Spectroscopy. HEPES buffer (25 mM, pH 7.0) containing 10% glycerol was used in the EPR sample preparations. X-band EPR spectra were obtained in perpendicular mode on a Bruker (Billerica, MA) EMX spectrometer at 100-kHz modulation frequency using a 4119HS high-sensitivity resonator. Low temperature was maintained with an ITC503S temperature controller, an ESR910 liquid helium cryostat, and a LLT650/13 liquid helium transfer tube (Oxford Instruments, Concord, MA). Measurements were conducted by keeping the frequency of the electromagnetic radiation constant while the magnetic field was swept.

Raman Spectroscopy. Resonance Raman spectra were collected on an Acton AM-506 spectrophotometer (1200 groove rating) using Kaiser Optical Systems holographic supernotch filters and a Princeton Instruments liquid N_2-cooled CCD detector (LN-1100PB) with 4 cm^{-1} spectral resolution. HEPES buffer (25 mM, pH 7.0) containing 5% glycerol was used in the sample preparations. Spectra were collected at 100 mW laser power using a Spectra Physics 1060-KR-V krypton ion laser (568.2 and 647.1 nm) or a 2030-15 argon ion laser (488.0 and 514.5 nm). Raman frequencies were referenced to indene standard with an accuracy of ± 1 cm^{-1}.

Raman spectra were collected at 4 °C by 90° scattering from a flat bottomed NMR tube with 30-60 minutes to collect each window of spectra. The full spectra were collected in two overlapping windows; the florescent background was subtracted before normalizing spectra at nonresonance enhanced protein phenylalanine ring vibration at 1004 cm^{-1} and splicing normalized spectra together. Baseline corrections (polynomial fits) were applied using Gram/32 Spectral Notebook (ThermoGalactic). Excitation profiles were constructed by comparing peak intensities of resonance enhanced bands of normalized spectra to nonresonance enhanced vibration at 1004 cm^{-1}. Difference resonance Raman spectra of Fe(III)-H228Y were obtained by subtracting the spectrum of WT Zn-ACMSD both normalized at 1004 cm^{-1} vibration followed by a base line correction.

X-ray Data Collection and Crystallographic Refinement. Co(II)-H228Y, Zn(II)-H228Y and Zn(II)-H228G ACMSD were crystallized following the condition previously established for WT ACMSD by hanging drop vapor diffusion in VDX plates (Hampton Research). Single crystals suitable for X-ray data collection were obtained from drops assembled with 1 μL protein solution layered with 1 μL reservoir solution containing 0.1 M Tris-Hcl (pH 8.75), 0.2 M $MgCl_2$, and 15% PEG 5000. Reservoir solution for Fe(III)-H228Y was modified and contained 0.1 M Tris (pH 7.0), 0.2 M $MgCl_2$, and 17% PEG 5000. Crystals were frozen in liquid nitrogen after being dipped into the cryoprotectant solution that contained 30% of glycerol or ethylene glycol in the mother liquid. X-ray diffraction data were collected at SER-CAT beamline 22-ID or 22-BM of the Advanced Photon Source (APS), Argonne National Laboratory, Argonne, IL. Data were collected at 100 K using a beam size matching the dimensions of the largest crystal face. The data were processed with HKL-2000 (137). Structure solutions were obtained by molecular replacement using MolRep (138) from the CCP4i program suite (139) with the entire WT Co(II)-

ACMSD structure (PDB entry 2HBX) for Co(II)/Fe(III)-H228Y ACMSD and WT Zn(II)-ACMSD structure (PDB entry 2HBV) for Zn(II)-H228Y/Zn(II)-H228G ACMSD as the search models. Refinement was carried out using REFMAC (140) in the CCP4i program suite (139) for Fe(III)- and Co(II)-H228Y and PHENIX software for Zn(II)-H228Y and Zn(II)-H228G, and model-building was carried out in COOT (141). Restrained refinement was carried out using no distance restraints between the metal center and its ligands. Residue Tyr228 and Gly228 were well-ordered and added to the model based on the $2F_o$-F_c and F_o-F_c electron density maps. Refinement was assessed as complete when the F_o-F_c electron density contained only noise.

2.4 Results

Biochemical properties of the His228 mutants—To probe its role in the enzyme mechanism, His228 was mutated to tyrosine, which also has a ring structure and is capable of hydrogen bonding, and glycine, which effectively deletes the side chain and thus eliminates hydrogen bonding. Both variants were expressed as soluble proteins. Surprisingly, the H228Y mutant purified from cells grown in LB medium exhibits a blue chromophore while wild-type (WT) Zn-ACMSD is colorless (Figure 2.2). Metal analysis of the blue H228Y-ACMSD by ICP-OES spectroscopy showed the enzyme contained 0.51 ± 0.013% Fe, 0.08 ± 0.003 Zn and 0.001 ± 0.0004 Cu ions per polypeptide chain, whereas the wild-type enzyme prepared under the same experimental conditions contained only zinc (80). Likewise, the as-isolated H228G mutant contained mainly Fe ion (0.32 ± 0.001 per polypeptide chain) and trace amount of zinc. A sharp ferric EPR signal is seen at $g = 4.27$ from the two as-isolated His228 mutants (not shown). These results suggest His228 is an important determinant of metal selectivity in this enzyme.

The as-isolated H228G and H228Y were catalytically inactive regardless of whether the Fe ion is in the ferrous or ferric oxidation state. Zinc- and cobalt-substituted variants were also tested. H228Y was inactive in both Zn(II) and Co(II) forms, as was Zn(II)-H228G. These results suggest that His228 is essential for the enzyme catalytic activity.

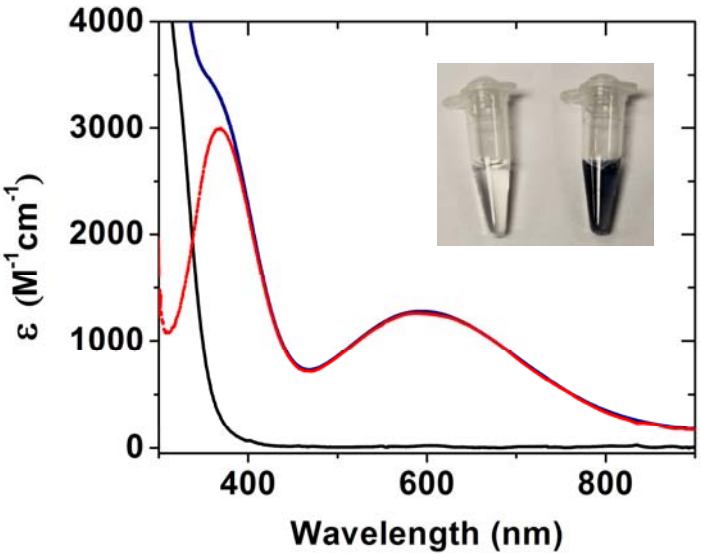


Figure 2.2. UV-Vis spectra of Fe-H228Y. The as-isolated Fe(III)-H228Y protein(dark blue trace), its reduced form (black) and the different spectrum (red) showing two absorption at 370 nm and 595 nm. Photographs of the as-isolated Fe(III)-H228Y (right) and reduced Fe(II)-H228Y (left) are shown in the inset.

The blue color in H228Y is not due to formation of a metal-bound dihydroxyphenylalanine—Two possibilities were considered for the origin of the blue chromophore of the as-isolated Fe-containing H228Y. The first option is an oxygenated Tyr may be produced by post-translational modification resulting from oxygen activation by the redox active iron. Examples can be found for a number of non-heme iron enzymes (135,142,143). However, this option is not supported by A quinone-based staining assay (135) using reversible ponceau S staining of a nitrocellulose membrane (See Supporting information Figure S2.1). Alternatively, it could be due to the coordination of Tyr228 to the active site metal, as tyrosinate-metal interactions have been

observed in many other enzymes, such as purple acid phosphatases (144). phospholipase D (145), and catalases (146).

Resonance Raman characterization suggests an Fe(III)-tyrosinate chromophore in as-isolated Fe-H228Y—The blue Fe-H228Y mutant exhibits two UV-Vis absorption bands at 370 and 595 nm, with respective extinction coefficients of 3000 and 1300 M^{-1} cm^{-1} (Figure 2.2). While the 370 nm feature is a common spectral characteristic of the Fe(III)-bound nonheme Fe proteins, the blue chromophore at 595 nm is within the range of the Fe(III)-tyrosinate LMCT band with an absorption maximum that is typically between 410 – 600 nm (1000 – 2500 $M^{-1}cm^{-1}$ per phenolate ligand) (147). In contrast, an iron(III)-catecholate LMCT chromophore typically exhibits two lower energy bands between 400 – 580 nm and 550 – 900 nm (2000 – 2500 $M^{-1}cm^{-1}$), these are not present in Figure 2.2. The blue chromophore vanished upon the addition of sodium dithionite or L-ascorbate to the protein, suggesting that reduction of the Fe(III) metal center to Fe(II) is linked with the loss of the intense ligand-to-metal charge-transfer (LMCT) transition.

Figure 2.3A shows resonance Raman spectra of the as-isolated blue Fe-H228Y acquired with 568.2 nm laser excitation using 90° scattering geometry compared to the spectra of reduced Fe-H228Y, WT Zn-ACMSD and the buffer (25 mM HEPES pH 7.0 with 5% glycerol by volume) collected under the same conditions. Vibrations associated with the buffer solution and protein were subtracted by taking the difference spectra of the as-isolated blue Fe-H228Y and WT Zn-ACMSD after normalizing both spectra by scaling to non-resonance-enhanced protein phenylalanine ring vibration at 1004 cm^{-1} (Figure 2.3B). The difference spectrum shows resonance-enhanced vibrations at 573, 792, 873, 1164, 1290, 1503, and 1604 cm^{-1}. The

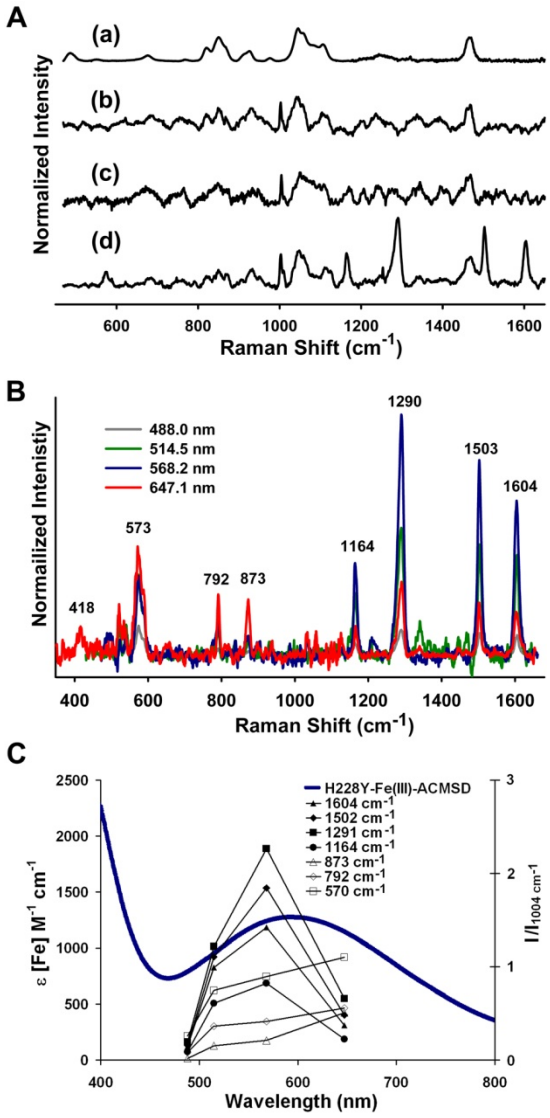


Figure 2.3. Resonance Raman spectra of H228Y-ACMSD. (A) Resonance Raman spectrum with 568.2 nm laser excitation of (a) as-isolated blue Fe(III)-H228Y, (b) native Zn-ACMSD, (c) Fe(II)-H228Y ACMSD, and (d) Fe(III)-H228Y ACMSD. (B) Difference resonance Raman spectra of Fe(III)-H228. (C) Excitation profiles for resonance enhanced bands.

excitation profile of the 1164, 1290, 1503, and 1604 cm^{-1} features shown in Figure 2.3C illustrates maximum enhancement of these vibrations near the absorption maximum of the blue chromophore at 595 nm, while the 573, 792, and 873 cm^{-1} vibrations are more enhanced at 647.1 nm. The spectrum of the reduced form of the colorless Fe(II)-H228Y does not exhibit any resonance-enhanced vibrations but closely resembles the WT Zn-ACMSD (Figure 2.3A).

The resonance enhanced vibrations of Fe(III)-H228Y at 573, 1164, 1290, 1503, and 1604 cm^{-1} are collectively regarded as the Raman signature for a Fe(III)-tyrosinate chromophore (Table 2.1) (147-159). These vibrations are distinct from those expected for a Fe(III)-catecholate chromophore, which has characteristic vibrations observed at 1270, 1320, 1425, and 1475 cm^{-1} (Table 2.1) (160-164). By analogy to previous work on nonheme Fe proteins the 573 cm^{-1} can be assigned to the Fe-$(O)_{Tyr}$ vibration, while the higher frequency vibrations are associated with aromatic ring modes. The 792 and 873 cm^{-1} vibrations may originate from a Fermi doublet formed by the mixing of the fundamental υ_1 symmetric ring-breathing vibrational mode and first overtone υ_{16a} non-planar ring vibration (165-168). Solid L-tyrosine exhibits a Fermi doublet with peaks at 803 and 845 cm^{-1} as well as the υ_{16a} fundamental at 419 cm^{-1} (166). The two peaks observed for the blue Fe-H228Y have a much larger splitting of 81 cm^{-1}. An even larger splitting has been observed for porcine uteroferrin (803 and 873 cm^{-1}), while the υ_{16a} fundamental was not detected (165,168). The large splitting of tyrosine Fermi doublet was rationalized as a result of strong covalence of the Fe(III)-tyrosinate metal ligand bond based on the fact that it was resonance enhanced and that its center of gravity (~ 838 cm^{-1}) is similar to that of L-tyrosine (~ 824 cm^{-1}) (165). Upon 647.1 nm laser excitation of the blue Fe-H288Y, a weak peak is observed at 418 cm^{-1} that may be assigned to the fundamental υ_{16a} vibration.

Table 2.1. Resonance Raman vibrations and LMCT bands of nonheme iron protein and model complexes with Fe(III)-phenolate and Fe(III)-catecholate chromophores.

Complex[a]	**λ_{max} (nm) (ε, $M^{-1}cm^{-1}$)**	**Resonance Raman Vibrations (cm^{-1})**									**Reference**
Fe(III)-phenolate complexes											
H228Y-Fe(III)-ACMSD	**370(3000), 595(1255)**		**573**	**792**	**873**	**1164**	**1290**		**1503**	**1604**	**this work**
Fe(II)-HPPD + O_2	595(2600)		583	751	881	1171	1290		1502	1600	(148,169)
Porcine uteroferrin	545(1000)		575	803	873	1173	1293		1504	1607	(165,168)
[Fe(salhis$_2$)]ClO_4	530(4100)					1159 1132	1337 1310		1476 1452	1625 1605	(153)

Fe(III)-serotransferrin	470(2500)			759	828	1174	1288		1508	1613	(156,159)
Fe(III)-lactoferrin	465(2070)					1170	1272		1500	1604	(157,158)
Fe(III)-ovotransferrin	465(2000)			759	860	1173	1264		1501	1603	(151)
Catechol 1,2-dioxygenase	450(3000–4000)			757	872	1175	1289		1506	1604	(153,154)
Protocatechuate 3,4-dioxygenase	435(3000), 525		592	756	826 854	1172 1180	1254 1266		1506	1604	(149,152)
[Fe(salen)(OC_6H-4-CH_3)]	410		568			1168	1272		1501	1603	(147,155)
Fe(III)-catecholate complexes											
Fe(III)TyrH-dopamine	415(1700), 695(2000)	528	592	631		1275	1320	1425	1475		(163)
Blue PMI	420(1500), 680(2100)		591	631		1266	1330	1428	1482		(161)
Fe(III)-MndD-DOPA	675(750)	530	569 586	646	666	1273	1318	1423	1464		(170)
Fe(II)TauD-αKG +O_2	550(700)	544	580	623	644	1261	1314	1425	1475		(162,171)

[a]Abbreviations used in this table: ACMSD, α-Amino-β-carboxymuconate-ε-semialdehyde decarboxylase from *Pseudomonas fluorescens*; salhis, *N*-[2-(4-imidazolyl)ethyl]salicylaldimine; salen, N,N'-ethylenebis (salicy1idenamine) dianion; MndD, homoprotocatechuase 2,3-dioxygenase from *Arthrobacter globiformis*; DOPA, dihydroxyphenylalanine; PMI, phosphomannose isomerase from *Candida albicans* expressed in *Escherichia coli*; TauD taurine/α-ketoglutarate dioxygenase from *Escherichia coli*; HPPD, (4-hydroxyphenyl) pyruvate dioxygenase from *Pseudomonas fluorescens*; α-KG, α-ketoglutarate.

Fe(III)-tyrosinate LMCT bands in the visible region typically range from 410–600 nm where the energy of the LMCT transition can be correlated with the Lewis acidity and the Fe(III/II) redox potential of the iron center and can give additional information about the nature of the other coordinating ligands (147). The low energy of the Fe(III)-tyrosinate LMCT band of Fe-H228Y at 595 nm and the observation that Fe-H228Y is easily reduced by mild reducing agents are consistent with a Lewis acidic metal center with weak field ligands (148). Post-translationally modified blue (4-hydroxyphenyl) pyruvate dioxygenase (HPPD) has a similarly low-energy Fe-tyrosinate LMCT band at 595 nm (148,169). From its crystal structure (172), unmodified HPPD has an iron center coordinated by two histidines and one glutamate, but no tyrosinate ligand. However it does have several phenylalanine amino acid residues in the second sphere, one of which has been proposed to be self-hydroxylated to form the tyrosinate ligand in blue-HPPD

(164,172). Based on the similar energies of their LMCT bands, blue Fe(III)-H228Y and the self-hydroxylated HPPD are likely to have comparable ligand environments with two or three histidine residues and a carboxylate group from either an aspartate or glutamate. Along with the fact that H228Y no longer picks up Zn(II), the Fe(III) is likely to be coordinated by His9, His11, His177, Asp294 and Tyr228 residues in the ACMSD active site.

Spectroscopic characterization of Co(II)-substituted H228Y— Co(II)-H228Y was obtained by adding supplemental $CoCl_2$ to the M9 minimum growth medium and purification buffer (see experimental procedures). ICP-OES spectroscopic analysis confirmed that H228Y isolated with supplemented cobalt contains 0.98 ± 0.02 cobalt and 0.05 ± 0.03 iron per protein monomer. The cobalt-substituted H228Y displays a brown chromophore that is darker than the color of the WT Co(II)-substituted ACMSD (Figure 2.4, top), The brown chromophore originates from forbidden metal d-d transitions, the intensities of which can be correlated to the metal coordination number. Decreasing the coordination number increases the intensity of the d-d bands due to greater p-d orbital mixing. Typically, six-coordinate high-spin Co(II) complexes have extinction coefficients of less than 50 $M^{-1}cm^{-1}$ above 500 nm, while five- and four-coordinate complexes have values between 50 – 250 M^{-1} cm^{-1} and greater than 300 $M^{-1}cm^{-1}$, respectively (173-175). The visible spectrum of the brown Co(II)-substituted H228Y is similar to that of WT Co(II)-ACMSD above 500 nm (Figure 2.4, top), consistent with a 5-coordinate Co(II) center (176). However, the spectrum below 500 nm is distinct from that of the wild-type protein (Figure 2.4, top). The absorption peaks at 355 (680 ± 20 $cm^{-1}M^{-1}$) and 420 nm (430 ± 10 $cm^{-1}M^{-1}$) in WT ACMSD are respectively red shifted to 380 (590 ± 30 $cm^{-1}M^{-1}$) and 458 (490 ± 20 $cm^{-1}M^{-1}$) in the Co(II)-substituted H228Y. These differences indicate that the Co(II) center in the H228Y mutant is in a weaker ligand field environment than in WT Co(II)-substituted ACMSD. This observation is

consistent with the substitution of a water ligand by a stronger π-donating ligand in Co(II)-substituted H228Y due to Tyr228 ligation to the metal as suggested by resonance Raman study for Fe-H228Y in as-isolated H228Y protein.

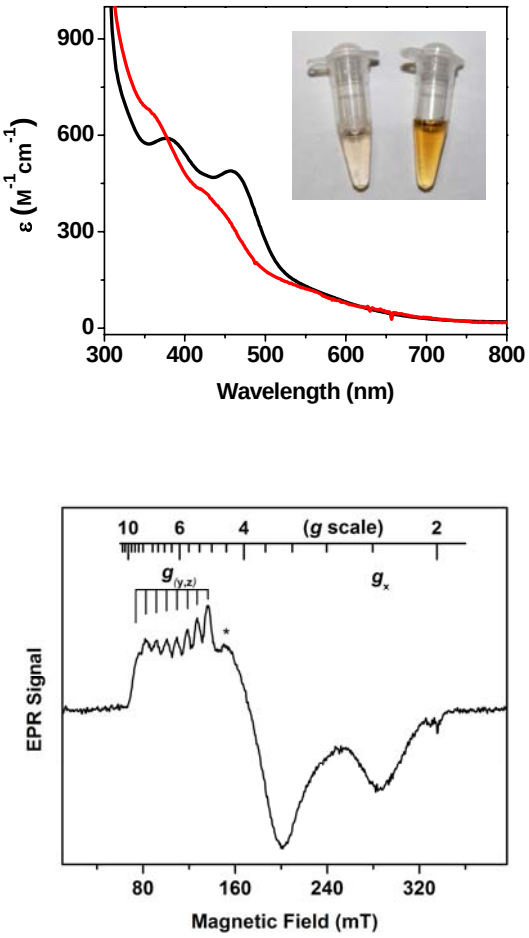


Figure 2.4. UV-Vis and EPR spectra of Co(II)-H228Y-ACMSD. Top panel: UV-Vis spectra of Co(II)-H228Y ACMSD (black) and Co(II)-substituted WT ACMSD (red). Photographs of the as-isolated Co-WT (left) and Co-H228Y (right) are shown in the inset. Bottom panel: EPR spectra of Co(II)-H228Y ACMSD obtained at 3G modulation, 0.25 mW microwave power, and 10K. The asterisk indicates a g=4.27 signal from a small amount of Fe ion present in the sample due to a change in the metals preference of H228Y.

Electron paramagnetic resonance (EPR) spectroscopy was employed to characterize the electronic structure of the metal center in Co(II)-substituted H228Y. The EPR spectrum of Co-

H228Y exhibits an $S = 3/2$ spectrum with a well-resolved eight-line ^{59}Co hyperfine interaction pattern associated with the low-field feature (Figure 2.4, bottom), consistent with a high-spin Co(II) ion bound specifically to the protein. An Fe(III) EPR signal is also present at $g = 4.27$, showing that a minor fraction of protein still binds iron even under the cobalt-rich conditions. The well-resolved ^{59}Co hyperfine structure of Co-H228Y is observable even at 10 – 20 K, but only a featureless cobalt signal can be observed in WT Co-ACMSD at this temperature (14). The *g* value of the high-field feature of the Co-H228Y mutant is also different from that of the WT protein, *i.e.*, 2.343 versus 2.613, indicating that the ligand geometry around the metal ion in H228Y is different from that of the WT enzyme (14). The optical and EPR data suggest that the high-spin cobalt center in the brown H228Y is distinct from the cobalt center in the wild-type enzyme. The spectral features are most reminiscent of high-spin five-coordinate cobalt centers with square pyramidal geometry, such as that found in inhibitor derivatives of cobalt carbonic anhydrase (177).

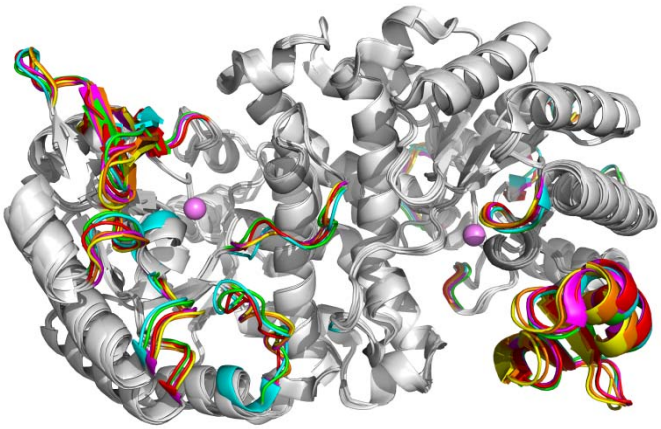


Figure 2.5. Superimposed overall structures of ACMSD containing different metals. WT Zn(II)-ACMSD in blue, WT Co(II)-ACMSD in magenta, Fe(III)-H228Y ACMSD in orange, Co(II)-H228Y in yellow, Zn(II)-H228Y in green, and Zn(II)-H228G in red.

X-ray crystal structure—While the spectroscopic data strongly suggest a tyrosinate ligation to Fe in the as-isolated H228Y protein, it remains unclear that if Tyr228 is a metal ligand in Co(II)-substituted H228Y. The zinc-containing mutants do not have chromophore or EPR features. In order to observe directly what changes at the enzyme active site arise from the point mutation and metal substitution, we obtained single crystals from the as-isolated Fe(III)-H228Y, Co(II)-substituted H228Y, Zn(II)-substituted H228Y, and Zn(II)-reconstituted H228G proteins. The X-ray crystal structures of Fe(III)-H228Y (2.8 Å, PDB entry: 4ERG), Co(II)-H228Y (2.4 Å resolution, PDB entry: 4ERA), and Zn(II)-H228G (2.6 Å, PDB entry: 4EPK) were solved in the $P4_22_12$ space group, while Zn(II)-H228Y (2.0 Å, PDB entry: 4ERI) was determined in the $C2$ space group. Data collection and refinement statistics are presented in Table 2.2. Figure 2.5 shows the superimposition of the four structures we determined against the previously reported WT Zn(II)-ACMSD and Co(II)-substituted ACMSD data. As previously seen in the wild-type enzyme, the mutants of *Pf*ACMSD exhibits a homodimeric quaternary structure. The substitution of His228 with either Tyr or Gly has little impact on the folding of the overall TIM-barrel scaffold, regardless of the metal identity. The circular dichroism spectra of these variants were nearly identical to that of the WT enzyme (not shown), consistent with the structural data. In Figure 2.5, structural variations relative to WT ACMSD are highlighted in different colors. These structural differences are mainly observed in the previously defined mobile insertion domain, which is more flexible compared with other parts of the structure (80). Limited structural differences in a few loop regions are observed due to changes of metal ligation in the metal-substituted mutants. Therefore, we conclude that no significant change in secondary structure was introduced due to the mutation and metal substitution. The organization of amino acids residues surrounding the metal center is very similar in all four structures with one

exception of one of the metal ligands, Asp294, which tilts towards the imidazole ring of His9 to various degrees (Figure 2.6).

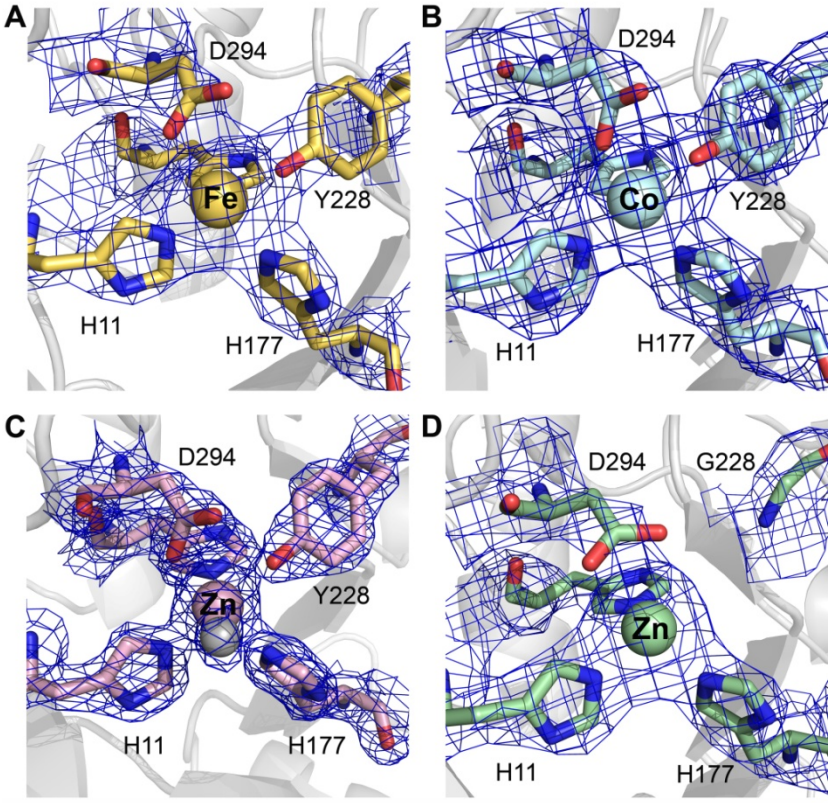


Figure 2.6. The active site structures of ACMSD containing different metals. (A) As-isolated Fe(III)-H228Y, (B) Co(II)-substituted H228Y, (C) Zn(II)-H228Y, and (D) Zn(II)-substituted H228G. The $2F_o - F_c$ electron densities are taken from chain A and contoured at 1.5σ. The larger spheres represent metal ions. The smaller gray sphere in panel C represents a water molecule, which is weakly bound to the zinc ion (2.44 Å) as compared to the zinc−water ligand distance of 2.04 Å in the wild-type enzyme (PDB entry 2HBV). Such a water ligand is missing from the metal center in all other H228G and H228Y structures.

Table 2.2. X-ray crystallography data collection and refinement statistics

Data collection	**Fe-H228Y**	**Co-H228Y**	**Zn-H228Y**	**Zn-H228G**
Detector type	MAR300 CCD	MAR225 CCD	MAR300 CCD	MAR300 CCD
Source	APS, Sector 22-ID	APS, Sector 22-BM	APS, Sector 22-ID	APS, Sector 22-ID
Space group	$P4_22_12$	$P4_22_12$	C2	$P4_22_12$
Unit cell lengths (Å)	a=b=90.43, c=167.53	a=b=90.03, c=167.31	a=153.84, b=48.56, c=110.20	a=b=91.50, c=170.12
Unit cell angles (˚)	α=β=γ=90º	α=β=γ=90º	α=γ=90º, β=126.89º	α=β=γ=90º

Wavelength (Å)	1.00	1.00	1.00	1.00
Temperature (K)	100	100	100	100
Resolution (Å) [a]	2.80 (2.85)	2.40 (2.44)	2.00 (2.03)	2.60 (2.64)
Completeness (%)[a]	99.8 (99.8)	99.2 (94.1)	92.7 (54.7)	98.4 (68.7)
R_{merge} (%)[a, b]	10.1 (69.7)	10.1 (61.0)	6.1 (22.8)	7.7 (65.6)
I/σI [a]	60.8 (2.5)	40.5 (2.9)	51.1 (6.0)	63.1 (1.9)
Multiplicity [a]	24.2 (9.8)	13.9 (9.6)	6.3 (3.4)	18.6 (7.2)
Refinement				
Resolution (Å)	2.8	2.4	2.0	2.6
No. reflections; working/test	17041/915	26155/1382	41014/2073	22859/1173
R_{work} (%)[c]	19.9	21.5	20.5	20.8
R_{free} (%)[d]	27.3	29.5	25.6	28.1
Protein atoms	5194	5194	5194	5153
Ligand atoms	2	2	3	3
Solvent sites	9	88	209	32
Ramachandran statistics[e]				
Preferred	93.48	92.73	94.52	91.78
Allowed (%)	5.61	5.61	4.72	7.61
Root mean square deviation				
Bond lengths (Å)	0.011	0.011	0.008	0.009
Bond angles (°)	1.299	1.464	1.098	1.204
PDB access code	4ERG	4ERA	4ERI	4EPK

[a] Values in parentheses are for the highest resolution shell.
[b] $R_{merge} = \Sigma_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \Sigma_i I_{hkl,i}$, where I is the observed intensity and $\langle I_{hkl} \rangle$ is the average intensity of multiple measurements.
[c] $R_{work} = \Sigma||F_o|-|F_c|| / \Sigma|F_o|$, where $|F_o|$ is the observed structure factor amplitude, and $|F_c|$ is the calculated structure factor amplitude.
[d] R_{free} is the R factor based on 5% of the data excluded from refinement.
[e] Based on values attained from refinement validation options in COOT.

Inspection of the metal center shows that Tyr228 is a metal ligand in all three H228Y mutant proteins (Figure 2.6), with Tyr-to-metal bond lengths of 2.13, 2.12, and 2.41 Å in chain A, and 2.26, 2.16, 2.44 Å in chain B in the Co-, Fe-, and Zn-H228Y structures, respectively. It would appear that the interaction between the Zn ion and the Tyr228 residue weaker than corresponding interactions with Fe and Co. Nonetheless, the angles of the Tyr228 plane, phenolic oxygen and the metal, including zinc, are well within the range of typical metal-ligand angular orientations (110 – 140°) reported in biological systems (130.2°, 137.3°, and 128.7° in chain A, and 129.7°, 132.5°, and 122.6° in chain B in the Co-, Fe-, and Zn-H228Y structures, respectively) (178). Hence, the prediction of Tyr228 ligation by our spectroscopic work is confirmed by these as-isolated and metal-substituted H228Y structures. It should be noted that the metal center in chain

B in the mutant structures shows a noticeable degree of disorder. This was the same as previously observed in the as-isolated and cobalt-substituted wild-type ACMSD structures (2HBV & 2HBX) (80).

The most significant structural difference with respect to the WT structure is that the water ligand is missing from the metal center in Fe(III)-H228Y, Co(II)-substituted H228Y, and Zn(II)-reconstituted H228G proteins. Modeling of the omit maps (Figure S2.2) show the electron density can only be fitted by the metal ion in these structures. However, the resolution is not sufficiently high to assign the solvent content in the Fe(III)-H228Y and Zn(II)-H228G data sets, the missing water ligand cannot be reliably concluded based on these structures. In contrast, Zn(II)-H228Y and Co(II)-substituted H228Y structures contain reasonable amount of water molecules. The water ligand is missing from the cobalt structure, whereas, the water ligand is still present in the Zn(II)-substituted H228Y structure but at a much longer distance from the metal, 2.44 Å in chain A and 2.38 Å in chain B from the zinc ion as compared to 2.05 and 2.07 Å in WT Zn-ACMSD structure. Our EPR, resonance Raman and optical data suggest a low coordination number in the Co(II)- and Fe(III)-H2228Y metal center. Taken together, it is likely that the water ligand is missing, or dissociated, from the metal center in Co(II)- and Fe(III)-H228Y mutants. Moreover, the hydrogen bonding network that potentially activates the water ligand is significantly altered. In the WT Zn-ACMSD crystal structure the water ligand hydrogen bonds to both His228 and Asp294, while in the H228Y mutant the distance between Asp294 and the water ligand is elongated to 3.60 Å in chain A and 3.67 Å in chain B, ,compared to 2.88 and 3.33 Å in the WT Zn-ACMSD. In the H228G structure, the metal center is still five-coordinate even though the water ligand is absent. Asp294 becomes a bidentate ligand to the metal ion.

From the H228G structure and the three H228Y structures, we conclude that His228 plays an important role in metal ion recruitment and maintaining the position of the water ligand.

2.5 Discussion

His228 is a major determinant of metal ion selectivity in PfACMSD—ACMSD was thought to be a cofactor-free enzyme for about 50 years until we found many divalent transition metal ions can effectively activate it (14). In the subsequent study, we found that the as-isolated protein contains only zinc ion (80). However, the molecular basis of metal selectivity in proteins is generally poorly understood and the preference for zinc within the active site of ACMSD has remained a mystery prior to this work. In the present study, a point mutation of a second coordination sphere His228 residue to Gly or Tyr is described to change *Pf*ACMSD's metal ion selectivity from Zn to Fe ion. In general, the primary control of metal selectivity in proteins should be the metal-binding ligand set. A single site mutation on a non-metal ligand residue changes metal selectivity in a protein is a surprising result. To our knowledge, the finding described in this work for metal selection role of a second coordination sphere residue is only the second example in literature. Previously, mutation of a conserved glutamate in *Escherichia coli* manganese superoxide dismutase has been shown to change metal preference to Fe ion (179). In all other documented cases the residue identified for affecting metal selectivity is a metal ligand. The lesson learned from *Pf*ACMSD His228 is anticipated to improve our understanding of metal ion selectivity in proteins and expand the scope of roles histidine plays in the enzyme active site.

The origin of blue color in H228Y—To understand the underlying reason for the loss of enzyme activity due to mutation of His228, a spectroscopic study was performed to interrogate the metal center's chemical and electronic structure in H228Y. The optical and NBT-staining data suggest the ligation of Tyr228 to the Fe ion in the as-isolated Fe(III)-H228Y, and this notion

is further supported by our resonance Raman results. These data also eliminate the possibility of post-translational modification of Tyr228 to a DOPA.

The role of His228 in maintaining the hydrolytic water ligand forcatalysis—His228 attracted our attention because it is a strictly conserved active site residue. It lies on the opposite side of the presumed substrate binding pocket (80,131). A previous structural study has revealed the conformational diversity of this active site residue (80). In the present work, His228 was mutated to tyrosine and glycine. In line with our expectations, neither of the protein variants had a measureable catalytic activity; regardless of which metal was incorporated into the active site. The biochemical and spectroscopic data suggest that the loss of catalytic activity is, in large part, due to the missing histidine that acts as an acid/base catalyst and, the loss of the water ligand. Thus, a dual role of His228 is revealed, *i.e*., it stabilizes the water ligand while plays an important role in metal selectivity of the enzyme.

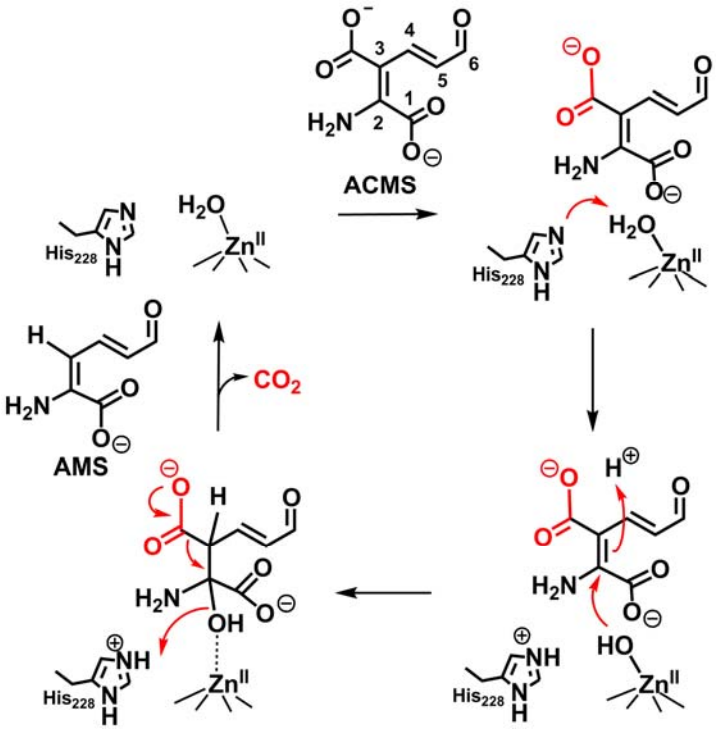
Tyrosinate ligation to the metal center was identified by resonance Raman spectroscopy and further supported by the EPR data. The UV-Vis absorption spectra of both WT and Co(II)-substituted H228Y suggest five-coordinate metal centers. Therefore, the water ligand in the wild-type enzyme is likely replaced by the Tyr228 residue in the mutant. In the crystal structures of H228Y proteins loaded with three different metal ions, Tyr228 is a metal ligand. The spectroscopic data and the crystallographic results are consistent to each other.

A water ligand is observed in all published structures of ACMSD, including Co(II)- and Zn(II)-*Pf*ACMSD, as well as Zn(II)-*h*ACMSD proteins. Such a water ligand is also present in other members of the ACMSD subfamily, including 4-oxalomesaconate hydratase, and γ-resorcylate decarboxylase (21,80,180). The structures of Fe(III)-H228Y, Co(II)-H228Y and H228G are the first structures determined that do not show a water ligand, while Zn(II)-H228Y shows a weakly

bound water ligand with a longer distance to the metal center compared to the WT protein. Because of its weak binding to the Zn center, the pK_a of the water ligand in Zn(II)-H228Y may not be decreased as effectively as in the WT proteins by coordination to the zinc ion. Cobalt and zinc ions are particularly strong Lewis acids and can dramatically reduce the pK_a of bound water ligands. In a well-established example of carbonic anhydrase, the bound Zn(II) ion decreases the pK_a of its water ligand from 15.7 to 7 (181-183). Catalysis ensues from this zinc activation of its bound water. His228 is strictly conserved within all known ACMSD amino acid sequences, and in all available structures of the enzyme this residue is in hydrogen bonding distance to the water ligand. The primary role of this His228 is proposed as deprotonation of the metal-bound water so that a hydroxide ion can be generated as an active site nucleophile at ca. pH 6 – 7, the optimum pH range for ACMSD. Although the hydroxide attack mechanism model shown in Scheme 2.1 for ACMSD is distinct from other metal-dependent decarboxylases, it shares common features with the established mechanisms in the amidohydrolase superfamily (19).

The results described in this work are consistent with our working model of the ACMSD reaction (Scheme 2.1). In the working mechanistic model shown in Scheme 2.1, the metal-bound water molecule is deprotonated to hydroxide anion (OH^-) with the assistance of His228, which is acting as a general acid/base catalyst to deprotonate the zinc-bound water molecule for attack on the substrate. The hydroxide ion performs a nucleophilic attack on the C2=C3 double bond of ACMS with concomitant protonation at C3. Essentially, a water molecule is added across the double bond with the hydroxyl group at C2 and the new proton at C3, generating a substrate-based tetrahedral intermediate. Collapse of the tetrahedral intermediate initiates the decarboxylation and produces α-aminomuconate-ε-semialdehyde and regenerate the metal

center. This decarboxylation model follows a hydrolytic mechanism consistent with the mechanistic paradigm of the amidohydrolase superfamily.



Scheme 2.1. The proposed catalytic role of His228 in ACMSD reaction.

2.6 Supporting Information

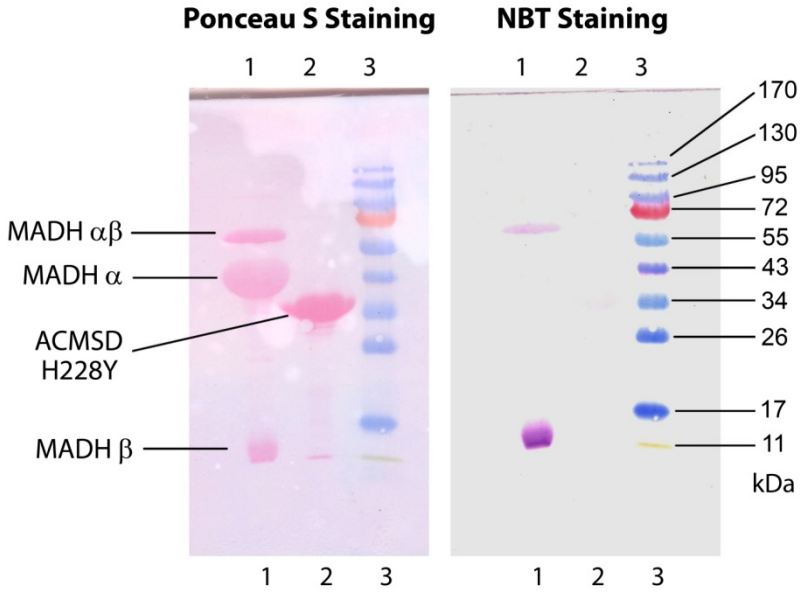


Figure S 2.1. Ponceau S (left) and nitroblue tetrazolium (NBT) quinone (right) staining of purified MADH and blue H228Y ACMSD from *Pseudomonas fluorescences*. Lane 1: MADH; lane 2: blue H228Y ACMSD; lane 3: pre-stained protein molecular weight ladder. The quinone-based staining assay was carried out to test the potential presence of the quinonoid compounds, including L-3,4-dihydroxyphenylalanine (DOPA). Methylamine dehydrogenase (MADH) from *Paracoccus denitrificans* was used as a control. Reversible ponceau S staining of the nitrocellulose membrane reveals three bands on the MADH lane corresponding to the β-subunit, α–subunit, and a small amount of αβ heterodimer, while there is only one single band on the H228Y lane under the SDS-PAGE conditions (left panel), demonstrating that MADH and blue Fe-H228Y were both purified to near homogeneity. After thoroughly rinsing away the ponceau S red dye, the same membrane was subjected to the quinone-based NBT staining. Only the small β-subunit and the partially dissociated αβ fraction of MADH were visualized by the NBT staining, while Fe-H228Y and the large α-subunit of MADH were both unstained (right panel). Blue Fe-H228Y was insensitive to specific staining. Hence, the NBT staining data led us to eliminate the possibility of a quinonoid compound in blue Fe-H228Y.

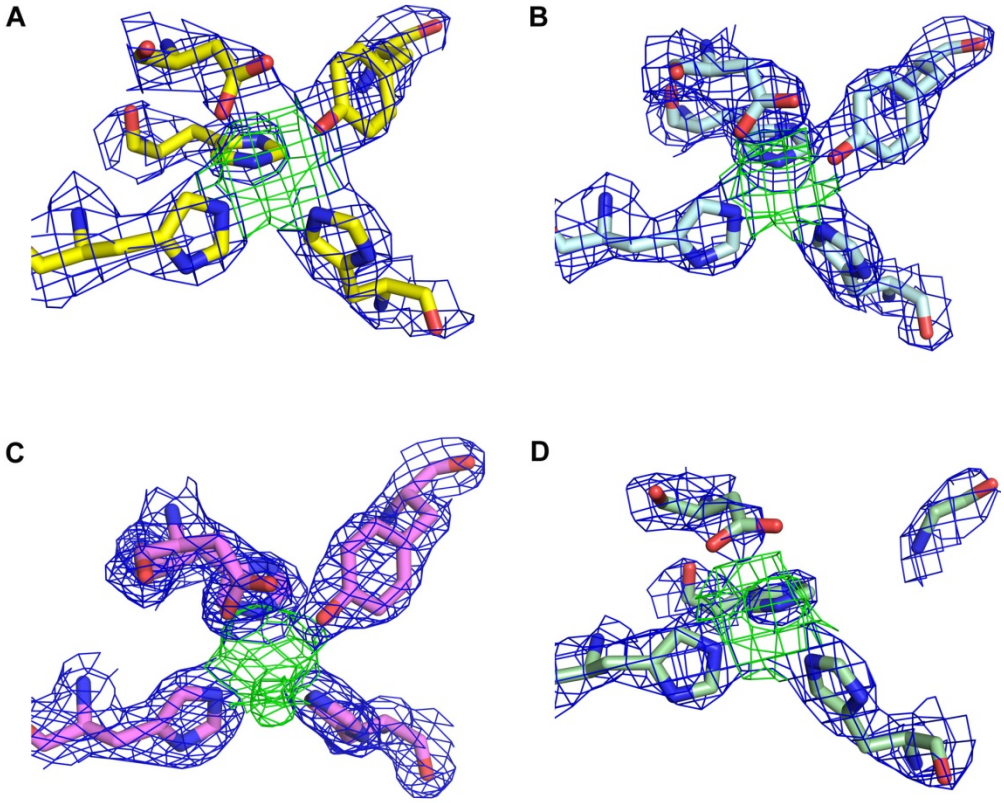


Figure S 2.2. Omit maps (green) generated from the $2F_o$-F_c maps (blue) after removing metal ion and any assigned water ligand(if any). The electron density is contoured at 1.5 σ. (A) as-isolated Fe(III)-H228Y (PDB entry: 4ERG), (B) Co(II)-substituted H228Y (4ERA), (C) Zn(II)-H228Y (4ERI), and (D) Zn(II)-substituted H228G (4EPK).

2.7 Acknowledgements

We thank Dr. Victor L. Davidson for the generous gift of a control sample used in the quinone-based NBT staining, and Drs. Xiaodong Cheng and John Horton for using the X-ray crystal screening facility at Emory University. X-ray data were collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID and 22-BM beamlines at the Advanced Photon Source, Argonne National Laboratory. We thank Dr. Bi-Cheng Wang and the SER-CAT staff at section 22 for assistance with remote data collections. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38.

CHAPTER 3 THE POWER OF TWO: ARGININE 51 AND ARGININE 239* FROM A NEIGHBORING SUBUNIT ARE ESSENTIAL FOR CATALYSIS IN α-AMINO-β-CARBOXYMUCONATE-ε-SEMIALDEHYDE DECARBOXYLASE

The section of chapter 3 is a published work on ACMSD: The Power of two: arginine 51 and arginine 239* from a neighboring subunit are essential for catalysis in α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase. Lu Huo, Ian Davis, Lirong Chen, and Aimin Liu. (2013) J. Biol. Chem., 288(43), 30862-30871. This work was supported in part or whole by NSF grant MCB-0843537 and Georgia Research Alliance Distinguished Scholar Program (AL). X-ray data was collected at Southeast Regional Collaborative Access Team (SER-CAT) 22-ID and 22-BM beamlines at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng. We thank Dr. B.-C. Wang and the SER-CAT staff scientists of the Section 22 of Argonne National Laboratory for assistance in beamline preparation and X-ray data collection.

3.1 Abstract

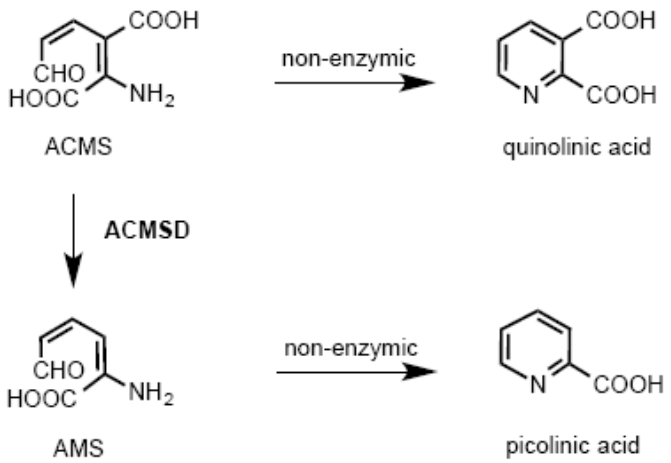
While the crystal structure of α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) from *Pseudomonas fluorescens* was solved as a dimer, this enzyme is a mixture of monomer, dimer, and higher order structures in solution. In this work we found that the dimeric state, not the monomeric state, is the functionally active form. Two conserved arginine residues are present in the active site: Arg51 and an intruding Arg239* from the neighboring subunit. In this study, they were each mutated to alanine and lysine, and all four mutants were catalytically inactive. The mutants were also incapable of accommodating pyridine-2, 6-dicarboxylic acid, a competitive inhibitor of the native enzyme, suggesting that the two Arg residues are involved in

substrate binding. It was also observed that the decarboxylase activity was partially recovered in a heterodimer hybridization experiment when inactive R51A/K and R239A/K mutants were mixed together. Of the 20 crystal structures obtained from mixing inactive R51A and R239A homodimers which diffracted to a resolution lower than 3.00 Å, two structures are clearly R51A/R239A heterodimers and belong to the *C*2 space group. They were refined to 1.80 and 2.00 Å resolutions, respectively. Four of the remaining crystals are apparently single mutants and belong to the $P4_22_12$ space group. In the heterodimer structures, one active site is shown to contain dual mutation of Ala51 and Ala239* while the other contains the native Arg51 and Arg239* residues, identical to the wild-type structure. Thus, these observations provide the foundation for a molecular mechanism by which the oligmerization state of ACMSD could regulate the enzyme activity.

3.2 Introduction

Arginine residues perform diverse functions in proteins due to their charge distribution, capability of forming multiple hydrogen bonds and their long flexible side chains. One popular function is the formation of salt bridges in which the positively charged guanidine head pairs with a negatively charged carboxylate group (184,185). In enzymes, arginine residues are often proposed to participate in recognizing, binding and maintaining the orientation of substrates and reaction intermediates (186-188). Arginine is also involved in RNA recognition because its guanidium group is able to form five hydrogen bonds with RNA acceptor groups (189). Although not widespread, arginine residues have also recently been proposed to act as a general base in several enzymes (190-192).

Two arginine residues in α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD) caught our attention because they appear to be in an optimal position to bind the substrate α-amino-β-carboxymuconate-ε-semialdehyde (ACMS) by interacting with its two carboxyl groups. ACMSD is a metal-dependent enzyme in the amidohydrolase superfamily (14,15), and its substrate, ACMS, is an unstable tryptophan catabolic intermediate in the kynurenine pathway (2,14,129). The conversion of ACMS to α-amino-muconate-ε-semialdehyde (AMS) is catalyzed by ACMSD, however ACMS spontaneously forms quinolinic acid (QA) in a non-enzymatic dehydration process in the absence of ACMSD (125,193) (Scheme 3.1). Thus, ACMSD is a key enzyme that determines the QA levels. QA is maintained only at the basal level for the de novo synthesis of a small fraction of NAD^+ in mammals. Reduced activity of ACMSD leads to accumulation of QA in body fluids, a condition which is known to be related to several neurodegenerative diseases including Alzheimer's disease (127), Huntington's disease (128), stroke (2), and epilepsy (2,126,129). However, the instability of ACMS and the compounds with similar structures makes studying the mechanism of ACMSD substrate binding technically challenging.



Scheme 3.1. ACMSD in metabolic pathway

The first crystal structure of ACMSD was solved as a homodimer from *Pseudomonas fluorescens* (80). Two strictly conserved arginine residues are present at the active site: Arg51 and Arg239* (Figure 3.1). Arg51 is located in the mobile insertion domain, which is proposed to undergo conformational changes upon substrate binding because of its unique position and high flexibility (80). Arg239* is an intruding residue which is inserted from the sixth α-helix of the other subunit and it is labeled with a * hereafter to distinguish its origin from the chain of a neighboring subunit. Because of its high flexibility, this residue is shown to be present in alternative conformations in chain B of the crystal structure.

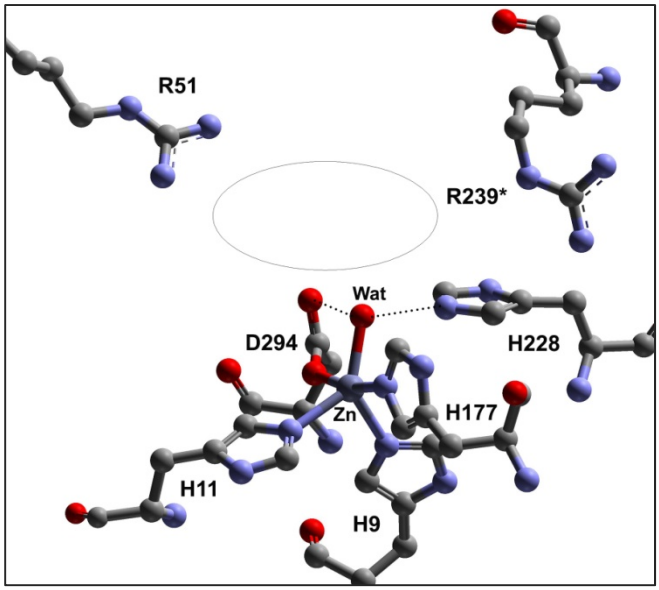


Figure 3.1. The active site metal center of *Pf*ACMSD and the putative substrate-binding pocket (illustrated with an ellipse).

The second ACMSD crystal structure is the human enzyme in complex with 1,3-dihydroxyacetone-phosphate. This inactive complex was assigned as a monomer by Rizzi and coworkers (131). The equivalent residue of Arg51 is shown in the active site, but the counterpart of Arg239 is far removed from the catalytic center. The same group previously reported that functional human ACMSD is a monomer because gel-filtration chromatography showed the asisolated enzyme with a molecular mass around 50 kDa (194), which is between monomer (38

kDa) and dimer (76 kDa) mass. In fact, it is still unclear whether the catalytically active form of human ACMSD in solution is a monomeric or dimeric enzyme at the present time. All of the *Pf*ACMSD structures determined thus far, including wild type and mutants charged with different transition metals, are dimeric (25,80). If the monomeric form is catalytically active, then only one conserved arginine residue, Arg51, is present in the active site of the enzyme in both human ACMSD and *Pf*ACMSD structures. Arg239 of the same subunit is not located near the enzyme active site in the monomeric state and would therefore be less likely to be an essential component of catalysis. Thus, point mutation at residue 239 provides a sensible means for probing the effects of protein oligmerization on the catalytic activity.

The present work represents the first attempt to investigate the molecular mechanism by which the oligmerization state of ACMSD could determine the enzyme activity. The quaternary state required for catalytic activity is reported, and the catalytic role of the two conserved arginine residues, Arg51 and the intruding Arg239* from the neighboring subunit, are proposed. Interestingly, we found that the inactive Arg51 and Arg239 mutants could form a crystallographically accessible, catalytically-active heterodimer through protein hybridization.

3.3 Experimental Procedures

Site-Directed Mutagenesis. R51A, R51K, R239A, and R239K single mutation mutants were constructed by using the PCR overlap extension mutagenesis technique (132) described previously (15) and by using the plasmid containing ACMSD from *Pseudomonas fluorescens* as a template (14). The mutants were verified by DNA sequencing to ensure that base changes were introduced correctly and that no random changes had occurred. After sequencing, the positive clone was used for protein expression in Escherichia coli BL21(DE3).

Protein Preparation, SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and Size Exclusion Chromatography. Expression and purification of ACMSD proteins followed the protocol established previously (10,14,15,25) with a few minor changes. Rather than using pH 8.0 buffer during affinity chromatography purification, pH 7.4 was used to increase protein stability of the mutants. Metal-free apo-enzyme was obtained by treating purified protein with 5 mM EDTA for 12 h followed by running a G50 desalting column. Zn(II) and Co(II) containing ACMSD were reconstituted as previously described (14). Purified protein samples were mixed with SDS loading buffer and heated for 10 min or with no heating prior to loading to 10% SDS-PAGE. ACMSD was loaded onto a Superdex 75 column pre-equilibrated with buffer containing 25 mM HEPES (pH 7.0) and 5% glycerol (v/v). Protein was eluted with the same buffer.

Enzyme Activity Assay. The substrate ACMS was generated enzymatically by catalyzing the dioxygenation of 3-hydroxyanthranilic acid in oxygen saturated buffer with ferrous 3-hydroxyanthranilate 3,4-dioxygenase containing no free transient metal ion as reported previously (14,15,25). Briefly, the ACMSD enzyme assay mixture contained 5 – 120 μM ACMS and an appropriate amount of ACMSD protein in 25 mM HEPES buffer with 5% glycerol, pH 7.0. Specific activities of ACMSD proteins were measured in triplicate at room temperature on an Agilent 8453 spectrophotometer by monitoring the decrease of ACMS absorbance at 360 nm using a molar extinction coefficient of 47,500 $M^{-1}cm^{-1}$ (195) when the substrate concentration was equal to or lower than 20 μM, or at 320 nm using a molar extinction coefficient of 9600 $M^{-1}cm^{-1}$ when the substrate concentration was higher than 20 μM (15).

Derivation of Equations for Analysis of the Correlation of Monomer/Dimer Ratio and Catalytic Activity. In order to show that dimerization is required for catalytic activity, a series of activity assays were performed in which the stock concentration of ACMSD is varied while

keeping the final assay concentration constant. If the rate of dissociation of the ACMSD dimers is sufficiently slow on the time scale of the assay, then the measured activity should conform to eq 1 where k_{obs} is the observed decarboxylation rate, k_{max} is the maximum rate, and %D is the mole ratio of dimer present in the assay.

$$k_{obs} = k_{max} \cdot \%D \quad (1)$$

In order to solve for the percent dimer, mass balance was invoked to define M_0 as the total protomer concentration, M as the monomer concentration, and D as the dimer concentration.

$$if\ M_0 = M + 2D,$$

$$then\ M = M_0 - 2D \quad (2)$$

The dimer percentage can then be written in terms of dimer concentration and total protomer concentration (a quantity which is known in an activity assay) alone.

$$\%Dimer = \frac{2D}{M_0} \quad (2')$$

The association constant for this equilibrium, can then be written in terms of dimer concentration and total protomer concentration by substitution of eq 2, as seen in eq 3.

$$K_a = \frac{D}{(M_0 - 2D)^2} \quad (3)$$

Distribution and collection of like terms then gives eq 4, and eq 4 can be multiplied by M_0^{-2} to give eq 5.

$$4K_aD^2 - (4K_aM_0 + 1)D - K_a{M_0}^2 = 0 \quad (4)$$

$$K_a\frac{4D^2}{{M_0}^2} - \left(2K_a + \frac{1}{2M_0}\right)\frac{2D}{M_0} - K_a = 0 \quad (5)$$

The quadratic equation can now be invoked on eq 5 to solve directly for the dimer percentage and is shown in eq 6. Multiplying the right side of eq 6 by a disguised 1, ($2M_0$/ $2M_0$), the expression for the dimer percentage can be simplified to eq 7.

$$\frac{2D}{M_0} = \frac{\left(2K_a+\frac{1}{2M_0}\right)\pm\sqrt{\left(2K_a+\frac{1}{2M_0}\right)^2-4K_a{}^2}}{2K_a} \quad (6)$$

$$\%\, D = 1+\frac{1\pm\sqrt{8K_aM_0+1}}{4K_aM_0} \quad (7)$$

Substitution of eq 7 back into eq 1 gives k_{obs} as a function of k_{max}, K_a, and M_0 (i.e., eq 8). Non-linear least squares fitting can then be used to obtain both K_a and k_{max} (though both positive and negative square roots should be considered; only the negative roots gave convergent fits of the measured data).

$$k_{obs} = k_{max}\cdot\left(1+\frac{1-\sqrt{8K_aM_0+1}}{4K_aM_0}\right) \quad (8)$$

Protein Hybridization. Arg51 and Arg239 mutants were mixed in a 1:1 ratio in 25 mM HEPES buffer, pH 7.0, containing 5% glycerol. The final protein concentration for each mutant was 12 μM. Specific activity of the mixtures was examined frequently after mixing by a spectrophotometric assay on an Agilent 8453 diode-array spectrophotometer as described above. *Identification of a Competitive Inhibitor.* Pyridine-2,6-dicarboxylic acid (PDC) inhibition pattern and constant were determined by measuring k_{cat} and K_m in the presence of varying inhibitor concentrations (0, 20, 40, and 80 μM). Apparent K_m values for each inhibitor concentration were plotted as a function of inhibitor concentration and inhibition is expressed as K_i values in micromolar (K_i = intercept/slope).

Analysis of PDC Inhibitor Interaction with the Arginine Mutants. PDC in HEPES buffer displays UV absorbance at 270 nm with extinction coefficient of 4,100 $M^{-1}cm^{-1}$. ACMSD proteins (1.5 ml, 40 μM in each sample type), including wild type, R51A, R239A, H228G, and premixed R51A and R239A for 48 h, were incubated with one equivalent of PDC for 5 min in separate experiments, followed by filtration in an Amicon ultra centrifugal filter tube with cellulose membrane of 10,000-dalton cut-off. PDC (40 μM) with no protein was used in the

control experiment. Samples were centrifuged in 3,000 × g for 8 min. The flow-through solution containing unbound DPA was quantified by Agilent 8453 spectrophotometer.

X-ray Data Collection and Crystallographic Refinement. Zinc-reconstituted mutant proteins were used for crystallization. All single crystals were obtained following the crystallization conditions described in the previous studies (25,80). For the heterodimer crystal, R51A and R239A were mixed for more than 48 h at 4 °C before being used for crystallization by hanging drop vapor diffusion in VDX plates (Hampton Research). Mother liquid containing 20% ethylene glycol was used for crystal mounting. X-ray diffraction data were collected at SER-CAT beamline 22-ID/BM of the Advanced Photon Source (APS), Argonne National Laboratory, Argonne, IL. Data were processed with HKL-2000 (137). Structure solutions were obtained by molecular replacement with MolRep (138) from the CCP4 suite (139) with wild-type ACMSD (PDB code 2HBV) as the search model (80). Refinement was carried out using PHENIX software (196), and the model-building was carried out in COOT (141). Ala51 and Ala239 were modeled based on the $2F_o$-F_c and F_o-F_c electron density maps. Refinement was assessed as complete when the F_o-F_c electron density contained only noise.

3.4 Results

ACMSD Does not Exist as a Single Species in Solution. As shown in Figure 3.2A, ACMSD presents several bands by SDS-PAGE that correspond to monomer, dimer, and higher order oligmerization states if the sample is not heated before loading onto the gel (Figure 3.2A lane 1). However, if the sample is heated before loading, the bands corresponding to the higher order states and most of the dimer are not observed (Figure 3.2A lane 2). The size-exclusion chromatography profile, on the other hand, presents two major peaks. The molecular weight of the protein eluted in the first peak is larger than 100 kDa, and the second peak exhibits a

molecular weight of ca. 51 kDa. The molecular weight of a *Pf*ACMSD monomer, however, is 38 kDa. Also, there is no difference in the specific activity of the two fractions as they elute from the column. The second peak shows two major bands corresponding to ACMSD monomer and dimer molecular weights on SDS-PAGE when not heated before loading (Figure 3.2A lane 3). The band which arises from the dimer is not observed upon heating of the sample before loading (Figure 3.2A lane 4). Hence, we assign the second peak as a combination of ACMSD monomer and dimer and the first peak as a mixture of higher order complexes which are not able to be resolved further.

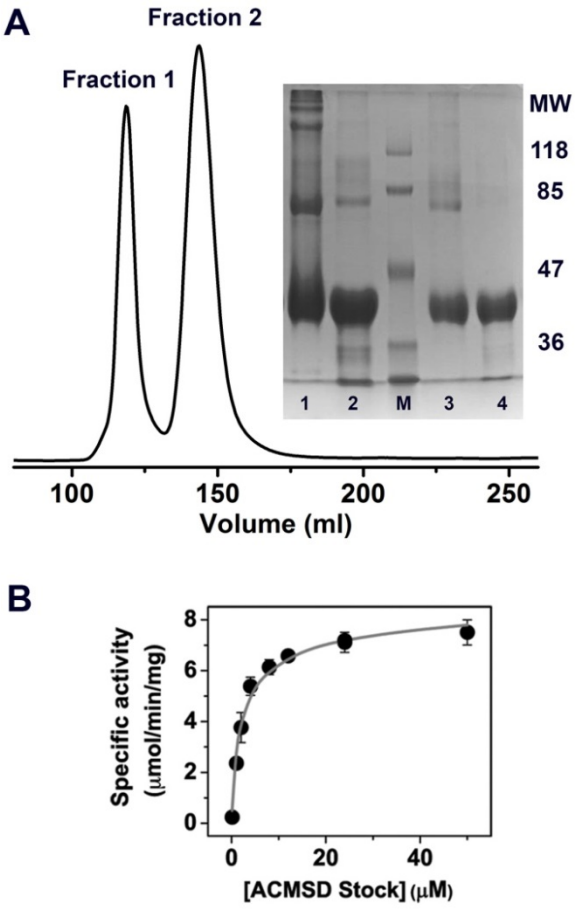


Figure 3.2. ACMSD is a mixture of monomer and higher oligormers in solution and is not active as monomer. (A) The UV absorbance at 280 nm of ACMSD as a function of the elution volume from size exclusion chromatography column: Superdex 75. The inset is an SDS-PAGE. Lane 1: As-isolated ACMSD with no heating, 2: as-isolated ACMSD heated for 10 min, M: molecular marker, 3: fraction 2 with no heating, 4: fraction 2 heated for 10 min. (B) The specific activity of ACMSD as a function of different enzyme stock concentration.

In order to estimate the dimer dissociation constant, ACMSD was prepared in stock concentrations ranging from 0.1 to 50 μM (protomer concentration) and allowed to equilibrate at 4 °C for 24 h. Specific activity of the different stock concentrations was determined by keeping the final enzyme and substrate concentrations constant. The specific activity increased with increasing stock concentration. Decarboxylation rates as a function of enzyme stock concentration were plotted in Figure 3.2B and data were fit to eq 8, which represents the equilibrium of ACMSD between inactive monomer and active dimer forms as described in the Experimental Procedures. Fitting of the data to eq 8 gives a K_a of 3.3×106 M^{-1}, and a maximum specific activity of 9.4 μmol•min^{-1}•mg^{-1}.

Mutational Analysis and Catalytic Activity Rescue Through Protein Hybridization. Arg51 andArg239 were each mutated to alanine and lysine. All of the four mutants were found to be completely inactive with either zinc or cobalt as the cofactor, the two metals which were previously found to be optimal for catalytic activity in the wild-type enzyme. Chemical rescue performed on R51A and R239A by guanidine derivatives was unsuccessful in restoring enzymatic activity. However, when the two mutants, R51A and R239A, were mixed together, enzyme activity was slowly restored (Figure 3.3A). Reaction rates were determined at several time points after mixing. Samples incubated for 1, 7, and 24 h on ice were analyzed by following decay of the substrate ACMS at 360 nm for 1 min. Reaction with no enzyme and the two inactive arginine mutants were used as negative controls. The rate of activity rescue was temperature dependent, and the rate of recovery was faster when the mixing reaction was carried out at room temperature rather than on ice. Specific activity was calculated and plotted as a function of incubation time as shown in Figure 3.3B and fit to a single exponential association,

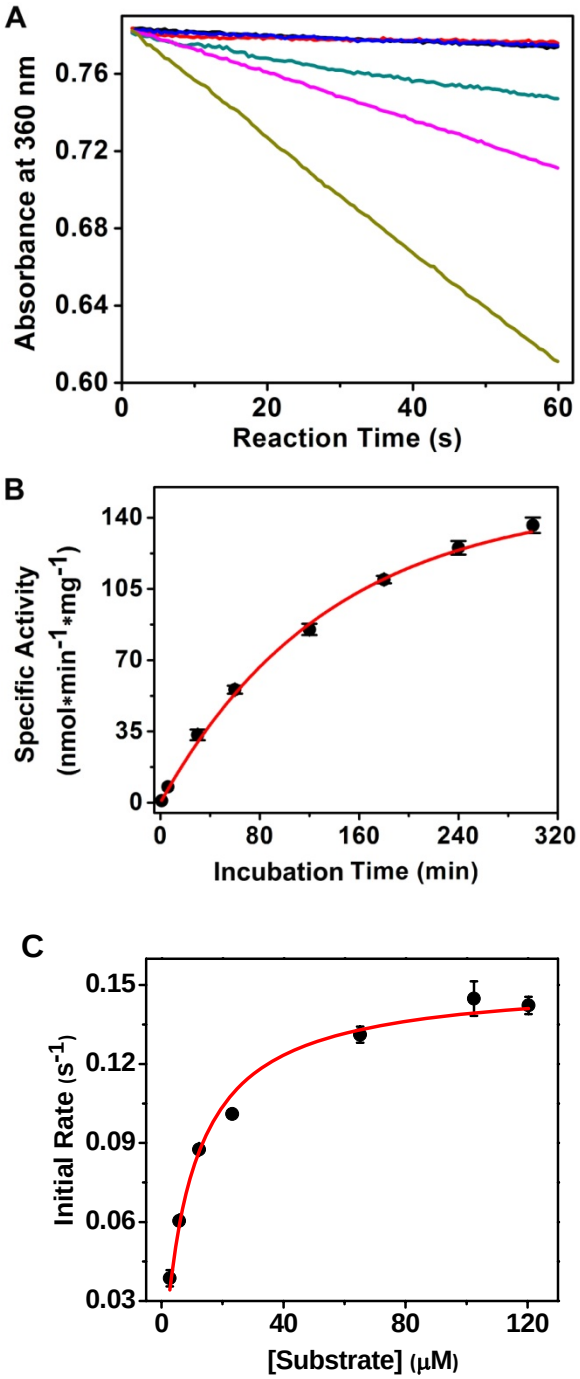


Figure 3.3. Activity rescue by mixing R51A and R239A. (A) Catalytic activity rescue of mixed R51A and R239A mutants. Decay of substrate ACMS in the first 1 min was monitored by UV-Vis at 360 nm. Autodecay of substrate (black trace), in the presence of R51A (red), and in the presence of R239A (blue), are three overlapped lines. R51A and R239A were pre-mixed on ice for 1 hour (green), 7 hour (pink), and 24 hours (tan), respectively. (B) The rescued activity increased as a function of incubation time of the two mutants. Data were fitted to a single exponential equation. The incubation and kinetic reaction are done in room temperature. (C) Michaelis-Menten kinetics of the 1:1 mixed R51A and R239A after 48 hours of equilibration.

eq 9, where A is the measured specific activity, Amax is the maximum activity, *k* is the associative rate constant, and t is time in minutes. According to the fitting, the maximum rate is 152 ± 5 nmol•min^{-1}•mg^{-1}, and the rate constant is 7.2 ± 0.5 × 10^{-3} min^{-1}. We also determined kinetic parameters for the mixed zinc-reconstituted mutants after 48h of equilibration on ice. Figure 3.3C shows the Michaelis-Menten plot. The K_m, 9.3 ± 1 μM, of the mixed mutants is similar to that of the wild-type Zn-ACMSD, 9.6 μM. The kcat, 0.15 ± 0.01 s^{-1}, of the mixed mutants, on the other hand, is dramatically decreased compared to the 6.5 s^{-1} of the WT Zn-ACMSD (80).

$$A = A_{max} \cdot (1 - e^{-kt}) \quad (9)$$

Structural Determination of R51A, R239A and the Hybridized Mixtures of the Two Mutants. All crystals of R51A and R239A ACMSD belong to the $P4_22_12$ space group. They diffracted poorly and only at relatively low resolutions. The best structures obtained were 2.50 Å resolution forR51A (Protein Data Bank entry: 4IFO) and 2.40 Å for R239A (4IFR). Interestingly, crystals from the mixture of these two mutants were crystallized in both the $P4_22_12$ and *C*2 space groups. A total of twenty data sets were collected from individual single crystals obtained from crystallization trays with equal molar ratio mixing of R51A and R239A mutant proteins. Among the twenty crystal structures that we have solved, three of them are R239A homodimer, and one of them is an R51A homodimer. All of these four structures are determined in the $P4_22_12$ space group. Two single crystals were found to belong to the *C*2 space group, and they were refined to 1.8 Å (Protein Data Bank entry: 4IG2) and 2.0 Å (4IFK) resolutions, respectively. The structures determined from these two crystals clearly differ from R51A or R239A, and they fit the proposed heterodimer as illustrated in Figure 3.4. Fourteen of the data sets in the $P4_22_12$ space group cannot be assigned to any homodimer or heterodimer due to incomplete electron density which

might be caused by the flexibility of the arginine side chain or low arginine occupancy. The data collection and refinement statistics for both the homodimer single mutants and the heterodimers are listed in Table 3.1.

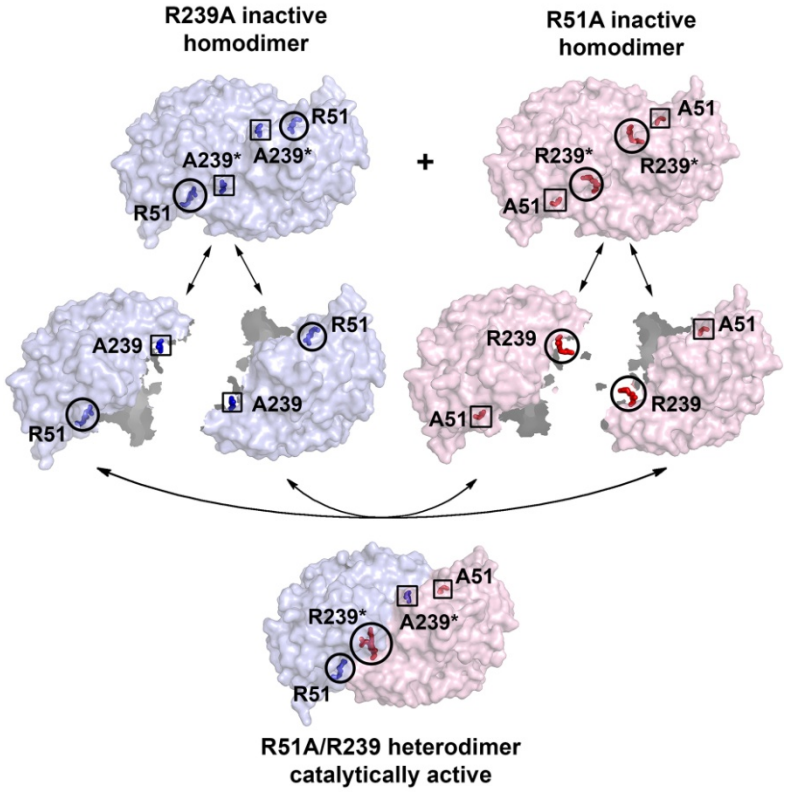


Figure 3.4. Molecular mechanism for catalytic activity rescue through protein hybridization of the two completely inactive mutants.

Both DNA sequencing and electron density map of the mutant crystal structures demonstrate the successful substitution of arginine residues with alanine in both R51A and R239A variants Figure 3.5A (top and middle panels). The alignment of R51A (4IFO) and R239A (4IFR) structures with the wild-type structure (2HBV) are shown in Figure 3.5B (top and middle panels), respectively. Both mutants exhibit homodimeric quaternary structures as previously seen in wild-type ACMSD. Secondary structure changes caused by mutation (colored in blue and cyan) can be observed in both subunits of the two structures and they are mainly observed in the loop regions. No significant structural deviations, however, were observed for the folding of the

Table 3.1. X-ray crystallography data collection and refinement statistics.

Data collection	**R51A-R239A Heterodimer #1**	**R51A-R239A Heterodimer #2**	**R51A Homodimer**	**R239A Homodimer**
Detector type	MAR225 CCD	MAR300 CCD	MAR300 CCD	MAR300 CCD
Source	APS, Sector 22-BM	APS, Sector 22-ID	APS, Sector 22-ID	APS, Sector 22-ID
Space group	*C2*	*C2*	$P4_22_12$	$P4_22_12$
Unit cell lengths (Å)	a=153.10, b=48.52,c=109.83	a=153.85, b=49.06,c=110.55	a=b=91.21, c=170.47	a=b=91.16, c=167.99
Unit cell angles (°)	α=γ=90, β=126.80	α=γ=90, β=127.10	α=β=γ=90	α=β=γ=90
Wavelength (Å)	1.00	0.80	0.80	0.80
Temperature (K)	100	100	100	100
Resolution (Å) a	35.00-1.80 (1.86-1.80)	35.00-2.00 (2.03-2.00)	35.00-2.50 (2.54-2.50)	35.00-2.40 (2.44-2.40)
Completeness (%)a	86.1(42.7)	97.6 (87.6)	99.9(100.0)	99.6(95.7)
Rmerge (%)a, b	5.0(60.0)	10.7(47.8)	7.8(64.0)	9.9(79.3)
I/σI a	37.2(1.5)	28.4(1.9)	34.9(2.9)	50.0(1.9)
multiplicity a	6.8(3.7)	5.9(2.9)	15.8(11.4)	22.0(7.5)
Refinement				
Resolution (Å)	1.80	2.00	2.50	2.40
No.reflections; working/test	51829/2590	42682/2143	25616/1304	28612/1452
Rwork (%)c	23.2	20.9	19.7	21.0
Rfree (%)d	28.2	26.3	24.4	29.5
No. of protein atoms	5174	5184	5162	5144
No. of ligand atoms	2	3	2	2
No. of solvent sites	141	229	100	26
Average B factor (Å2)				
Protein	49.0	39.7	54.0	73.5
Zn(II)	36.9	31.3	55.2	64.8
Mg(II)	N/A	60.4	N/A	N/A
Solvent	40.8	38.7	42.5	59.1
Ramachandran statisticse				
preferred (%)	93.01	94.07	93.76	91.32
allowed (%)	4.56	4.10	4.11	5.18
Root mean square deviation				
Bond lengths (Å)	0.009	0.008	0.010	0.009
Bond angles (°)	1.249	1.149	1.175	1.236
PDB entry	4IG2	4IFK	4IFO	4IFR

[a] Values in parentheses are for the highest resolution shell.
[b] $R_{merge} = \Sigma_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \Sigma_i I_{hkl,i}$, where $I_{hkl,i}$ is the observed intensity and $\langle I_{hkl} \rangle$ is the average intensity of multiple measurements.
[c] $R_{work} = \Sigma||F_o|-|F_c||/\Sigma|F_o|$, where $|F_o|$ is the observed structure factor amplitude, and $|F_c|$ is the calculated structure factor amplitude.

[d] R_{free} is the R factor based on 5% of the data excluded from refinement.
[e] Based on values attained from refinement validation options in COOT.

overall TIM-barrel scaffold, indicating that the impact due to mutation is limited to specific flexible regions. We further analyzed the dimerization interface using the ACMSD wild-type and mutant structures at an online server: Protein Data Bank Europe Protein Interfaces, Surfaces and Assemblies (PDBePISA) (197). The results indicate the interface contains a total of 13 H-bonds and 13 salt bridges. Nevertheless, these analyses would not be sufficient to draw a definitive conclusion if there were no intruding residue Arg239 catalytically important which allowed us to generate an active heterodimer from mixing two inactive dimers. In the wild-type structure, Arg239 contributes two Hbonds by interacting with the main chain molecules of the neighboring monomers to help to stabilize the dimer. In the R239A structure, the dimerization is not interrupted because these two hydrogen bonds are substituted by the following two new H-bonding interactions: Glu252 of chain A interacts with Lys189 of chain B and Arg247 of chain A interacts with Met191 of chain B.

Figure 3.5B (bottom panel) shows the superimposition of the heterodimer structure (4IFK) and the wild-type Zn-ACMSD structure (2HBV). The structural deviations of the mutant relative to the wild-type are highlighted in red. The comparison clearly shows that major changes can only be observed in chain B. This observation is expected because in subunit A of the heterodimer, Arg51 and Arg239* are both in place with electron density similar to that in the wild-type structure (Figure 3.5A bottom). The electron density for both arginine residues' side chains in subunit B is missing, indicating that both residues 51 and 239* are alanines (Figure 3.5A bottom). Hence, subunit A represents a complete native-like ACMSD subunit, but subunit B is a "double mutant" containing two alanine residues. It is noted in the structure of the native enzyme that Arg239* adopts two conformations because of its high flexibility (80). Likewise,

this residue in subunit A has alternative conformations which are only partially covered by the electron density map.

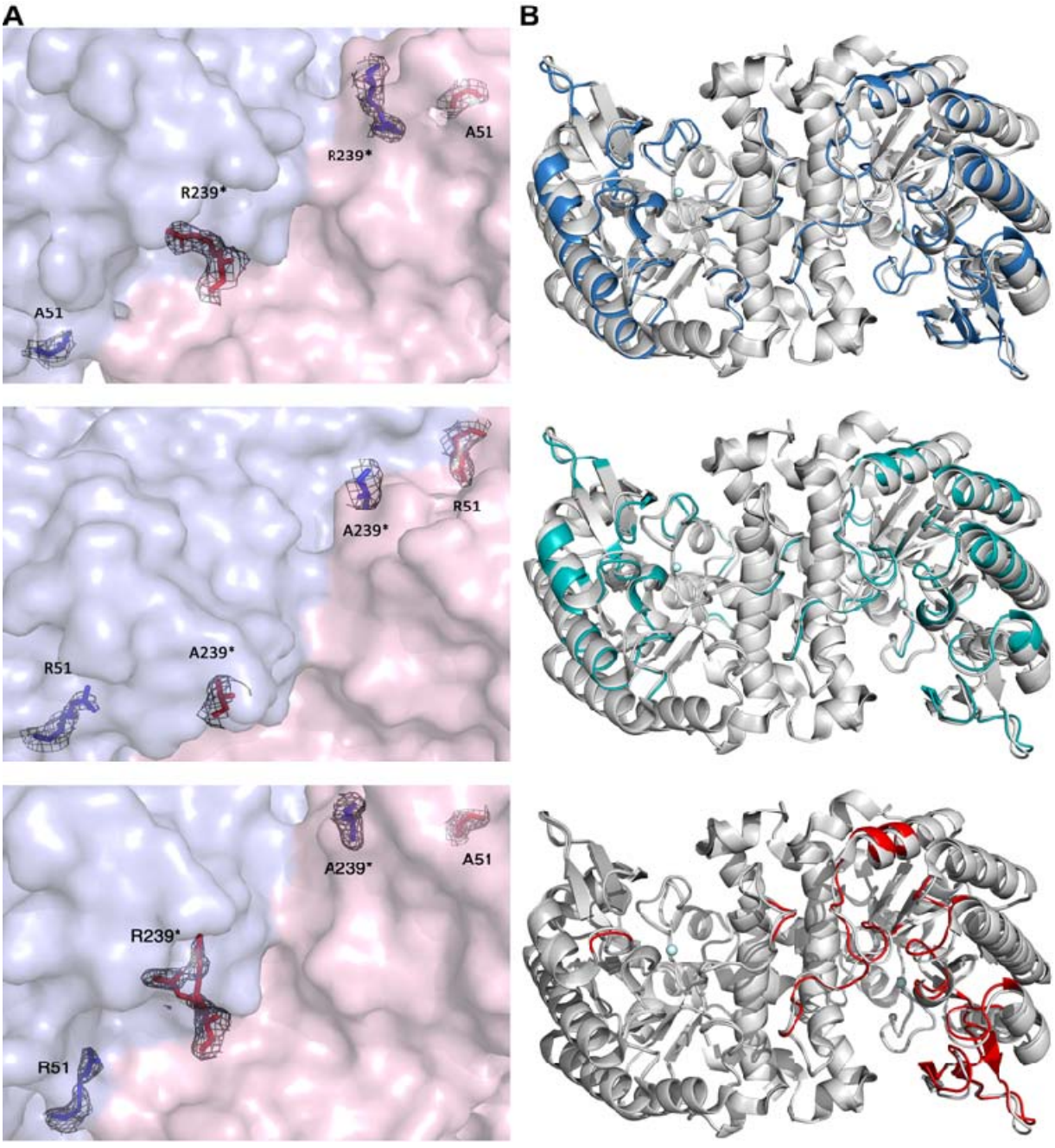


Figure 3.5. Active site and overall structures of R51A, R239A, and heterodimer. (A) Electron density of position 51 and 239 in the R51A homodimer (top panel), R239A homodimer (middle), and R51A-R239A heterodimer (bottom). (B) Superimposed overall structures of WT Zn-ACMSD with Zn-R51A (top) and Zn-R239A (middle), and heterodimer (bottom). The well-aligned structural components are shown in gray color, while structural changes for R51A and R239A are highlighted in blue and cyan colors, respectively.

Arg51 and Arg239 are Essential for Substrate Binding. Based on the current and previous studies, we propose that the two arginines are critical for proper substrate binding at the enzyme active site and for preventing its autocyclization to QA. Since the substrate, ACMS, is unstable and all arginine mutants are catalytically inactive, studying the postulated role of the two arginines in substrate binding is challenging. Therefore, an efficient competitive inhibitor needs to be found. Unfortunately, the inhibitors reported from previous *in vivo* studies, phthalate esters for instance (198,199), were not effective on the pure enzyme in our experiments. Acyclic ACMS analogue compounds tend to be unstable in solution, however, 2,6-pyridine-dicarboxylic acid (PDC), a cyclic compound, is structurally similar to ACMS (Figure 3.6A). Both compounds contain conjugate repeating single-bond double-bond structures and two carboxylate groups. The difference is that PDC is a stable compound due to its cyclic, aromatic structure, while ACMS tends to dehydratively autocyclize. Figure 3.6B shows that PDC behaves kinetically as a competitive inhibitor for the ACMS decarboxylation. The K_i value was determined to be 22.7 ± 1.8 μM. This allows for use of PDC as an analog of ACMS to study substrate binding.

The catalytic activity-based kinetic approach, unfortunately, cannot be employed to assess the binding of substrate or substrate analog compounds to the inactive mutant proteins. We took advantage of the fact that PDC has an electronic absorption band at 270 nm (ε270 nm = 4100 $M^{-1}cm^{-1}$, determined in 25 mM HEPES buffer, pH 7.0) and designed an experiment to test PDC's binding to the wild-type and mutant proteins of ACMSD utilizing separation by centrifuge filters. The threshold for membranes in centrifuge filters with a 10,000 Da cut-off is smaller than ACMSD but much larger than PDC. If the PDC is bound to ACMSD it will not be able to pass the filter membrane while unbound PDC molecules can pass freely. The amount of PDC in the flow-through was quantitated by UV absorbance. Wild-type ACMSD was employed

as a positive control while the same buffer containing no protein was used as one negative control. A previously well-characterized mutant H228G (25), was also employed as a secondary positive control. His228 is a strictly conserved residue sitting in the interior side of the metal center which is previously shown to function as a general base to deprotonate the water ligand during catalysis and is not related with substrate binding. The properties and crystal structure of H228G have been reported previously (25). Figure 3.6C shows that WT and H228G ACMSD proteins exhibit similar binding of PDC. In contrast, R51A and R239A are unable to effectively bind PDC as exhibited by the similar absorbance of their flowthroughs to the negative control. This experiment suggests that the two arginine residues play an important role in binding of the substrate analogue, which implies that they play a role in substrate binding in the native enzyme.

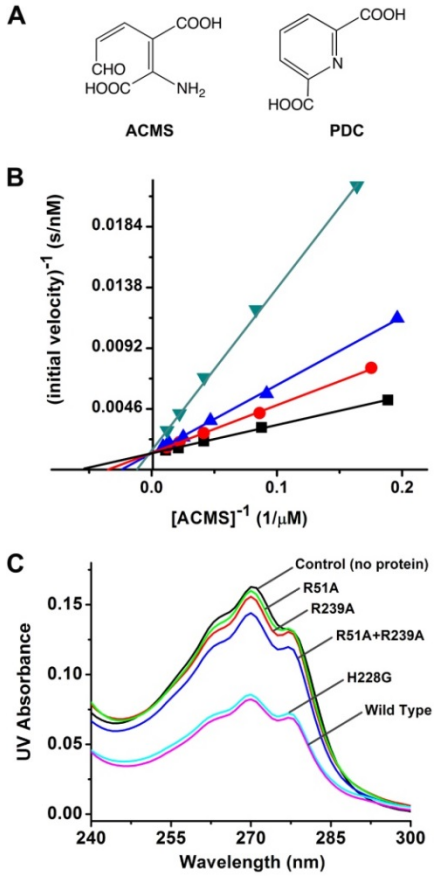


Figure 3.6. Arginine mutants cannot bind the competitive inhibitor PDC. (A) Chemical structure of ACMS and PDC. (B) Competitive inhibition of wild-type ACMSD by 0, 20, 40, and 80 μM of PDC. (C) PDC in flow through of centrifuge detected by UV-Vis spectroscopy.

3.5 Discussion

It is often a challenging task to determine whether or not an experimentally determined structure represents the biologically-relevant assembly. It is also often difficult to pin down the role of a critical residue in catalysis when its mutants are completely inactive. An additional complexity in the present case is that the two conserved arginine residues which are being studied are proposed to bind the substrate, ACMS, which is unstable, so no binding constant can be easily measured. We have crafted a series of biochemical tests to show that *P. fluorescens* ACMSD must function as a dimer, and the two conserved arginine residues, including one intruded into the catalytic center from a neighboring subunit, are required for substrate binding in catalysis.

The molecular mechanism by which the catalytic activity is rescued through hybridization is illustrated in Figure 3.4. During dimer–monomer equilibrium, three kinds of dimers form in solution when the two inactive homodimers of R51A and R239A are mixed: R51A homodimer, R239A homodimer, and R51A-R239A half-active heterodimer, in which one of the monomers has both Arg51 and Arg239 to constitute a complete active site identical to that of the native enzyme. If it is assumed that all homodimers are completely dissociated and rearranged, the percentage of formation of R51A homodimer and R239A homodimer will be both 1/4. Consequently, the rest 1/2 will be the half-active heterodimer. As a result, the maximum activity that can be rescued theoretically will be 25% of the wild-type ACMSD assuming there is no cooperative interaction and that each of the dimers has the same association constant. By comparing the k_{cat} values of the equilibrated mixed mutant ACMSD with WT ACMSD, 9.2% of the theoretical maximum activity is recovered. The inability to recover more activity can likely be attributed to unequal equilibrium constants for each of the species (inactive homodimer

formation may be thermodynamically preferred to half-active heterodimer formation), or some protein dynamics and cooperativity may have been lost due to the mutations of important active site residues.

To our knowledge, this X-ray crystallographic study provides the first unambiguous experimental evidence for the formation of hybridized active heterodimer by mixing of two inactive homodimers. These structures allow for a molecular level understanding of the correlation of enzymatic activity with the protein quaternary structure and the mechanism of activity rescue from two completely inactive mutants. Together with our previous work, we now can describe the functionally-active assembly of the enzyme and the role of the two arginines in a more definitive manner as discussed below.

The homodimer structures contain a ~2,500 $Å^2$ dimer interface area. Prior to this work, the question of whether dimerization is an artificial aggregation or required for enzyme activity was not known. We showed here two major bands corresponding to monomer and dimer on the polyacrylamide gel, indicating the existence of both dimer and monomer forms in aqueous solution. Enzyme specific activities were measured under identical assay conditions using protein at different stock concentrations. The finding, *i.e.* increasing specific activity with increasing protein stock concentrations, indicates that ACMSD is active when in the dimeric form and loses activity as it dissociates to the monomeric form.

Dimer formation as a mechanism for enzyme activation is common in nature. Cytochrome c peroxidases (200), cell death protease caspases (201,202), diaminopimelate epimerase (203), and histidine kinase CheA (204) are just a few examples similar to ACMSD. The reason behind this dimerization-induced activation is usually the generation or extension of binding faces at the dimer interface (205). Hence, protein activation is triggered by an increase in the local protein

concentration. As calculated by the dependence of specific activity on stock enzyme concentration, the dissociation constant of ACMSD is 0.3 μM. This is within the K_d range for "weak" transient complexes that show a dynamic mixture of both monomer and dimer states *in vivo* which could be tuned by physiological environmental factors such as pH and local protein or ligand concentrations (206).

The mutagenesis and reaction rescue experiments described in this study show that both Arg51 and Arg239 are required for enzyme activity. However, their role in catalysis was unknown due to the silent nature of their activity in all the point mutants, including the conservative Lys mutants, until the identification of an effective competitive inhibitor of the enzyme, PDC. By using this competitive inhibitor, the problems which arise from use of an unstable substrate can be mitigated. The inability of PDC to bind to the arginine mutants is demonstrated by the filtration experiment. With proper positive and negative controls, the designed experiment provides a straightforward and reliable conclusion. Hence, we propose both arginine residues R51 and R239* are involved in substrate binding during catalysis (Figure 3.7). It is noteworthy that this scenario resembles a well-documented example in a highaffinity nitrate transporter where two perfectly conserved arginine residues are required for substrate binding (207). Likewise, arginines play crucial roles in the active sites and subunit interfaces of the ATPase domains of AAA and AAA^+ proteins (208).

Heterodimer-induced activity restoration has also been previously observed, although not structurally characterized, in histidine protein kinase CheA, a protein essential for bacterial chemotaxis' stimulus-response coupling by phosphorylating another two chemotaxis proteins: CheY and CheB (209). CheA must be autophosphorylated at its own His48 residue before this phosphoryl group can be transferred to CheY and CheB (210,211). *Escherichia coli* expresses

two kinds of CheA proteins: the full length CheAL and the N terminally truncated CheAS, which lacks His48, the site of autophosphorylation (210). But in the presence of both CheAS and kinase defective CheAL mutant, CheA470GK, the mutant can still be phosphorylated. Thus, the autophosphorylation is accomplished within the CheAS-CheA470GK heterodimer is suggested even in the absence of structural proof as shown in the present study (204).

The results described in this work helped us further refine our working model for the ACMSD catalytic cycle. Previously, the roles of His228 and the metal bound water were proposed and demonstrated (25). The two catalytically essential arginine residues Arg51 and Arg239* are now added to the model (Figure 3.7). They are proposed to work as a team to recognize the substrate and stabilize both the substrate and possibly catalytic intermediates. Because the two residues at each metal center belong to two different subunits, the requirement of protein dimerization for enzyme activity becomes evident.

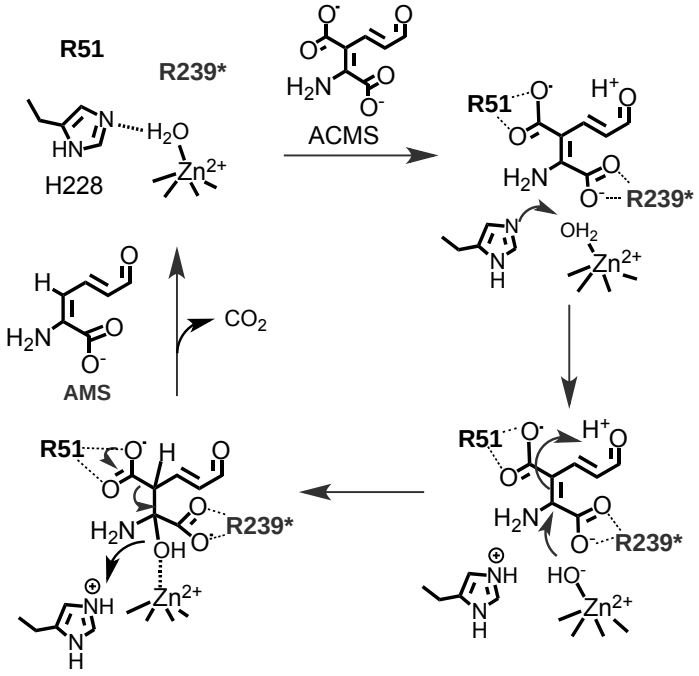


Figure 3.7. Proposed role of Arg51 and Arg239* in ACMSD.

Among cellular metabolic proteins, monomers are a small fraction compared to multimeric proteins. Homodimers and homotetramers are two major forms of oligomerized proteins (205). Compared to monomeric proteins, oligomerized proteins have some apparent advantages. Firstly, oligomerization enables higher-order protein regulation, such as allosteric and concentration sensing regulations. Secondly, protein oligomerization can help to increase protein stability and the flexibility of activity by subunitcooperation. Lastly, homo-oligomers are more economic because they allow for a smaller genome while still building large proteins. Hence, the oligmerization of ACMSD may function as a mechanism of protein concentration sensing and enzyme activity adjustment *in vivo*.

CHAPTER 4 HUMAN α-AMINO-β-CARBOXYMUCONATE-ε-SEMIALDEHYDE DECARBOXYLASE (ACMSD): A STRUCTURAL AND MECHANISTIC UNVEILING

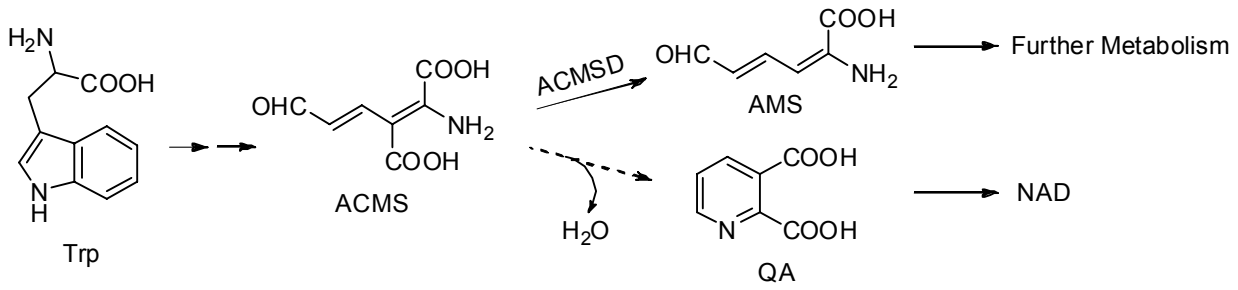
The section of chapter 4 is our manuscript under preparation: Human α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (ACMSD): A structural and mechanistic unveiling. Lu Huo, Fange Liu, Hiroaki Iwaki, Tingfeng Li, Yoshie Hasegawa, and Aimin Liu. We thank Dr. Shin-Ichi Fukuoka for providing cDNA clone of human ACMSD, Dr. Lirong Chen for stimulating discussions in the analyses of the structural data.

4.1 Abstract

Human α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase determines the fate of tryptophan metabolites in the kynurenine pathway by controlling the quinolinate levels for de novo NAD biosynthesis. The unstable nature of its substrate has made gaining insight into its reaction mechanism difficult. Our electron paramagnetic resonance (EPR) spectroscopic study on the catalytically active, Cu-substituted enzyme suggests that the native substrate does not directly ligate to the metal ion. Substrate binding did not result in a change of either the hyperfine structure or the super-hyperfine structure of the EPR spectrum. We also determined the crystal structure of the enzyme in its native state (at 1.99 Å resolution), a substrate analogue-bound form (2.50 Å resolution), and a selected active site mutant form with one of the putative substrate binding residues altered (2.32 Å resolution). These structures illustrate that each asymmetric unit contains 3 pairs of dimers. Consistent with the EPR findings, the ligand-bound complex structure shows that the substrate analogue does not directly coordinate to the metal ion but is bound the active site and stabilized by positively charges residues through non-covalent interactions.

4.2 Introduction

α-Amino-β-carboxymuconate-ε-semialdehyde (ACMS) is a metabolic intermediate found in two catabolic pathways: the kynurenine pathway (Scheme 4.1), which is responsible for tryptophan catabolism, and the 2-nitrobenzoic acid biodegradation pathway (14). ACMS is unstable and non-enzymatically dehydrates to quinolinic acid (QUIN) with a $t_{1/2}$ of 33 min at 20 °C, pH 7.0 (10). QUIN is the universal precursor for the de novo biosynthesis of NAD, providing the pyridine ring (193,212). In humans, QUIN levels must be tightly regulated (< 100 nM) because it is also an agonist of N-methyl-D-aspartate receptors, and its overproduction can cause overexcitement of neurons and cell death in the central nervous system (2,5,8,9,213). Elevated QUIN concentrations in bodily fluids have been observed in an exceptionally wide range of neuropsychiatric diseases including anxiety, depression, epilepsy, and neurodegenerative conditions such as Alzheimer's and Huntington's diseases (3-7,127,128). A 300-fold elevation of QUIN has also been detected in the brain of patients with HIV infection (214). Moreover, over the past two decades, an increasingly accumulating volume of evidence unanimously suggest that abnormal activity of the kynurenine pathway participates in the initial phases of neuropathological processes (7,129), the pathogenesis of AIDS-dementia complex (215,216), and Alzheimer's disease (2-4,217). When QUIN was injected into healthy rats, the characteristic neuropathological features of Huntington's disease started to develop (5). Thus, the connection between QUIN and related diseases flows both ways, *i.e*. in symptomatic patients QUIN levels are elevated, and in healthy mammals QUIN injection induces symptoms of disease. Notably, ACMS decarboxylase (ACMSD) directs the vast majority of metabolites to further catabolic steps controlled by enzymes and, by doing so, avoids the overproduction of QUIN (Scheme 4.1) (14,193). Thus, the decarboxylase ACMSD is of significant biomedical importance.



Scheme 4.1. ACMSD in the Kynurenine pathway

Prior to this work, the structure of human ACMSD (*h*ACMSD) in its catalytically active form was unavailable. The elucidation of this structural information is crucial for the design of small molecule regulators and for the interpretation of biochemical and spectroscopic studies. We are also interested in the structural and mechanistic divergence between *h*ACMSD from the tryptophan kynurenine pathway and *Pseudomonas fluorescens* (*Pf*ACMSD) from the 2-nitrobenzoic acid metabolic pathway. However, *h*ACMSD has proven to be a considerable challenge to express in prokaryotic systems. Although expression of *h*ACMSD as a weakly active enzyme has been achieved in mammalian cell lines (13) and the eukaryote, *Pichia pastoris* (194), prior attempts to express *h*ACMSD in bacterial systems were all unsuccessful (13,131,194).

In the present work, we successfully expressed *h*ACMSD in *E. coli* and obtained catalytically active form of the enzyme. By using substrate analogue and metal-substituted *h*ACMSD, we employed a combined structural and spectroscopic approach to tackle the question of how the unstable substrate interacts with the enzyme's active site.

4.3 Experimental Procedures

Expression of Human ACMSD. To construct an Nterminally His10-tagged *h*ACMSD expression plasmid, the *h*ACMSD cDNA (13), a generous gift from Dr. Fukuoka, was amplified

by the polymerase chain reaction with the forward primer 5'-GGAATTCCATATGAAAATTGACATCCATAG-3' and the reverse primer 5'-CCGCTCGAGTCATTCAAATTGTTTTCTCTC-3' (built-in NdeI and XhoI sites are underlined). The PCR product was purified from a 0.8% agarose gel, digested with NdeI and XhoI, and ligated in the equivalent sites of pET16b (Novagen). After sequencing, the positive clone was used for transformation into E. coli BL21 (DE3) containing the GroEL/ES overexpression plasmid pGro7, pColdII-pGro7, pColdII-pG-Tf2, and pET16b-pG-Tf2, respectively (Takara Corp., Japan).

Cell Culture. Single colonies of *E. coli* containing the appropriate pET-*h*ACMSD plasmid were used to inoculate 50 mL of M9 growth medium (1×M9 medium supplement with 1 mM $MgSO_4$, 0.1 mM $CaCl_2$, 2 μg/mL vitamin B1, 0.04% casamino acids, 0.4% D-glucose, and a divalent metal ion described below). To obtain *h*ACMSD with desired metal at the active site, a divalent metal ion of either $CoCl_2$, CuSO4, $Fe(NH_4)_2(SO_4)_2$, or $ZnCl_2$ was added to a final concentration of 0.05 to 0.5 mM prior to induce expression. The cell culture medium also contained 100 μg/mL ampicillin, 40 μg/mL chloramphenicol and 500 μg/mL L-arabinose, the latter of which induces the expression of the protein chaperone GroES-GroEL. The cells were grown at 37 °C with shaking until the OD600 reached 0.4, at which time the temperature was lowered to 25 °C. *h*ACMSD expression was induced by 0.1 mM isopropyl β Dthiogalactopyranoside at an OD600 of 0.6. The cultures were allowed to further grow at 25 °C for 16 h before being harvested by centrifugation at 8000 × *g* for 15 min at 4 °C. The wet cells were washed with 50 mM potassium phosphate buffer, pH 8.0, twice before storage at -80 °C.

Protein Isolation. Human ACMSD was purified using a modified procedure described for the bacterial analogue (25,26). Frozen cells were resuspended in 50 mM potassium phosphate, pH 8.0, containing 300 mM NaCl and 0.1 mM protease inhibitor phenylmethylsulfonyl fluoride. The cell slurry was passed through an M-110P Microfluidics cell disruptor and the debris was removed by centrifugation at 27,000 × g for 30 min at 4 °C. The supernatant was loaded onto a HiPrep immobilized metal ion affinity chromatography (Ni-NTA) 26/20 column (80 mL resin) for isolation of the native enzyme or FF 16/10 columns charged with either Co^{2+}, Cu^{2+} or Zn^{2+} depending on the corresponding metalspecific *h*ACMSD preparations in order to minimize possible metal cross-contamination in purified protein. The proteins were eluted in a two-step gradient with increased concentration of imidazole in the buffer. The major fractions with ACMSD decarboxylase activity were pooled, concentrated, and further purified on a Superdex 75 column (26/60) using an ÄKTA FPLC protein purification system with a 25 mM HEPES buffer, pH 7.0, containing 5% glycerol. Protein concentrations were determined using Coomassie Plus protein assay reagent (Pierce). The expression level and enzyme purity were determined by SDS-PAGE on 12% polyacrylamide gels.

Metal Analysis. The metal content of ACMSD was assayed in triplicate by inductively coupled plasma optical emission spectrometry (ICP-OES) using a Spectro Genesis spectrometer (Spectro Analytical Instruments GmbH & Co. KG, Germany) as previously described (25). Metal ions loosely or nonspecifically bound to the enzyme were removed by an overnight incubation of 0.5 mM EDTA with the protein solution at 4 °C followed by desalting on a 5-mL HiTrap column with 10 mM Tris-HCl (pH 7.4) and buffer wash with the ICP buffer and ultrafiltration.

EPR Spectroscopy. The interaction between ACMS and the enzyme-bound metal center was measured by EPR spectroscopy using Cucontaining ACMSD in the absence or presence of 3 equivalents of ACMS in the assay buffer. The EPR samples were obtained by mixing the protein solution with the substrate solution at 25 °C using a rapid-freeze-quenching apparatus, Update Instruments, system 1000. The reaction mixture was shot directly to an EPR tube and frozen in liquid nitrogen. The entire mixing and freezing process was completed in less than 5 seconds. In a parallel EPR experiment, the response of the addition of 3 equivalents of ACMS to copper sulfate was determined with the same buffer system. The same X-band EPR instrumentation is used as those previously described (14,15).

Steady State Kinetics. The enzyme activity of the recombinant *h*ACMSD was determined according to the procedures described in detail in earlier reports (14,15,25). Briefly, ACMS was generated from 3-hydroxyanthranilate by an enzymatic method (14). The ACMSD enzyme assay mixture contained 0 – 120 μM ACMS and an appropriate amount of ACMSD protein in 25 mM HEPES buffer, pH 7.0 with 5% glycerol. Specific activities were calculated from the initial velocities of ACMS decay monitored by the loss of absorption at 360 nm using an absorbance coefficient constant of 47,500 $M^{-1}cm^{-1}$ when the substrate concentration was equal to or lower than 20 μM, or at 320 nm using an absorbance coefficient constant of 9,600 $M^{-1}cm^{-1}$ when the substrate concentration was between 20 – 120 μM. In the enzyme concentration dependence assay, 800 μM Zn-*h*ACMSD was diluted to 200 nM by the reaction buffer. The initial rate of the reaction was monitored from 1 to 240 min frequently in triplet after dilution: a saturated substrate concentration of 15 μM was used. The inhibition pattern of pyridine-2,6-dicarboxylic acid (PDC) acting on *h*ACMSD was obtained by determining the k_{cat} and K_m in the presence of varying concentrations of PDC (0, 20, 40, and 80 μM) in both buffer and substrate solution.

Apparent *K*m values for each inhibitor concentration were plotted as a function of inhibitor concentration and inhibition is expressed as K_i values in micromolar (K_i = y-intercept/slope).

X-ray Protein Crystallography. Catalytically active form of Zn-*h*ACMSD was used to screen crystallization conditions in Art Robbins 96-well Intelli-Plates using an ARI Gryphon crystallization robot. The initial hit was obtained from the Index Screening Kit (Hampton Research). After optimization, crystals were obtained from drops assembled with 1.5 μL of 28 mg/mL *h*ACMSD mixed with 1.5 μL of a reservoir solution containing 0.2 M lithium sulfate monohydrate, 0.1 M Tris-HCl pH 6.5 - 7.5, and 20 - 25% polyethylene glycol 3,350, by hanging drop diffusion in VDX plates (Hampton Research). PDC-*h*ACMSD co-crystals were obtained by preincubating *h*ACMSD with 10 equivalents of PDC for 30 min prior to crystallization. Crystals suitable for X-ray diffraction were obtained ca. 10 days after crystal growth at 18 °C. Crystallization mother liquor containing 20% glycerol was used as a cryoprotectant. All crystals obtained under these conditions belong to the $P2_12_12_1$ space group and data were processed with HKL-2000 (137). The three dimensional structures of *h*ACMSD were solved by molecular replacement using our previously published *Pseudomonas fluorescens* (*Pf*ACMSD) structure as a search model (PDB code: 2HBV) using PHENIX (196). Electron density analysis and model building were carried out in Coot (141). The unit cell images were generated with PyMol (218) with the aid of the SuperSym plugin available at http://supersym.sourceforge.net.

4.4 Results

*Overexpression of the Human Enzyme in E. coli. h*ACMSD expressed from the constructed vectors, pColdII-pGro7, pColdII-pG-Tf2, and pET16bpG-Tf2 was only found in inclusion bodies.Fortunately, the constructed pET16b-pGro7 plasmid expressed *h*ACMSD as a soluble and

enzymatically active protein. The enzyme expressed from this system was purified to near homogeneity as described under "Experimental Procedures". Five to ten mg of pure enzyme was obtained from each liter of cell culture when gown in the LB medium; however the yield was significantly reduced when a minimal medium was used. Purified *h*ACMSD is stable at 4 °C for weeks and can endure freeze/thaw cycles with liquid nitrogen multiple times without a significant loss of activity.

Metal Content and Catalytic Activity. Metal analysis of *h*ACMSD purified from cells grown in the LB medium contained several metals. Iron, zinc and copper were present and the total metal content was substoichiometric. Metal analysis of *h*ACMSD purified from the cells cultured in the M9 minimal media supplemented with 100 μM of varying metal ions is presented in Table 4.1. Activity assays demonstrated that *h*ACMSD expressed in the presence of Zn^{2+} exhibited activity comparable to what was found for the most active forms of *Pf*ACMSD bound with a Zn^{2+} or Co^{2+} ion (14). Co^{2+}-containing *h*ACMSD, on the other hand, exhibited very little specific enzymatic activity relative to Co-*Pf*ACMSD. It should be noted that the contribution of Zn ion to the activity of Co-, Cu-, and Fe-*h*ACMSD was subtracted based on Table 4.1 metal analysis data with regard to the Zn-*h*ACMSD activity.

Table 4.1. Percentage of metal occupation and corresponding activity of *h*ACMSD from different medium cultures.

	LB	M9	M9 + Co^{2+}	M9 + Cu^{2+}	M9 + Fe^{2+}	M9 + Zn^{2+}
Co	0	0	96.8	0.4	0.8	0
Cu	2.2	2.9	0.7	47.9	3.6	9.7
Fe	28.7	9.8	1.0	10.0	55.1	7.5
Zn	13.8	9.8	0.8	0.6	7.7	64.4
Specific activity (nmol/min/mg)	910	800	190	90	400	3540

Unlike *Pf*ACMSD (14), the metal ions in *h*ACMSD could not be removed by the metal chelators EDTA, 1,10-phenanthroline or 8-hydroxy-quinoline-5-sulfonic acid, and the enzyme activity was not affected by the addition of these metal chelators. Moreover, the addition of one to ten equivalents of Co^{2+}, Cu^{2+}, Fe^{2+}, or Zn^{2+} to purified *h*ACMSD did not increase enzyme activity, indicating that the purified human enzyme cannot be reconstituted *in vitro*. Since *in vitro* metal reconstitution was not successful for the human enzyme, we attempted to add zinc ions to the M9 medium at various concentrations prior to induction of the cells, so that a variety of zinc/*h*ACMSD ratios could be obtained. Although we had no control over the final zinc content in the purified protein using this method, we did successfully purify the human enzyme with different amounts of zinc ion. Figure 4.1 shows that the specific activity of *h*ACMSD is proportional to the zinc content of the enzyme.

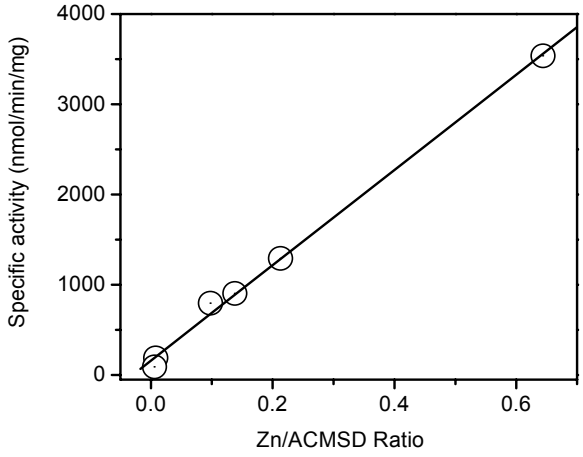


Figure 4.1. The specific catalytic activity of *h*ACMSD as a function of Zn content in the protein reveals that *h*ACMSD is a zinc-dependent enzyme.

Kinetic Properties. Zn-*h*ACMSD yielded steady-state kinetic parameters similar to those of Zn^{2+}-containing *Pf*ACMSD (31). A k_{cat} value of 4.8 s^{-1} was measured for *h*ACMSD with 64.4%

zinc ion occupancy. According to Figure 4.1, this value is anticipated to be 7.5 s^{-1} with a fullyloaded zinc ion. The K_m (5.8 μM) and the k_{cat} (7.5 s^{-1}) values of *h*ACMSD are comparable to the K_m (9.6 μM) and the k_{cat} (6.5 s^{-1}) reported for Zn-*Pf*ACMSD with *ca.* 80% zinc occupancy (80). When *h*ACMSD was diluted from 800 μM to 200 nM, it lost 90% of activity after 4 h (Figure 4.2). Since proteins prefer to form higher oligomer state at higher concentration, the *h*ACMSD dimer will start to dissociate once its concentration drops. Loosing catalytic activity after dilution implies that *h*ACMSD is unlikely to be active in the monomeric state. Similar observations have been reported for the bacterial ACMSD (26).

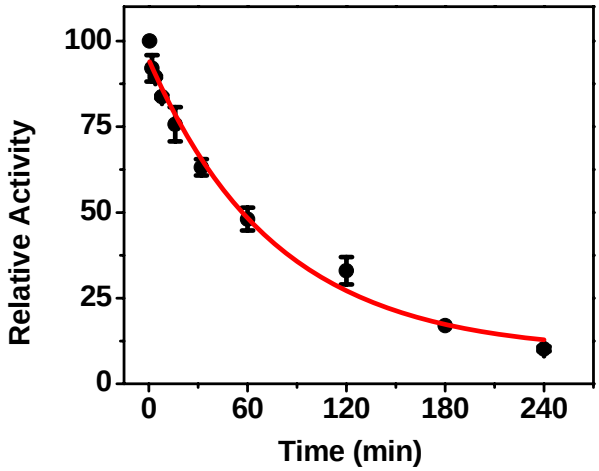


Figure 4.2. The catalytic activity of *h*ACMSD decays after dilution, indicating that *h*ACMSD has better activity at higher oligomerization sate. The relative activity of concentrated and diluted samples was measured under the same conditions as described in the text. The experimental data was fitted with first exponential decay (red line).

Probing the Nature of the Enzyme-Substrate Interactions. Although ACMS is an unstable molecule, we find that it tends to form stable complexes with various metal ions. Figure 4.3A shows that ACMS binds to a cupric ion in solution. The EPR signal intensity of $CuSO_4$ solution is weak due to the coupling of copper ions. Addition of unstable ACMS to $CuSO_4$ led to the formation of stable Cu^{2+}-ACMS complex which gives rise to well-resolved type II copper EPR

spectrum (219). Since this technique is highly sensitive to the changes of the electronic/chemical structure of a Cu^{2+} ion, especially when the superhyperfine structure is resolved, we analyzed ACMS perturbation to the enzyme-bound copper center by EPR spectroscopy.

To probe whether the substrate ACMS is ligated directly to the metal, we obtained Cu-substituted *h*ACMSD. Since the metal ion cannot be extracted and reconstituted, we circumvented this problem by isolating the human enzyme from cell culture of a metal-depleted minimal medium supplemented with cupric ion. The Cu-substituted protein is catalytically active although it is a relatively poor catalyst. Its k_{cat} value was 0.1 ± 0.02 s^{-1} and K_m value was 3.3 ± 0.2 μM.

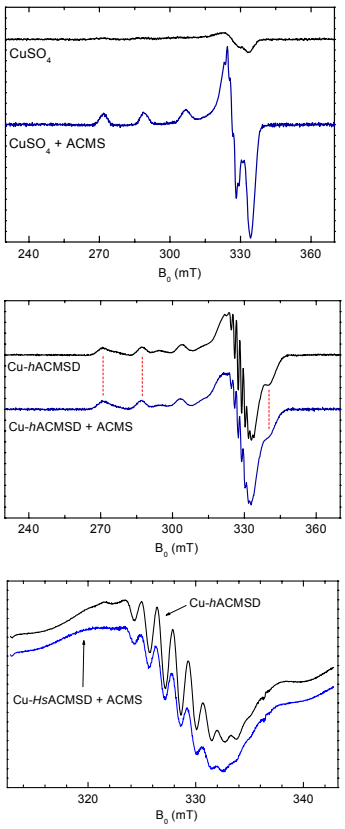


Figure 4.3. Low-temperature X-band EPR spectroscopic study confirms that the substrate binds to the free metal ion but does not directly coordinate to the catalytic metal ion in ACMSD. (A) Free $CuSO_4$ and $CuSO_4$-ACMS, (B) as-isolated Cu-*h*ACMSD and Cu-*h*ACMSD-ACMS, and (C) a magnified view of superhyperfine structure of Cu-*h*ACMSD in the absence/presence of ACMS. The spectrometer conditions: temperature, 20 K; modulation amplitude, 0.3 mT; microwave power, 0.5 mW; time constant, 40.96 ms; and sweep time 1.79 mT/s for the field from 225 to 375 mT.

Both hyperfine splitting due to the nuclear spin of copper ($I = 3/2$) and nitrogen ($I = 1$)-induced superhyperfine interactions are present in the EPR spectrum of Cu-*h*ACMSD (Figure 4.3B). The EPR parameters obtained from a simulation of the spectrum of wild-type Cu-*h*ACMSD have an $A//$ value of 18.7 mT and a $g//$ value of 2.218, both typical for a type II Cu^{2+} center. In the perpendicular region, a multiple-line superhyperfine structure with a splitting of about 1.5 mT is well resolved in the spectrum (Figure 4.3C). The superhyperfine structure originates from the spin-spin interaction between the nuclear spin of the nitrogen atoms (nuclear spin $I = 1$) and the electron spin of the Cu^{2+} ion. When 3 molar equivalents of ACMS were added to Cu-*h*ACMSD, frozen in less than 5 s, and subsequently measured by EPR at 20 K, no spectral change was observed. In sharp contrast to the free copper ion in solution, the presence of ACMS is unable to induce any disturbance in the hyperfine and superhypefine structures of the Cu center in *h*ACMSD. Thus, there is no direct ligation between ACMS and the Cu center.

Crystal Structure. The structure of *h*ACMSD was obtained from native, catalytically competent form of the enzyme and was refined to 1.99 Å resolution (Table 1, PDB entry: 4OFC). Each unit cell contains four asymmetric units in which six protomers are found as a trimer of dimers, Figure 4.4A. The dimer interface contains a surface area of about 2,400–2,460 $Å^2$ and is primarily formed by three helical regions which are composed of residues 191 to 201, 230 to 244, and 271 to 281 of each subunit. The central area of the dimer interface is strictly hydrophobic. Analysis of the dimer interface reveals a total of 24-33 direct hydrogen bonding interactions and 18-21 salt bridges for each dimer. The detailed interactions are listed in Table 4.3.

Due to the instability of ACMS, the substratebound ACMSD structure is not yet available. However, pyridine-2,6-dicarboxylic acid is a heterocyclic, stable ACMS analogue that

competitively inhibits the bacteria ACMSD (26). We found that it is also an effective competitive inhibitor of *h*ACMSD (Figure 4.5A) with a K_i value of 15.2 ± 0.5 μM. We successfully co-crystallized this compound with *h*ACMSD (Table 1, PDB entry: 4IH3) and the overall structure is nearly identical to the ligand-free enzyme. PDC is bound at the active site and forms ionic interactions with Arg47 and Arg235*, where the asterisk indicates this residue is from a neighboring subunit (Figure 4.5B). The PDC-bound structure shows that the inhibitor is fairly close to the zinc ion with the C4 *ca*. 2.6 Å from the zinc ion. However, the closest oxygen is 5.6 Å away from the metal ion. Thus, the substrate analogue does not directly coordinate to the metal ion.

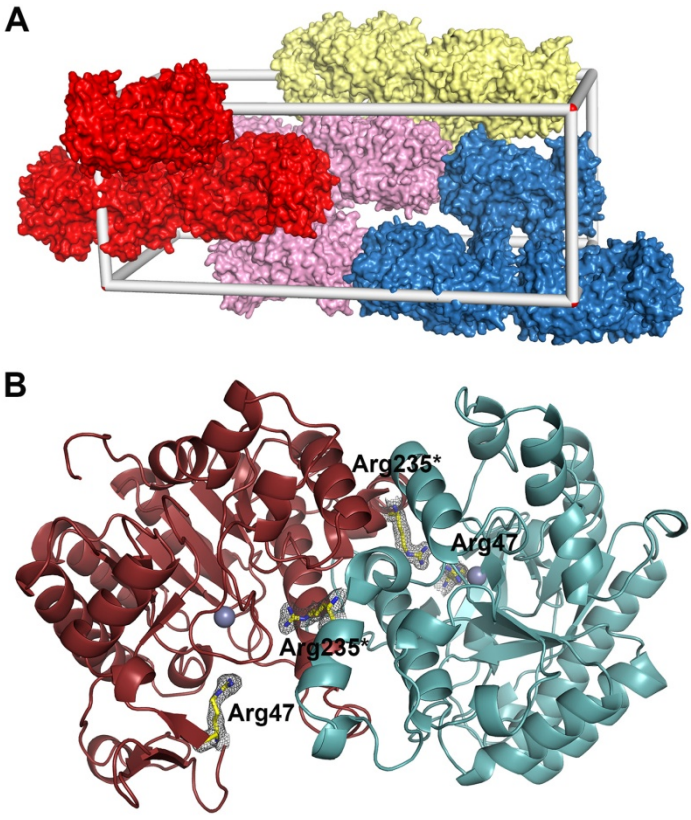


Figure 4.4. The crystal structure of *h*ACMSD is solved as a homodimer. (A) Each unit cell contains 4 asymmetric units of 3 pairs of dimers. (B) The catalytic center contains two arginine residues, Arg47 and Arg235* from a neighboring subunit.

Table 4.2. X-ray crystallography data collection and refinement statistics.

Data collection	*h*ACMSD	PDC-*h*ACMSD	R47A *h*ACMSD
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell lengths (Å)	a=88.704, b=101.082, c=232.820	a=88.368, b=101.655, c=233.117	a = 89.142, b=101.686, c = 232.613
Unit cell angles (˚)	α=γ= β=90	α=γ= β=90	α = β = γ = 90
Wavelength (Å)	1.0	1.0	0.8
Temperature (K)	100	100	100
Resolution (Å)[a]	45.00-1.99 (2.03-1.99)	45.00-2.50 (2.50-2.54)	50.00-2.33 (2.37-2.33)
Completeness (%)[a]	87.6 (81.3)	75.4 (65.7)	98.4 (82.8)
R_{merge} (%)[a, b]	15.3 (78.2)	9.9 (29.3)	13.2 (48.6)
$I/\sigma I$ [a]	20.6 (2.3)	10.3 (1.5)	17.9 (2.0)
Multiplicity [a]	10.1 (8.0)	9.7 (8.5)	25.3 (13.7)
Refinement			
Resolution (Å)	1.99	2.50	2.32
No. reflections; working/test	130135/6497	55192/2809	85220/4492
R_{work} (%)[c]	19.88	20.20	20.8
R_{free} (%)[d]	24.07	28.28	27.6
No. of protein atoms	15984	15810	15774
No. of ligand atoms	6	78	6
No. of solvent sites	994	286	821
Average B factor (Å^2)			
Protein	29.7	33.3	30.0
Zn(II) or PDC	26.3	39.4	23.6
Solvent	31.3	28.4	32.2
Ramachandran statistics[e]			
Preferred (%)	96.15	91.82	97.7
Allowed (%)	2.60	5.86	2.3
Root mean square deviation			
Bond lengths (Å)	0.008	0.029	0.015
Bond angles (˚)	1.155	1.262	1.785
PDB entry	4OFC	4IH3	4IGN

[a] Values in parentheses are for the highest resolution shell.
[b] $R_{merge} = \Sigma_i |I_{hkl,i} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \Sigma_i I_{hkl,i}$, where $I_{hkl,i}$ is the observed intensity and $\langle I_{hkl} \rangle$ is the average intensity of multiple measurements.
[c] $R_{work} = \Sigma||F_o|-|F_c|| / \Sigma|F_o|$, where $|F_o|$ is the observed structure factor amplitude, and $|F_c|$ is the calculated structure factor amplitude.
[d] R_{free} is the R factor based on 5% of the data excluded from refinement.
[e] Based on values attained from refinement validation options in COOT.

Table 4.3. Hydrogen bonds stabilizing the dimeric surface in the 2.0 Å resolution crystal structure of human ACMSD (4OFC.pdb) calculated from PDBePISA.

H-Bond #	Subunit B	Distance (Å)	Subunit C
1	ASP 151[OD1]	3.01	ARG 183[NE]
2	ASP 151[OD1]	2.89	ARG 183[NH2]
3	ASP 151[OD2]	2.45	LYS 186[NZ]
4	ASP 151[OD2]	3.06	ARG 183[NE]
5	TYR 187[O]	3.64	CYS 247[SG]
6	TYR 187[OH]	3.12	ASN 153[ND2]
7	TRP 191[O]	2.8	ARG 235[NE]
8	TRP 191[O]	3.02	ARG 235[NH2]
9	MET 195[SD]	3.29	ARG 235[NH2]
10	GLU 198[OE2]	2.88	ARG 183[NH1]
11	GLU 198[OE2]	2.95	ARG 183[NH2]
12	MET 208[O]	3.52	TYR 187[OH]
13	PHE 231[O]	2.62	SER 273[OG]
14	ASP 250[OD2]	2.5	TYR 187[OH]
15	ASP 270[OD2]	2.64	LYS 256[NZ]
16	PRO 295[O]	2.94	HIS 238[NE2]
17	GLY 297[O]	2.96	HIS 238[NE2]
18	ASN 153[ND2]	3.34	TYR 187[OH]
19	ARG 183[N]	3.9	GLU 149[OE1]
20	ARG 183[NE]	2.99	ASP 151[OD2]
21	ARG 183[NH1]	2.83	GLU 198[OE2]
22	ARG 183[NH2]	3.36	GLU 198[OE1]
23	ARG 183[NH2]	2.78	ASP 151[OD1]
24	LYS 186[NZ]	2.75	ASP 151[OD2]
25	TYR 187[OH]	2.67	ASP 250[OD2]
26	ARG 235[NE]	2.83	TRP 191[O]
27	ARG 235[NH2]	2.96	TRP 191[O]
28	ARG 235[NH2]	3.51	MET 195[SD]
29	HIS 238[NE2]	2.94	PRO 295[O]
30	HIS 238[NE2]	3.01	GLY 297[O]
31	CYS 247[SG]	3.68	TYR 187[O]
32	LYS 256[NZ]	2.82	ASP 270[OD1]
33	SER 273[OG]	2.71	PHE 231[O]

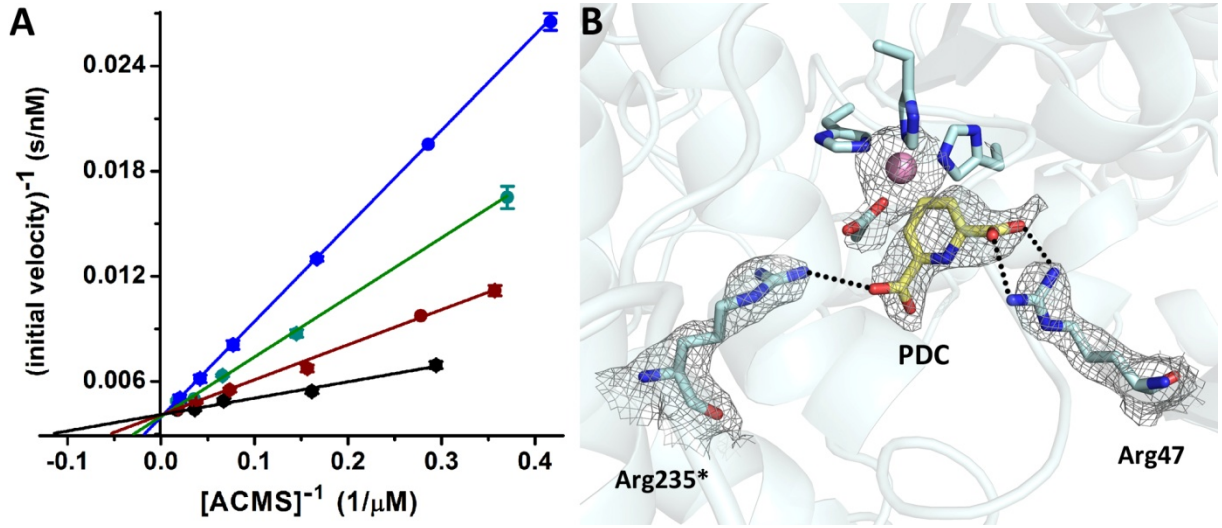


Figure 4.5. The crystal structure of *h*ACMSD in complex with a substrate analogue sheds light on the enzyme-substrate interaction mode. (A) Competitive inhibition of *h*ACMSD by 0, 20, 40, and 80 μM of PDC. (B) The co-crystalized structure shows that the competitive inhibitor PDC is bound by Arg47 and Arg235* near the zinc ion.

To further prove that the unstable substrate binds to the enzyme by the active site arginines, we mutated Arg47 and Arg235* to alanine. Both R47A and R235A variants showed no detectable catalytic activity, proving that they are important residues for catalysis. The result of this analysis is consistent with our structural findings shown in Figure 4.4B. While R235A eludes crystallization, the crystal structure of the R47A mutant was solved at 2.32 Å resolution (Table 4.2). Similar to the wild-type enzyme structure, six protomers in three dimers are present in each asymmetric unit (Figure 4.6). The same interface region is observed between the protomers as observed in the wild-type enzyme. The RMSD value is only 0.254 Å from 14,034 out of 14,040 atoms for the structures of R47A and wild type.

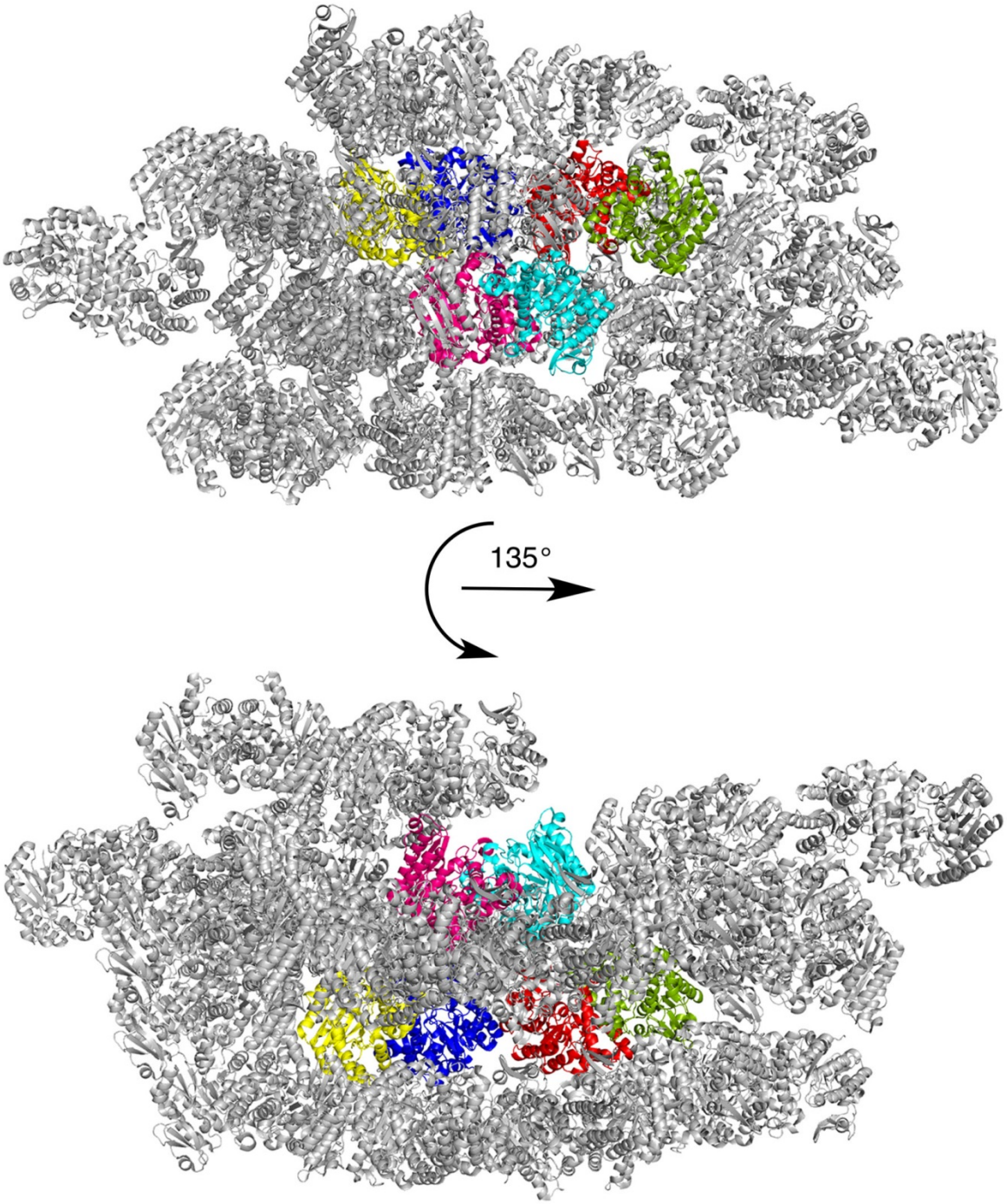


Figure 4.6. Crystal structure of R47A *h*ACMSD. A symmetric unit of the structure (shown in color) contains three pairs of dimers.

4.5 Discussion

Our work on human ACMSD solved several longstanding mysteries. The first pertains to the cofactor identity. Our previous work on the analogous bacterial system indicates that *Pf*ACMSD is a metal-dependent enzyme (14,15,80). Thus, one would expect *h*ACMSD is a metalloprotein. However, all prior studies on *h*ACMSD suggest that the catalytic activity of human enzyme is independent of metal ions (13,194). This is likely due to the observation that neither a metal chelator, nor external metal affects the activity of purified *h*ACMSD. The first crystal structure of *h*ACMSD was later determined with a zinc ion bound at the enzyme active site, but how the zinc ion is associated with catalytic activity was not discussed (131), leaving a questionable cofactor-free conclusion for *h*ACMSD in literature. Here, by showing protein expressed in a minimal medium supplemented with increasing concentrations of zinc ions, giving higher decarboxylase activity together with the data present in Table 4.1, we found that *h*ACMSD has the best affinity for zinc ion and the zinc enzyme gives rise to the highest catalytic activity. The catalytic efficiency of Zn-*h*ACMSD is significantly higher than that previously reported for the same enzyme expressed in *Pichia pastoris*, which has a similar K_m but a lower k_{cat} value of 1.0 s^{-1} (194). Thus, it becomes evident that the *h*ACMSD is a zinc-dependent enzyme.

Since chelators are demonstrably to be able to extract the metal ion from *Pf*ACMSD but not the human enzyme, *h*ACMSD apparently binds its metal more tightly. The crystal structures solved provide a molecular understanding for this nonconformity. An aspartate and three histidine residues in *h*ACMSD coordinate the zinc ion. Notably, His174 is a metal ligand in all six protomers of each asymmetric unit of the structure, with a bond distance of 2.2-2.3 Å. In contrast, the equivalent histidine in *Pf*ACMSD, His177, presents significant flexibility in the structure we have solved (PDB entry: 2HBV). In subunit A, His177 coordinate to the zinc ion

with distance of 2.2 Å. In subunit B, His177 is no longer a metal ligand and is 4.2 Å away from the zinc ion. Thus, His177 is a weaker ligand in *Pf*ACMSD. The observed metal affinity deviation therefore becomes understood.

It is challenging to predict the cofactor dependency of a protein base on the primary structure if the cofactor is a single atomic metal. Although characteristic sequences can be used to identify some metal binding motifs, including zinc finger, calcium hand, and iron-sulfur cluster, drawing conclusions based solely on sequence can be misleading. For example, LigI is a member of the amidohydrolase superfamily, whose members were all previously demonstrated to be divalent metal dependent including ACMSD. Based on sequence alignment, LigI has all four conserved residues, three histidines and one aspartate, for metal coordination. However, later mechanistic and structural studies revealed that LigI is actually metal independent and the four metal ligand residues are with new catalytic roles (30). Sometimes, even with the assistance of biochemical study, prediction of metal cofactor can still be misjudged. The enzyme urinate isomerase catalyzes the isomerization reaction of D-glucuronate and D-fructuronate was initially thought to be metal independent because metal chelators do not affect the enzyme activity (212). Later on, an improved study suggested that uronate isomerase is actually a mononuclear zinccontaining enzyme (220). Hence, a combined knowledge of sequence, biochemical assay, and structure are needed to correctly identify a metal cofactor. In the present case, it is fortunate that *h*ACMSD holds its metal tightly, so its preference for zinc ion is visible in isolated protein in our hand although this property also caused misjudgment in some of the earlier studies for the metal cofactor-dependency (13,194).

Since its initial identification from rat liver, ACMSD has been known for 58 years (130) and this enzyme has arisen significant biomedical interest. Following our success in obtaining the

bacterial enzyme structure in 2006 (80), a structure of the enzyme in its inactive form, complexed with 1,3-dihydroxyacetone-phosphate (DHAP), was obtained through molecular replacement in 2009 by Garavaglia et al. (PDB entry 2WM1) (131). It should be noted that the DHAP-bound structure is monomeric with only one subunit in each asymmetry unit, whereas the three *h*ACMSD structures present in this study, including a substrate-free and ligand-bound structure, and all of the *Pf*ACMSD structures determined thus far (25,26,80), are homodimers. Although all these coordinates are highly valuable data for future structure-based design of bioactive small molecule targeting of *h*ACMSD, a quaternary structure variety is apparently present. We think that the diversity brings up the following questions: (1) is the quaternary structure of ACMSD linked to the catalytic activity, and (2) is a ligand-induced structural change a potential mechanism of catalytic activity regulation? While the latter question will be a subject for future study, the results presented in this work and the comparison with our previous studies on the bacterial enzyme shed insightful clues, as discussed below.

We have recently shown that *Pf*ACMSD is a mixture of monomer and dimer in solution, and that the former is catalytically inactive while the latter is functionally competent (26). The decarboxylase activity is dependent on the presence of two arginine residues (Arg51 and Arg239*, star indicates that Arg239 is from the neighboring subunit), which play a key role in substrate binding (26). Likewise, when *h*ACMSD is in the dimeric form, the corresponding arginine residues (Arg47 and Arg235*) are both present in the active site in a manner resembling what has been described for *Pf*ACMSD (Figure 4.7). When *h*ACMSD is in the monomeric form bound with the DHAP, Arg235 is 25 Å from the catalytic metal and cannot be a residue involved in substrate binding (131). Thus, these conserved arginine residues are effective probes for determining the functional assembly of ACMSD.

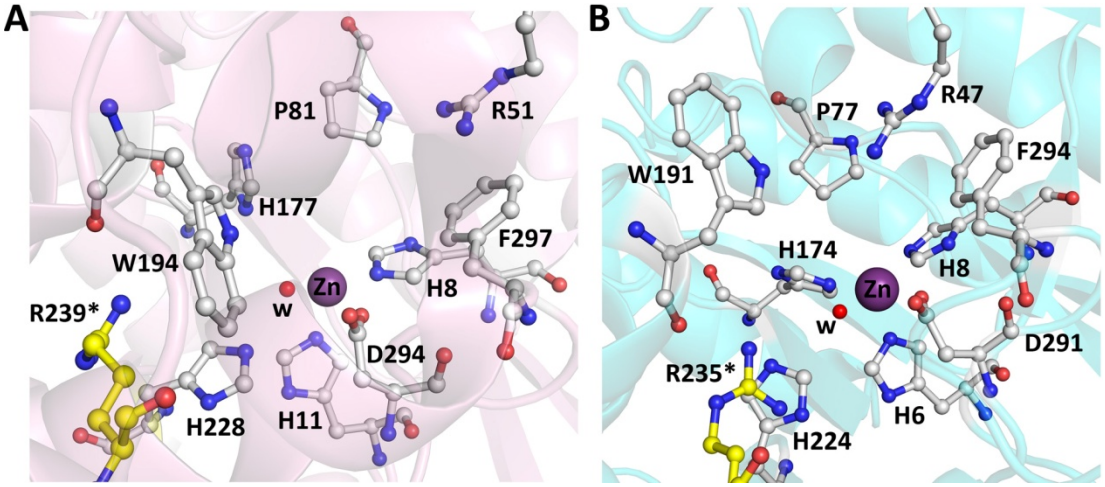


Figure 4.7. A side-by-side comparison of the active site architecture of (A) *Pf*ACMSD and (B) *h*ACMSD shows similar substrate binding environment.

It is evident that alteration of the decarboxylase quaternary structure, between monomer and homodimer, is a possible avenue to regulate the catalytic activity. 5-Carboxyl-uracil decarboxylase (IDCase) is a member of the ACMSD subfamily under the amidohydrolase superfamily and has very high sequence similarity to ACMSD (18,81). This enzyme is the most closely related neighbor of ACMSD in the reconstructed evolution tree. Recently, the crystal structures of IDCase from both Cordyceps militaris and Metarhizium anisopliae were solved as homodimers (221). The overall structure and the active site of IDCase mostly resemble ACMSD. Interestingly, in both of the substrate 5-carboxyl-uracil and inhibitor 5-nitro-uracil bound binary complex structures; the catalytic metal ion does not directly interact with the substrate and inhibitor. The two conserved arginine residues, Arg68 (equivalent to Arg47 of *h*ACMSD) and Arg262* (equivalent to Arg235* of *h*ACMSD) from the neighboring subunit play major roles in substrate/inhibitor binding. Specifically, Arg68 forms two hydrogen bonds with the N1 and O2 of the substrate pyrimidine ring and Arg262* binds to the leaving carboxyl group in the substrate bound structure (PDB entry: 4LAM). This observation further suggests that Arg47 and Arg235*

are responsible for substrate binding not only in ACMSD but also in proteins from the same subfamily. As a result, a dimeric quaternary structure is most likely required for enzyme activity through the ACMSD subfamily.

Since mammals have no mechanism to store free amino acids not used for protein synthesis, tryptophan is primarily directed towards catabolism. During starvation, when tryptophan levels are low, the priority of catabolism is to preserve the kynurenine pathway intermediates for NAD biosynthesis and thus, the activity of ACMSD must be lessened. Using a glycolic intermediate, DHAP, to force ACMSD to change its structure to the catalytically inactive monomeric form is a possibly unprecedented metabolic interrelation between the tryptophan kynurenine catabolic pathway and glycolysis.

4.6 Conclusion

The biochemical, spectroscopic, and structural results led us to conclude that *h*ACMSD is a zinc dependent dimeric enzyme that harnesses its unstable, negatively charged substrate by its two positive charged residues rather than by the metal ion during substrate positioning at the early stage of the catalytic cycle. The quaternary structure variety, previously demonstrated in the bacterial analogous enzyme (80) and now implicated in the human enzyme, is directly linked to the enzyme catalytic activity. Thus, the structural arrangement of the catalytic center offers a potential regulatory mechanism at a critical junction of the downstream tryptophan kynurenine pathway.

CHAPTER 5 EVIDENCE FOR A CATALYTIC ROLE OF THE WATER LIGAND IN α-AMINO β-CARBOXYMUCONATE ε-SEMIALDEHYDE DECARBOXYLASE REVEALED BY KINETICS AND CRYSTAL STRUCTURES IN COMPLEX WITH AZIDE AND BROMIDE

The section of chapter 5 is our manuscript under preparation: Evidence for a catalytic role of the water ligand in α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase revealed by kinetics and crystal structures in complex with azide and bromide. Lu Huo, Vesna de Serrano, Lirong Chen, and Aimin Liu.

5.1 Abstract

We found that the catalytic activity of *Pseudomonas fluorescens* α-amino-β-carboxymuconate-ε-semialdehyde decarboxylase (*Pf*ACMSD) is inhibited by common anions. Bromide and azide were used to study the anion inhibition mechanism. Inhibition assay suggests a competitive inhibition by azide and a mixed-type inhibition by bromide. Co-crystallized structures of the enzyme with these two anions were determined at 2.15 and 2.50 Å resolutions, respectively, and each with two ligand molecules in the active site of the enzyme. Both of the azide molecules are shown to bind in the putative substrate-binding pocket, consistent with its competitive inhibition. In contrast, the bromide-bound ACMSD structure reveals that one bromide is bound at the site where azide molecules bind, and the other substitutes the metal bound water ligand. These results help establish that the metal-bound water ligand plays a major catalytic role and that the Arg51 residue is an important component in the second metal coordination sphere involved in substrate binding.

5.2 Introduction

Metal-bound water molecules have been long thought to be critical for playing a catalytic role in several enzyme families, including amidohydrolase superfamily (19,81), carbonic anhydrase family (181), metallo β-lactamases (222), and restriction endonucleases (223). However, it is very difficult to experimentally prove the catalytic role of a specific water molecule in enzymes. Until now, the only widely accepted and well-documented example for a definitive catalytic role of metal-bound water is found in carbonic anhydrase, in which biochemical, modeling and crystal structural studies all come together to demonstrate that the catalytic driving force is indeed derived from the water ligand (181-183,224). In the past 50 years, hydrolytic water as a catalytic main power in enzymes, especially in many of those zinc-dependent hydrolytic enzymes, has increasingly been appreciated. Thus, it is imperative to explicitly establish a second example from an enzyme other than the members of the carbonic anhydrase family the catalytic role of metal bound water. It has become a textbook knowledge that the metal coordination significantly reduces the pK_a value of water from 15.7 to nearly neutral, making it possible to become a hydroxide anion for subsequent nucleophilic attacks on the substrate. However, this type of mechanism has never been proposed for a decarboxylase prior to our study on an enzyme from the amidohydrolase superfamily. Here we present biochemical and structural evidence for a catalytic role of the water ligand in a decarboxylase.

α-Amino-β-carboxymuconate-ε-semialdehyde (ACMSD) is an important enzyme that controls the biosynthesis of quinolinic acid in kyneurine pathway in mammals (125) and also directs the metabolic flux to energy production in 2-nitobenzoic acid degradation pathway in bacteria (12). ACMSD is a metal-dependent enzyme (14) belonging to the amidohydrolase superfamily (15). The members of this superfamily are believed to initiate catalysis by a water molecule (19,81). In

all but one characterized members the water is a metal ligand. The catalytic role of the water molecule was initially suggested by a pH-dependence study on phosphotriesterase (225) and the crystal structure of adenosine deaminase in complex with a transition state analogue (226), both of which are the well-studied members in the amidohydrolase superfamily. In a most recent study, the enzyme Ligl is shown to be a very unique member of the superfamily. It does not contain a metal ion in its active site (30). However, a water molecule is found in the active site where substrate binds and a nucleophilic attack of the substrate by the water molecule is still believed to be the critical step in catalytic cycle of Ligl.

5.3 Results and Discussion

We found in this work that ACMSD from *Pseudomonas fluorescens* is inhibited by ammonium sulfate during protein purification, following the procedures previously established for substrate generation and kinetic assays (15). We further carried out inhibition assays in the presence of 25 mM following salts in 25 mM HEPES buffer (pH 7.0), NH_4Cl, KCl, NaCl, NaBr, NaI, $NaNO_2$, NaN_3, $NaHCO_3$, CH_3COONa, Na_2SO_4, NaH_2PO_4, and KH_2PO_4. The enzyme assay results showed that NaI, NaN_3, NaBr, and $NaNO_2$ had the strongest inhibition power, whereas CH_3COONa, NaH_2PO_4, and KH_2PO_4 had no inhibition effect. NH_4Cl, KCl, NaCl, $NaHCO_3$ and Na_2SO_4 had mild inhibition effects. Furthermore, NH_4Cl, KCl and NaCl had the same inhibitory effect while NaH_2PO_4 and KH_2PO_4 did not inhibit ACMSD, and therefore we conclude that the cations we have tested are not potential inhibitors for the decarboxylation activity of ACMSD. Hence, our subsequent study was placed on the anion inhibition mechanisms.

Two inhibition patterns were observed among different anioins, *i.e*. competitive inhibition and a mixed-type inhibition. We focused on sodium bromide and sodium azide, the representatives of the two observed inhibitory patterns. Figure 5.1A shows that in the presence of

0, 1.25 mM, 2.5 mM, and 5 mM sodium azide, it inhibits the decarboxylation activity follows a nearly pure competitive inhibition pattern with an inhibition constant of 0.76 ± 0.23 mM. In contrast, in the presence of 0, 2.5, 5 mM and 7.5 mM sodium bromide, the ACMSD k_{cat} values decrease, whereas the corresponding K_m values increase, suggesting a mixed-type, i.e., competitive and noncompetitive, inhibition (Figure 5.1B) with the competitive inhibition constant of 1.30 ± 0.02 mM, and noncompetitive inhibition constant of 10.67 ± 2.66 mM.

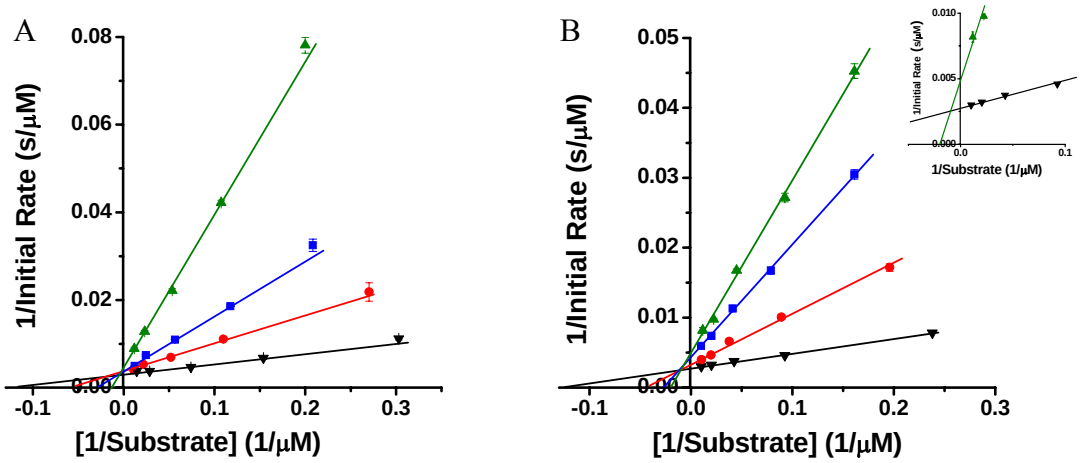


Figure 5.1. ACMSD activity inhibition by azide and bromide salts. Lineweaver-Burk plots of ACMSD inhibited by azide ion competitively (A). Symbols of black, red, blue, and green indicate sodium azide at 0, 1.25, 2.5, and 5 mM. And bromide ion (B), the axis crossing is shown in the inset, indicating a mixed-type inhibition. Symbols of black, red, blue, and green indicate inhibitors at 0, 2.5, 5 and 7.5 mM concentrations, respectively.

The ligand-bound crystal complexes were obtained by co–crystallizing ACMSD with 25 mM NaN_3 or NaBr. Crystals were grown by hanging drop vapor diffusion method at 18 °C. A total drop volume of 3 μL was set up at 1:1 protein to crystallization solution ratio and equilibrated against 500 μL of crystallization solution containing 0.2 M $MgCl_2$, 0.1 M Tris, pH 8.75 and 15% PEG 5000. X-ray diffraction data for N_3–ACMSD were collected at the SER-CAT beamline 22–ID of the Advanced Photon Source (APS), Argonne National Laboratory, Argonne, IL. Data for the NbaD-Br complex were collected using a Rigaku MicroMax–007 HF copper rotating anode

generator operated at 40 kV and 30 mA with Saturn944$^+$ CCD detector at the Emory University X–ray Crystallography Core Facility (Atlanta, GA). All data were indexed, integrated, and scaled using the HKL-2000 suite (137). Structure solutions were obtained by molecular replacement using MolRep (138) from the CCP4i program suite (139) with the published ACMSD structure (PDB entry 2HBV) as a search model. Structures were then refined by iterative cycles of model building in COOT (141) and positional and isotropic B factor refinement using Refmac5 (140) in the CCP4i suite of programs (139) and PHENIX software. The X–ray data collection and refinement statistics are summarized in Table 5.1.

Table 5.1. Data Collection and Refinement Statistics

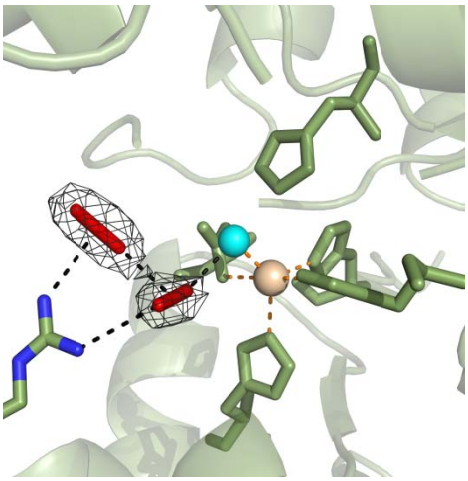
Data collection	**ACMSD-Br^- Complex**	**ACMSD-N_3^- Complex**
detector type	Saturn944$^+$ CCD	MAR300 CCD
space group	*C*2	*C*2
unit cell lengths (Å)	a=152.89, b=48.40, c=109.93	a=152.19, b=44.77, c=110.21
unit cell (°)	α=γ=90, β=126.74	α=γ=90, β=127.76
temperature (K)	100	100
wavelength (Å)	1.54	1.00
Resolution (Å) [a]	35.00-2.15(2.19-2.15)	50.00-2.50(2.54-2.50)
Completeness (%) [a]	95.3(92.9)	94.8(62.8)
R_{merge} (%) [a, b]	12.7(67.3)	7.0(28.7)
$I/\sigma I$ [a]	21.9(2.5)	21.2(2.9)
multiplicity [a]	7.1(7.0)	3.4(1.7)
Refinement		
resolution (Å)	2.15	2.50
no. reflections; working/test	32085/2216	19234/1923
R_{work} (%) [c]	22.6	19.8
R_{free} (%) [d]	30.9	27.9
no. of protein atoms	5298	5190
no. of ligand atoms	4	8
no. of solvent sites	99	60
Average B factor (Å^2)		
protein	50.3	50.7
Zn(II)	45.0	43.0
Br^- or N_3^-	56.1	47.4
solvent	41.3	35.8
Ramachandran statistics [e]		
preferred (%)	88.6	91.6
allowed (%)	10.5	6.7
Root mean square deviation		
bond lengths (Å)	0.014	0.009
bond angles (°)	1.668	1.250

[a] Values in parentheses are for the highest resolution shell.

[b] *R*merge = Σi |Ihkl,i - ‹Ihkl›|/Σhkl Σi Ihkl,i, where Ihkl,i is the observed intensity and ‹Ihkl› is the average intensity of multiple measurements.
[c] *R*work = Σ||F_o|-|F_c||/Σ|F_o|, where |F_o| is the observed structure factor amplitude, and |F_c| is the calculated structure factor amplitude.
[d] *R*free is the *R* factor based on 5% of the data excluded from refinement.
[e] Based on values attained from refinement validation options in COOT.

The overall structures of ACMSD bound anions remain unchanged except some loop movements compared to the ligand-free ACMSD structure (PDB code: 2HBV). Amino acids arrangement in the active site, including metal ligands and second sphere residues, are not interrupted, with the exception of Arg51, which moved away 1.1 Å from its original position for accommodating azide ions and 0.9 Å for accommodating bromide ions. In the N_3–bound ACMSD structure, extra electron density in the substrate binding pocket of chain A revealed the presence of two azide molecules (Figure 5.2A). The two azide molecules are in close proximity to each other (2.8 Å), and both of them are at a distance of forming weak interaction with Arg51 residue (3.3 and 3.4 Å). The position of metal bound water is not disrupted by the azide anions compared to the ligand-free ACMSD structure. The water ligand interacts with one of the azide molecules that is localized just 2.6 Å away from the water ligand (Figure 5.2B). The substrate ACMS is an unstable molecule, autocyclizing to quinolinic acid, and its binding site in ACMSD has not been previously established. Azide inhibits ACMSD following a pure competitive pattern, and accordingly we believe that the binding pocket for azide anions represents the putative substrate binding site. Binding at this pocket is favored partly because that the positively charged Arg 51 provides strong ionic interactions with the compounds with negative charges including ACMS, which possess two carboxylate groups. The maximum rate was also slowed down when higher concentration sodium azide was involved (above 5 mM in this work). It may due to hydrogen bonding interaction between the metal bound water and one of the azide ions is becoming stronger when higher concentration sodium azide salt is used (Figure 5.2B).

A



B

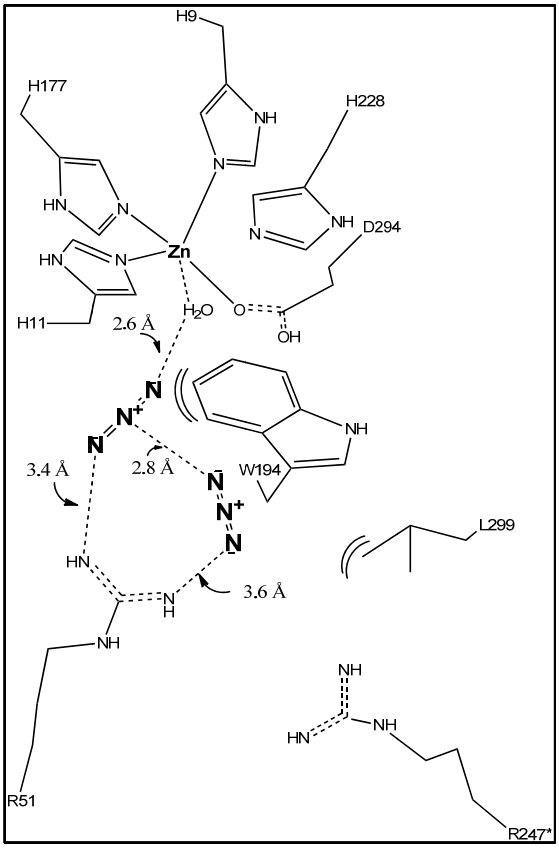
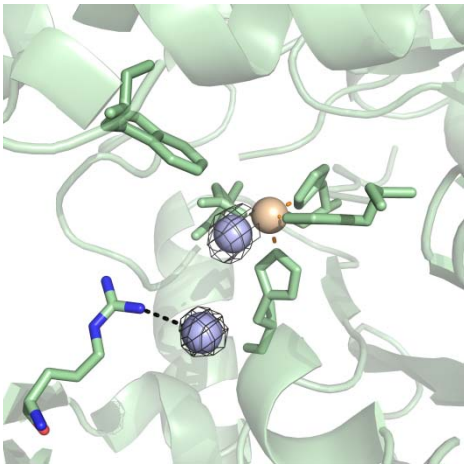


Figure 5.2. Crystal structure of ACMSD inclomplex with azide ions. (A) The structure of *Pf*ACMSD in complex with azide ions. Golden sphere represents zinc metal, blue sphere represents water molecule, and red sticks represent azide ions. The $2F_o$-F_c electron densities are countered at 1.0 σ. (B) The H-bond network around the active site. * indicates the residue R247 is intruded from chain B.

In order to locallized bromide atom in the structure, we took advantage of anomalous scattering of this atom, and for that purpose we collected 360 degrees of diffraction data from the crystals that belong to the C_2 space group. Even though Br atom has an absorption edge at 0.9202 Å, it still has scattering signal of 1.4 electrons at 1.54 Å. Hence, in order to help identify the Br binding site, anomalous difference Fourier electron density maps were calculated with phases from the final refined model of Br^- derivatized structure and coefficients $|F^+| - |F^-|$, corresponding to wavelength–dependent Bijvoet differences. Two Br atoms were placed into the electron density in chain A (Figure 5.3A). One of the bromide ions is ligated to the metal Zn as the fifth ligand which was previously served by a water molecule in the wild type structure. The

other bromide ion was found at the salt bridge distance to the conserved R51 residue (4.0 Å) (Figure 5.3B). No other positively charged residues can be found within 4 Å to this bromide ion.

A



B

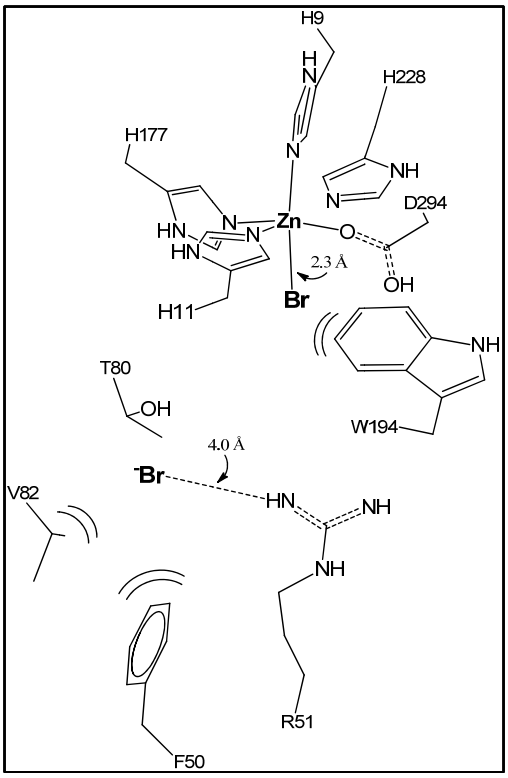


Figure 5.3. Crystal structure of ACMSD in complex with bromide ions. (A) The structure of *Pf*ACMSD in complex with bromide ions. Golden sphere represents zinc metal, purple spheres represent bromide ions. The $2F_o$-F_c electron densities are countered at 2.0 σ. (B) The H-bond network around the active site.

Sodium bromide inhibits the decarboxylation activity of ACMSD in a mixed pattern. The competitive portion is attributed to the binding of one of the bromide ions in the putative substrate binding pocket. Its ionic interaction with Arg51 and hydrophobic interaction with nearby Phe50 and Val82 make it compete with substrate to bind to the active site. The noncompetitive portion is very possibly caused by the ligation of bromide ion to the zinc metal, thus replacing the metal bound water. This is likely due to the loose of the water ligand that serves as a major catalytic driving force, already shown in the mechanism of human carbonic anhydrase (224,227).

CHAPTER 6 CAPTURE OF TETRAHEDRAL THIOHEMIACETAL AND THIOACYL INTERMEDIATES IN AN AMINOMUCONATE SEMIALDEHYDE DEHYDROGENASE

The section of chapter 6 is our manuscript under preparation: Capture of tetrahedral thiohemiacetal and thioacyl intermediates in an aminomuconate semialdehyde dehydrogenase. Lu Huo, Ian Davis, Fange Liu, Babak Andi, Shingo Esaki, Hiroaki Iwaki, Yoshie Hasegawa, Allen M. Orville, and Aimin Liu. This work was supported, in whole or in part, by National Institutes of Health grant GM108988, National Science Foundation grant CHE-0843537, and Georgia Research Alliance Distinguished Scientist Program (A.L.), Molecular Basis of Disease Area of Focus graduate fellowship (L.H. and I.D.), Center for Diagnostics and Therapeutics (F.L.), Georgia State University Dissertation Award (L.H.), and funds from Mext Haiteku (Y.H.), Offices of Biological and Environmental Research award FWP BO-70 of the US Department of Energy and NIH grant P41GM103473 (A.B. & A.M.O.). We thank Dr. Siming Wang for assistance of Mass spec analysis. X-ray data was collected at the Southeast Regional Collaborative Access Team (SER-CAT) 22-ID beamline at the Advanced Photon Source, Argonne National Laboratory. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. W-31-109-Eng-38. Single-crystal spectroscopy data was obstained at beamline X26-C of the National Synchrotron Light Source (NSLS), Brookhaven National Laboratory with the support of the U.S. Department of Energy under Contract No. DE-AC02-98CH10886.

6.1 Abstract

Aldehydes are ubiquitous intermediates in metabolic pathways, and their innate reactivity can often make them quite unstable. There are several aldehydic intermediates in the metabolic

pathway for tryptophan degradation which can decay into neuroactive compounds that have been associated with numerous neurological diseases. An enzyme of this pathway, 2-aminomuconate-6-semialdehyde dehydrogenase, which is responsible for 'disarming' the final aldehydic intermediate, has yet to be studied at the molecular level. Its activity steers the vast majority of metabolic flux to further catabolism and prevents overproduction of a metabolic dead-end, picolinic acid. Here we show crystal structures of a bacterial analogue enzyme in five catalytically relevant forms: resting state, one binary and two ternary complexes, and a thioacyl intermediate. We also report the structures of a tetrahedral thiohemiacetal intermediate and a thioacyl intermediate from an active site mutant. These intermediates were characterized by single-crystal, solution electronic absorption spectroscopy, and the mutant intermediate was confirmed by mass spectrometry. The crystal structures also reveal that the substrate must undergo an *E*/*Z* isomerization prior to oxidation. This study provides an in-depth mechanistic understanding of this dehydrogenase.

6.2 Introduction

The dominant route of tryptophan catabolism, the kynurenine pathway, has recently garnered increased attention given its apparent association with numerous inflammatory and neurological conditions, *e.g*. gastrointestinal disorders, depression, Parkinson's disease, Alzheimer's disease, Huntington's disease, and AIDS dementia complex (127,128,228-231). Though the precise mechanism by which the kynurenine pathway influences these diseases has not yet been fully elucidated, it has been determined that several metabolites of this pathway are neuroactive. Notably, the concentration of quinolinic acid, a non-enzymatically derived decay product of an intermediate of the kynurenine pathway used for NAD^+ biosynthesis, is elevated over 20-fold in patients' cerebrospinal fluid with AIDS dementia complex, aseptic meningitis, opportunistic

infections, or neoplasms (232), and more than 300-fold in the brain of HIV-infected patients (214). This NAD^+ precursor has also been shown to be an agonist of *N*-methyl-D-aspartate receptors, and an increase of its concentration may lead to over-excitation and death of neuronal cells (2,129).

The apparent medical potential of the kynurenine pathway warrants detailed study and characterization of its component enzymes and their regulation. One enzyme in particular, 2-aminomuconate-6-semialdehyde dehydrogenase (AMSDH), is responsible for oxidizing the unstable metabolic intermediate 2-aminomuconate-6-semialdehyde (2-AMS) to 2-aminomuconate (2-AM). Based on sequence alignment, AMSDH is a member of the hydroxymuconic semialdehyde dehydrogenase (HMSDH) family under the aldehyde dehydrogenase (ALDH) superfamily (86). Aldehyde dehydrogenases are prevalent in both prokaryotic and eukaryotic organisms and are responsible for oxidizing aldehydes to their corresponding carboxylic acids. They use $NAD(P)^+$ as a hydride acceptor to harvest energy from the primary substrate and generate NAD(P)H, which provides the major reducing power to maintain cellular redox balance (233,234). In addition to being commonly occurring metabolic intermediates, aldehydes are reactive electrophiles, making many of them toxic. Enzymes of the ALDH superfamily are typically promiscuous with regards to their substrates, however in recent years, this superfamily has had several new members identified with greater substrate fidelity, especially when the substrate is identified as a semialdehyde (235).

The putative native substrate of AMSDH, 2-AMS, is a proposed metabolic intermediate in both the 2-nitrobenzoic acid degradation pathway of *Pseudomonas fluorescens* KU-7 (11) and the kynurenine pathway for L-tryptophan catabolism in mammals (2,14,129). In the presence of NAD^+ and AMSDH, 2-AMS is oxidized to 2-AM (Fig. 6.1A), however it can also spontaneously

decay to picolinic acid and water with a half-life of 35 s at neutral pH (10). Due to its instability, 2-AMS has not yet been isolated, leaving its identity as the substrate of AMSDH an inference based on decay products and further metabolic reactions. There are several reasons for the poor understanding of this pathway: it is complex with many branches, some of the intermediates are unstable and difficult to characterize, and several enzymes of the pathway, including AMSDH, are not well understood. Hence, the structure of AMSDH will help to address questions such as what contributes to substrate specificity for the semialdehyde dehydrogenase, how 2-AMS is bound and activated during catalysis, and to help identify its counterpart in the human genome.

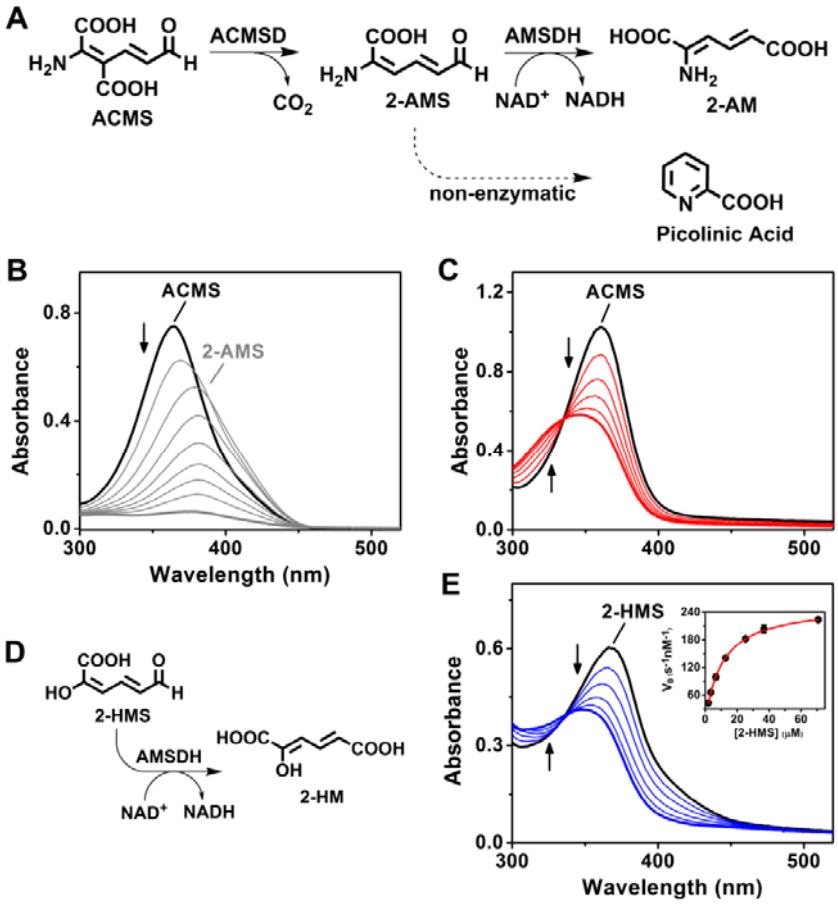


Figure 6.1. Activity of AMSDH. (A) Reaction scheme showing the enzymatic generation of 2-AMS. (B) Representative assay showing the ACMSD catalyzed consumption of ACMS. (C) Coupled-enzyme assay in which AMSDH oxidizes 2-AMS to 2-AM. (D) Reaction scheme showing 2-HMS oxidation by AMSDH. (E) Representative assay showing the activity of AMSDH on 2-HMS. The inset is a Michaelis-Menten plot.

6.3 Methods

Expression and Purification of PfAMSDH and PfACMSD. To construct a His_{10}-tagged AMSDH expression plasmid, *nbaE* gene from *P. fluorescens* (accession: AB088043.2) encoding AMSDH was amplified by the polymerase chain reaction (PCR) using genomic DNA of *P. fluorescens* strain KU-7 as a template and primers 5'-GGAATTCCATATGAATACCTTACCAAGTCAAG-3' and 5'-CCCTCGAGTTAAATTTTTATGCAGATGTTGG-3' (built-in *Nde*I and *Xho*I sites are underlined). The PCR product was purified from a 0.8% agarose gel, digested with *Nde*I and *Xho*I, and ligated in the equivalent sites of pET-16b (Novagen). Ligation product was transformed to *Escherichia coli* BL21(DE3) for protein expression. A single colony was introduced to 10 mL of autoclaved LB medium containing 100 μg/mL ampicillin and cultured at 37 °C. When cells reached ca. 0.6 OD at 600 nm, 1.5 mL of cells were diluted into 500 mL autoclaved LB medium containing ampicillin. The cells were cultured in 37 °C until the optical density reached ca. 0.8 at 600 nm. Isopropyl β-D-1-thiogalactopyranoside was then added to a final concentration of 0.6 mM, and the temperature was lowered to 28 °C for 12 hours to induce AMSDH expression before the cells were harvested by centrifugation at 8,000 g. The harvested cells were then resuspended in 50 mM potassium phosphate buffer, pH 8.0, containing 300 mM NaCl and 5% glycerol. The cell slurry was passed through an M-110P Microfluidics cell disruptor and the debris was removed by centrifugation at 27,000 g for 30 min at 4°C. The supernatant containing AMSDH was purified using a Ni-NTA affinity column on an ÄKTA FPLC system (GE Healthcare). The major fraction with AMSDH activity was eluted by increased imidazole concentration. Purified protein was concentrated and desalted on a pre-packed HiTrap desalting column (GE Healthcare) using buffer containing 50 mM HEPES (pH

7.5), 150 mM NaCl, and 1 mM DTT. Expression, purification and protein re-constitution of ACMSD were performed as described previously (26).

Site-directed Mutagenesis. C302S, E268A, R120A, and R464A single mutation variants were constructed by the PCR overlap extension mutagenesis technique (132). Plasmid containing AMSDH from *P. fluorescens* was used as a template. The forward primers used in the site-directed mutagenesis are 5'-CAACTCGGGGCAGGTCagcCTGTGTTCCGAACG-3' for C302S, 5'-GAAAGAAGTGTCTTTCgcgTTGGGGGGCAAGAACG-3' for E268A, 5'-GGACCCTCGATATTCCTgcgGCCATTGCCAACTTTC-3' for R120A, and 5'-GAACACCTGGTACTTGgcgGATCTGCGTACGCC-3' for R464A. The insert of each mutant was verified by DNA sequencing and the positive clone was transformed to E. coli BL21(DE3). The expression and purification of the mutants are the same as wild-type AMSDH.

Preparation of ACMS and 2-HMS. ACMS was generated by catalyzing the insertion of molecular oxygen to 3-hydroxyanthranilic acid by purified, Fe^{2+} reconstituted 3-hydroxyanthranilate 3,4-dioxygenase as described previously (14,26). 2-HMS is generated non-enzymatically from ACMS following a previously established method (236). The pH of solutions containing ACMS was adjusted to ~2 by the addition of hydrochloric acid. 2-HMS formation was monitored on an Agilent 8453 diode-array spectrophotometer at 315 nm. The solutions were then neutralized with sodium hydroxide once the absorbance at 315 nm stopped increasing. 2-HMS at neutral pH has a maximum UV absorbance at 375 nm (236).

Enzyme Activity Assay Using 2-HMS as Substrate. Steady-state kinetics analyses were carried out at room temperature on an Agilent 8453 diode-array spectrophotometer. Reaction buffer contains 25 mM HEPES and 1 mM NAD^+, pH 7.5. Consumption of the substrate 2-HMS by 200 nM AMSDH was detected by monitoring the decrease of its absorbance at 375 nm with a

molar extinction coefficient of 43,000 $M^{-1}cm^{-1}$ for 15 s with a 0.5 s integration time. For mutants, 700 nM protein and a wavelength of 420 nm, ε_{420} 11,180 $M^{-1}cm^{-1}$, was used. UV absorbance at 375 nm decreased and blue shifted to 295 nm, the maximum UV absorbance for the product, 2-hydroxymuconic acid. This is consistent with previously reported data in which the ending compound was purified and verified as the correct product (236). The pre-steady state spectra were obtained by using Applied Photophysics Stopped-Flow Spectrometer SX20 (UK) with the mixing unit hosted inside an anaerobic chamber made by Coy Laboratory Products (MI, USA). Pre-steady state activity used the same reaction buffer but with 23 μM AMSDH or E268A and 25 μM 2-HMS and were carried out at 10 °C. The change in absorbance was monitored for 1.0 s.

Crystallization, Data Collection, and Refinement. Purified AMSDH samples at a final concentration of 10 mg/mL containing no NAD^+ or 10 equivalents of NAD^+ were used to set up sitting-drop vapor diffusion crystal screening trays in Art Robbins 96-well Intelli-Plates using an ARI Gryphon crystallization robot. The initial crystallization conditions were obtained from PEG-Ion 1/2 (Hampton Research) screening kits at room temperature. The screened conditions were optimized by increasing protein concentration to 40 mg/mL and lowering crystallization temperature to 18 °C. NAD^+-bound AMSDH crystals were obtained from drops assembled with 1.5 μL of protein (pre-incubated for 10 minutes with 10 equivalents of NAD^+) mixed with 1.5 μL of a reservoir solution containing 20% polyethylene glycol 3350 and 0.2 M sodium phosphate dibasic monohydrate, pH 9.1, by hanging drop diffusion in VDX plates (Hampton Research). Pyramid shaped crystals which diffract up to ~1.9 Å appeared overnight. The reservoir solution for crystallizing the cofactor-free AMSDH crystals contains 12% polyethylene glycol 3350, 0.1M sodium formate, pH 7.0. Crystals belonging to the same space group formed within 2-3 days with an irregular plate shape and diffracted up to ~2.2 Å. NAD^+-AMSDH crystals were

used for substrate-soaking experiments. Crystals were transferred to mother liquor solution containing ~1 mM 2-HMS and incubated for 10 – 180 minutes before flash-cooling in liquid nitrogen. Soaking 2-AMS as substrate is more complicated because of its instability. Crystallization solution containing ~1.5 mM ACMS were used for soaking. After transferring several crystals to the soaking solution (8 μL), 2 μL of 1 mM purified ACMSD was included to catalyze the conversion of ACMS to 2-AMS. Crystals were flash-frozen after a 5 min-incubation. Crystallization solution containing 20% glycerol or ethylene glycol was used as cryoprotectant. X-ray diffraction data were collected on SER-CAT beamline 22-ID or 22-BM of the Advanced Photon Source, Argonne National Laboratory.

Ligand Refinement and Molecular Modeling. The first AMSDH structure, the cofactor NAD^+ bound structure, was solved by the molecular replacement method with the Advanced Molecular Replacement coupled with Auto Model Building programs from the PHENIX software using 5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (PDB: 2D4E) as a search model, which shares 39% of amino acids sequence identity with *P. fluorescens* AMSDH. The ligand-free, mutant and ternary complex structures were solved by molecular replacement using the refined NAD^+-AMSDH as the search model. Refinement was conducted using PHENIX software (196). The program Coot was used for electron density map analysis and model building (141). NAD^+/NADH, substrates 2-AMS and 2-HMS, and Cys-substrate covalent adduct intermediate were well-defined and added to the model based on the $2F_o$ - F_c and F_o - F_c electron density maps. Refinement was assessed as complete when the F_o - F_c electron density contained only noise. The structural figures were generated using PyMOL software (http://www.pymol.org/).

Single-Crystal Spectroscopy. Electronic absorption spectra from single crystals held at 100 K were collected at beamline X26-C of the National Synchrotron Light Source (NSLS) (237). The electronic absorption data were typically obtained between 200 – 1000 nm with a Hamamatsu (Bridgewater, N.J.) L10290 high power UV-Vis light source. The lamp was connected to one of several 3-meter long solarization-resistant optical fibers with an internal diameter of 115, 230, 400, or 600 μm (Ocean Optics, Dunedin, FL). The other end was connected to a 40 mm diameter, 35 mm working distance 4x, Schwardchild design reflective microscope objective (Optique Peter, Lentilly France). The spectroscopy spot size is a convolution of the optical fiber diameter and the magnification of the objective, which in this case produced 28, 50, 100, or 150 μm diameter spots, respectively. Photons that passed through the crystal were collected with a second, aligned objective that was connected to a similar optical fiber or one with a slightly larger internal diameter. The spectrum was then recorded with either an Ocean Optics USB4000 or QE65000 spectrometer. Anisotropic spectra and an image of the crystal/loop were collected as a function of rotation angle in five-degree increments. These were analyzed by XREC (238) to determine the flat face and optimum orientation.

Mass Spectrometry. To prepare samples for ESI mass spectrometry (MS), as-isolated E268A AMSDH was buffer-exchanged to 10 mM Tris (pH 8.0) by running through a desalting column (GE Healthcare). Intermediate bound E268A was obtained by mixing E268A with 3 eq. of 2-HMS. Excess 2-HMS was removed by desalting chromatography using the same buffer. Desalted proteins were concentrated to a final concentration of 20 μM. Freshly prepared samples were rinsed by acetonitrile and 0.1% formic acid (1:1 ratio) before injection. MS experiments were conducted using a Waters (Milford, MA) Micromass Q-TOF micro (ESI-Q-TOF) instrument operating in positive mode. The capillary voltage was set to 3500 V, the sample cone

voltage to 35 V, and the extraction cone voltage to 2 V. The source block temperature and the desolvation temperature were set to 100 and 120 °C, respectively. The samples were introduced into the ion source by direct injection at a flow rate of 5 μL/min. The raw data containing multiple positively charged protein peaks were deconvoluted and smoothed using MassLynx 4.1.

6.4 Results

Catalytic activity of wild-type AMSDH. Due to the unstable nature of its substrate, 2-AMS, activity of AMSDH was detected using a coupled enzyme assay that employed its upstream partner, α-amino β-carboxymuconate ε-semialdehyde decarboxylase (ACMSD), to generate 2-AMS *in situ*. ACMSD transforms α-amino β-carboxymuconate ε-semialdehyde (ACMS) (λ_{max} at 360 nm) to 2-AMS (λ_{max} at 380 nm) (10,14). As seen in Figure 6.1B, in an assay which uses only ACMSD, the absorbance peak of its substrate, ACMS, red-shifts to 380 nm as 2-AMS is formed. The absorbance at 380 nm then quickly decays as 2-AMS decays to picolinic acid, which has no absorbance features above 200 nm. In a coupled-enzyme assay, ACMSD, AMSDH, and NAD^+ are included in the reaction system. As shown in Figure 6.1C, ACMS is still consumed, however there is no red shift observed because 2-AMS is enzymatically converted to 2-AM (λ_{max} at 325 nm) rather than accumulating and decaying. The production of 2-AM requires that an equimolar amount of NAD^+ be reduced to NADH (λ_{max} at 339 nm). A stable alternative substrate, 2-hydroxymuconate-6-semialdehyde (2-HMS), was used to pursue kinetic parameters (Figure 6.1D), when using saturating NAD^+ concentrations (≥ 1 mM), the k_{cat} and K_m of AMSDH for 2-HMS were 1.30 ± 0.01 s^{-1} and 10.4 ± 0.2 μM, respectively (Figure 6.1E).

Crystal structures of AMSDH. Five crystal structures of wild-type AMSDH, including the ligand-free (2.20 Å resolution), NAD^+-bound binary complex (2.00 Å), ternary complex with NAD^+ and substrate 2-AMS (2.00 Å) or 2-HMS (2.20 Å), and a thioacyl intermediate (1.95 Å),

have been solved by molecular replacement. All five structures belong to space group $P2_12_12_1$. Data collection and refinement statistics are listed in Table 6.1. The complete AMSDH model includes four polypeptides per asymmetric unit describing one homo-tetramer (Figure 6.2A). Each monomer of AMSDH contains three domains: a subunit interaction domain, a catalytic domain, and an NAD^+ binding domain (Figure 6.2B).

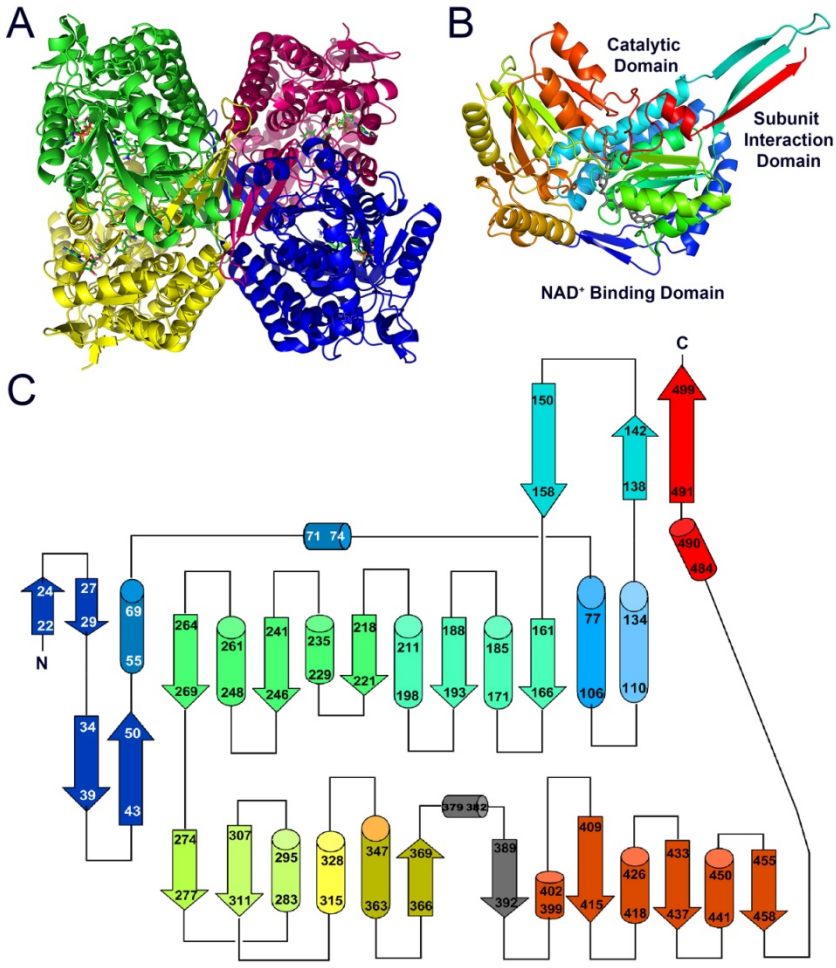


Figure 6.2. Overall crystal structure of AMSDH. (A) View of the tetramer. (B) Structure of one AMSDH subunit with NAD^+ and substrate 2-AMS. (C) Topology diagram showing the AMSDH secondary structure.

Each polypeptide contains 500 amino acids, but the first 16 - 17 amino acids are not included in our model due to missing electron density even though full-length AMSDH was used for crystallization. It is most likely that these N-terminal residues belong to a random coil as in most

other ALDH structures. The first 135 residues comprise a cap which surrounds the cofactor binding domain. This cap region starts with two β-hairpin motifs (residues 22 - 50) and is followed by four α-helixes (residues 51 - 134). The sequence then extends to the subunit interaction domain with two beta strands (residues 138 -158). The central strand of the cofactor binding domain starts at residue 161 and stops at residue 266 and resembles a distorted Rossmann fold. The catalytic domain (residues 267 - 476) is based on a topologically related βαβ polypeptide fold and contains a thiol, Cys302, in the catalytic center. The sequence ends with a C-terminal helix and a beta strand (residue 477 - 500), which is part of the oligomerization domain. The active site is located in the region between the NAD^+ binding domain and the catalytic domain with entrances for NAD^+ and substrate on two separate sides.

In the structure of the co-crystallized binary complex, an NAD^+ molecule is present in an extended, anti-conformation in the N-terminal, co-substrate binding domain of each monomer (Figure 6.3A). The electron density map of NAD^+ is well defined, and the interactions between the protein and NAD^+ are equivalent in all four subunits as shown in Figure 6.3E. The NAD^+-bound AMSDH structure is similar to the ligand-free structure with an aligned RMSD of 0.239 Å. Residues which belong to the NAD^+ binding pocket are also well aligned with the exception of Cys302, Arg108, and Leu116 (Figure 6.4). Upon binding NAD^+, the thiol moiety of Cys302 rotates so that the sulfur is 2.3 Å closer to the substrate binding pocket and away from the nicotinamide head of NAD^+.

The adenine ribose ring and the two phosphate groups are the main components that stabilize the NAD^+ position by interacting with protein residues which all belong to the surrounding loops, except Thr250, which belongs to the 8^{th} α-helix. The nicotinamide half of NAD^+ has less interaction with local residues. Two oxygen atoms (O2 and O3) of the ribose ring form H-bonds

with OE2 and OE3 of Glu404, respectively. On the other ribose ring, the O3 atom forms an H-bond with the Nz of Lys192 while its O2 atom forms H-bonds with both Lys192 (Nz) and Glu195 (OE1). The O1 belonging to the phosphate group nearer the nicotinamide is H-bonded to NE1 of Trp168. The anionic O2 of the other phosphate group forms H-bonds with both Thr250 (OG1) and Glu247 (N) (Figure 6.3E).

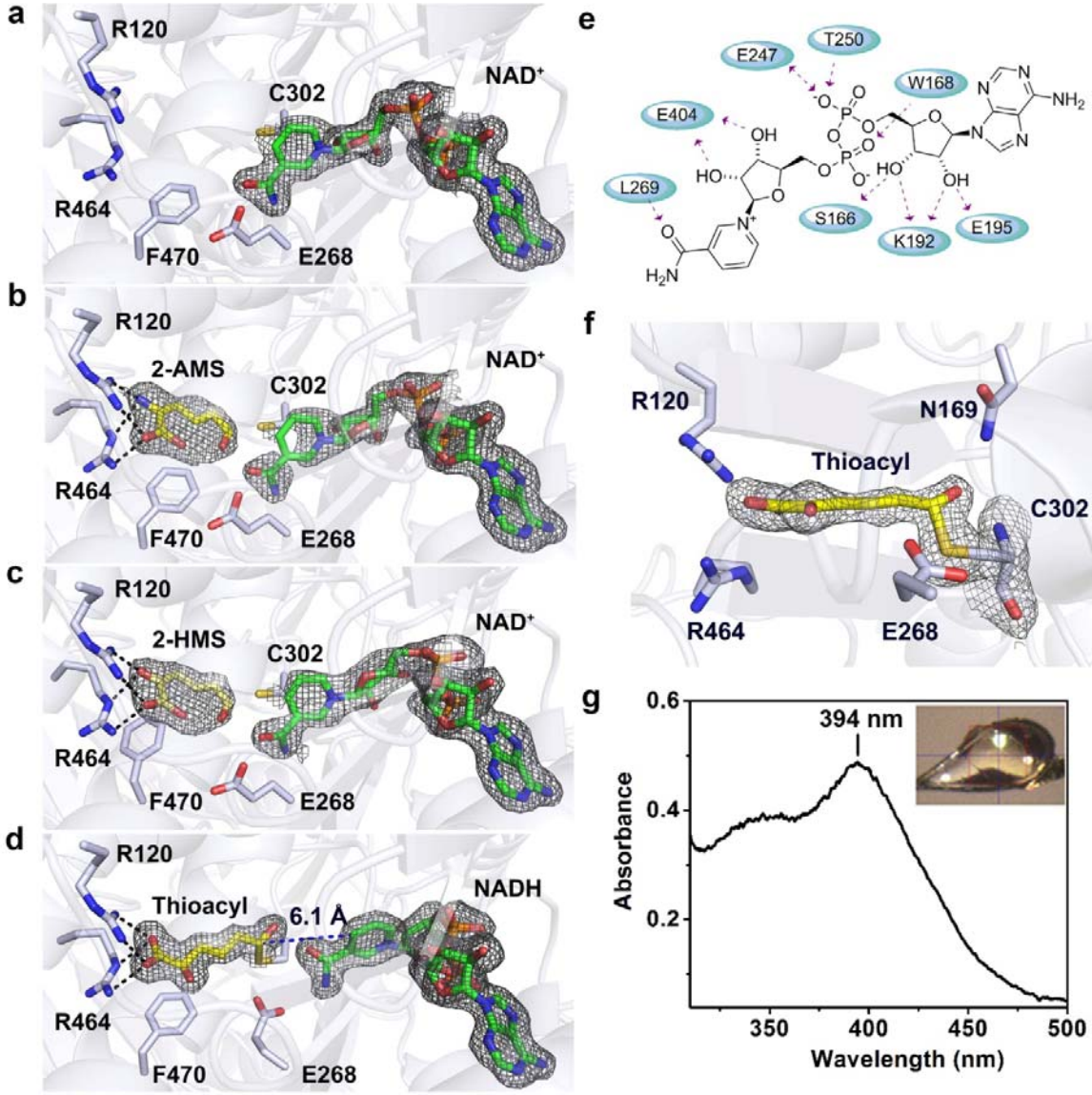


Figure 6.3. Crystal structures of wild-type AMSDH and single-crystal electronic absorption spectrum of a catalytic intermediate. Active site structure of AMSDH-NAD^+ complex (A), ternary complex of AMSDH-NAD^+-2-AMS (B), ternary complex of AMSDH-NAD^+-2-HMS (C), and thioacyl intermediate (D). (E) 2D-interaction diagram for NAD^+ binding. (F) Close-up of the thioacyl intermediate in D. (G) Single-crystal electronic absorption spectrum of D.

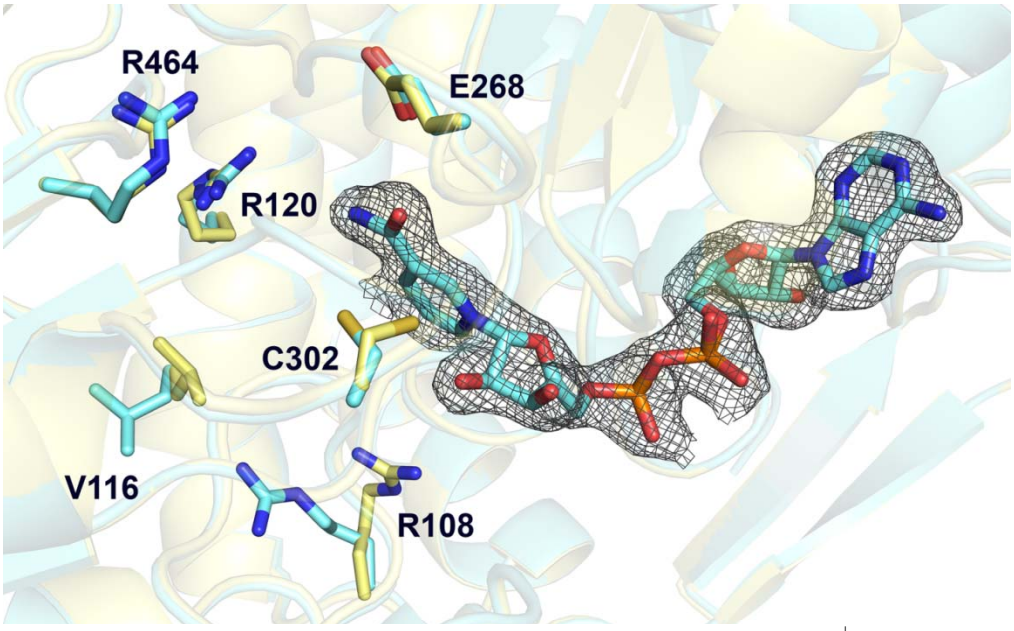


Figure 6.4. Local changes at the active site of AMSDH induced by NAD^+ binding.

Crystal structures of enzyme substrate ternary complexes. Structures of AMSDH in ternary complex with co-substrate NAD^+ and its primary substrates, 2-AMS or 2-HMS, were obtained by soaking co-crystallized AMSDH-NAD^+ crystals with substrate for 5 and 10 min for 2-AMS and 2-HMS, respectively. Extra density which fits with the corresponding substrate molecule was observed in the active site of each subunit. The co-substrate NAD^+ in the ternary complex structures is bound in the same manner as in the binary complex. Binding of the primary substrates introduced minimal change to the protein structure; the RMSD for the superimposed structures of substrate-free with 2-AMS- and 2-HMS-bound ternary complex structures are 0.170 and 0.276 Å, respectively. These two primary substrates bind to AMSDH in an identical fashion, with two arginine residues, Arg120 and Arg464, playing important roles in stabilizing the substrate by forming four H-bonds with one of the carboxyl oxygens and the 2-amino or hydroxy group of 2-AMS (Figure 6.3B) or 2-HMS (Figure 6.3C), respectively. Mutation of Arg120 to alanine causes a moderate decrease of the k_{cat} to 0.7 ± 0.2 s^{-1} from 1.30 ± 0.01 s^{-1} and a dramatic increase of the K_m with a lower bound of 446.3 ± 195.9 μM (an accurate determination of the K_m

is hindered by insufficient 2-HMS concentrations) compared to 10.4 ± 0.2 μM in the wild-type (Figure 6.5, top). Mutation of Arg464 to alanine decreased the k_{cat} to ~ 0.3 s^{-1}, and not only increased the K_m to ~ 170 μM, but also leads to a substrate inhibition effect with a K_i of ~ 6 μM (Figure 6.5, bottom). This substrate inhibition effect is likely caused by unproductive binding of a second substrate molecule in the space created by the deletion of arginine.

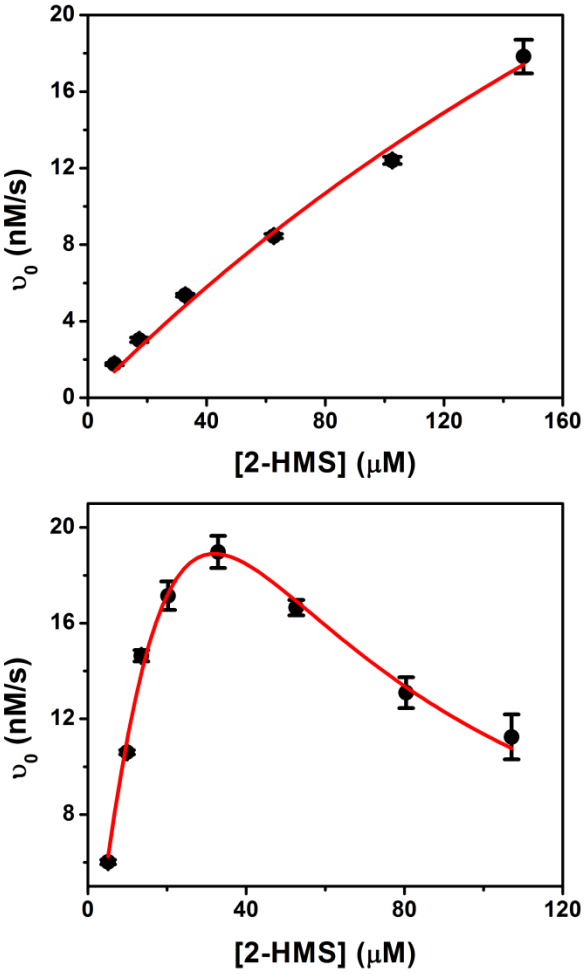


Figure 6.5. Kinetic assays of R120A and R464A with 2-HMS. Top, activity of R120A fit with the Michaelis-Menten. Bottom, activity of R464A fit with the Michaelis-Menten equation with substrate inhibition.

Catalytic intermediates trapped after ternary complex formation. Enzyme-NAD^+ binary complex crystals were soaked in mother liquor containing 2-HMS for a range of time points from 5 minutes to more than 3 hours before flash-cooling in liquid nitrogen. In a crystal which

was soaked for 40 min, an intermediate was trapped and refined to the resolution of 1.95 Å (Figure 6.3D). Crystals soaked for longer time points gave a similar intermediate with poorer resolution. In this structure, 2-HMS has isomerized from the (2*E*, 4*E*)-2-hydroxymuconate-6-semialdehyde of the substrate-bound ternary structure to a (2*Z*, 4*E*) conformation and interacts with Arg120 and Arg464 with both of its carboxyl oxygens rather than one carboxyl oxygen and the 2-hydroxy oxygen as shown in the 2-HMS ternary complex structure. Fitting this density with the 2*E*, 4*E* conformation resulted in unsatisfactory $2F_o$ - F_c and F_o - F_c density maps as shown in Figure 6.6A. Likewise, attempting to fit the 2*Z*, 4*E* isomer to the ternary complex structure did not produce satisfactory results (Figure 6.6B). Upon *E* to *Z* isomerization, the carbon chain of the substrate extends, and the distance between its 6th carbon and Cys302's sulfur is now at 1.8 Å, which is within covalent bond distance for a carbon-sulfur bond. Also, the continuous electron density between Cys302-SG and 2-HMS-C6 indicates the presence of a covalent bond (Figure 6.3F).

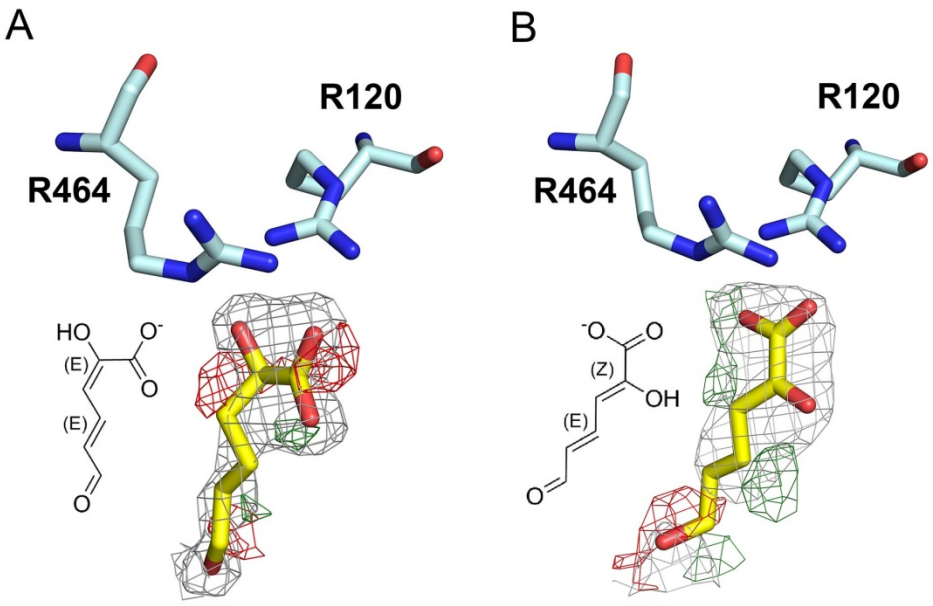


Figure 6.6. Alternate fitting of substrate-bound ternary complex and thioacyl intermediate.

Another feature of this intermediate is that the nicotinamide ring of NAD^+ has moved 4.6 Å away from the active site and adopted a bent conformation (Figure 6.3D) compared to the position in the binary or ternary complex structures (Figure 6.3A, B, and C), indicating that it has been reduced to NADH. Thus, this intermediate is assigned as a thioacyl enzyme-substrate adduct. The single crystal electronic absorption spectrum of the sample has an absorbance maximum at 394 nm (Figure 6.3G). The same absorbance band was observed in crystals soaked

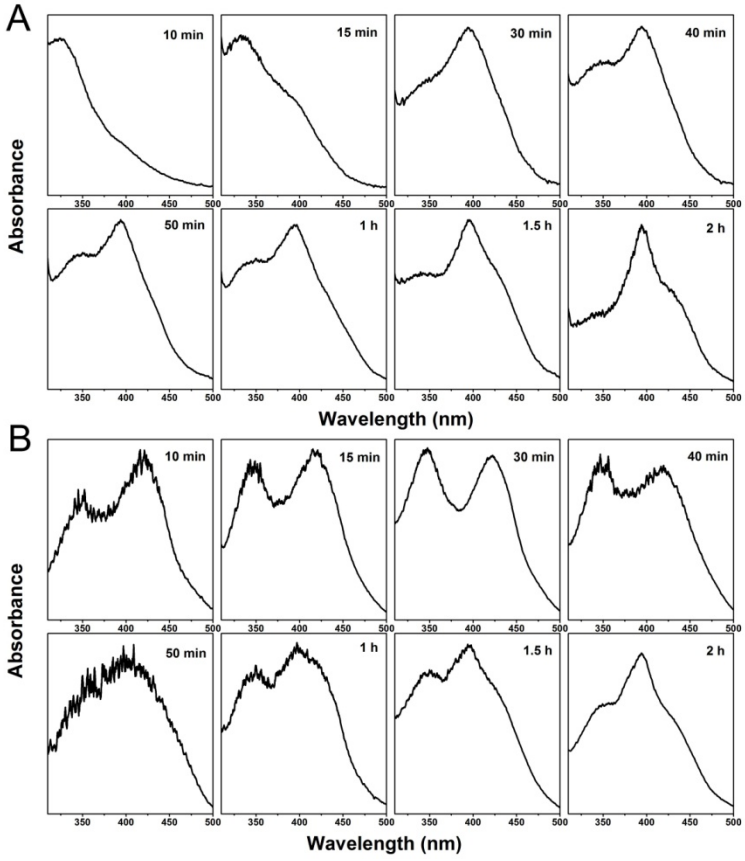


Figure 6.7. Single-crystal electronic absorption spectra of wtAMSDH and E268A AMSDH co-crystallized with NAD^+ and soaked with 2-HMS before flash-freezing.

with 2-HMS from 30 min to 2 h (Figure 6.7A). However, this long-lived intermediate in the crystal was not observed in solution with millisecond to second time resolution in stopped-flow experiments (Figure 6.8A). Thus, it is either present in an earlier time domain (sub-

milliseconds), or alternatively, it may not accumulate in solution because NADH can readily dissociate in solution, whereas it may be trapped in the active site when in the crystalline state.

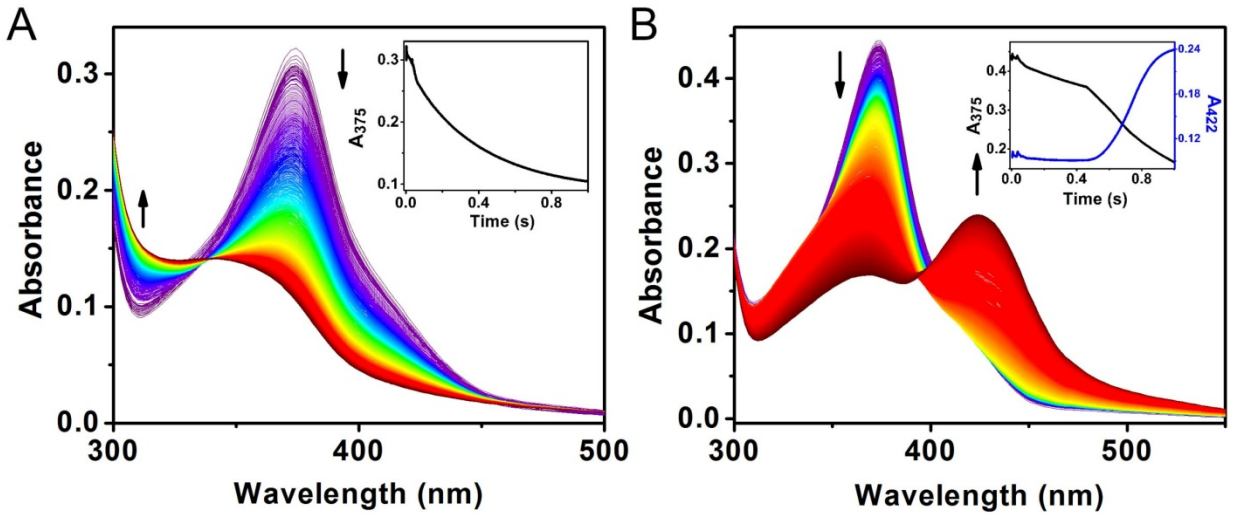


Figure 6.8. Time-resolved, stopped-flow UV-Vis spectra of the reactions catalyzed by wtAMSDH and E268A for 1 s.

Another notable change in the intermediate structure is the movement of the side chain of Glu268, which rotates 73° towards the active site. In order to probe the function of Glu268, we constructed an alanine mutant and found that it exhibited no activity in steady-state kinetic assays. Interestingly, E268A exhibits completely different pre-steady state activity than the wild-type enzyme. As shown in Figure 6.8B, a peak at 422 nm was formed concomitant with decay of the 2-HMS peak within the one second of reaction. This new species is generated stoichiometrically upon titration of 2-HMS with E268A (Figure 6.9D). The moiety which gives rise to this new absorbance band is stable for minutes at room temperature and cannot be separated from the protein by membrane filtration-based methods (26), suggesting that it is covalently bound to the protein. The formation of an enzyme-substrate adduct in the E268A mutant was investigated by mass spectrometry. For as-isolated E268A, the resultant multiply charged states (Figure 6.10) were deconvoluted to obtain a molecular weight (MW) of 56,252 Da (Figure 6.11, top). This value is in good agreement with the predicted MW of E268A AMSDH

plus an N-terminal His-tag and linking residues, 56,251 Da. The second largest peak in the deconvoluted spectrum has a MW 177 Da greater than that of the most abundant signal. This is likely due to post translational modification of the His-tag; α-N-Gluconoylation of His-tags has been observed in *E. coli* expressed proteins which cause 178 Da extra mass (239). The mutant protein was then treated with the alternate substrate, 2-HMS, and the mass spectrum shows a new major peak at 56,390 Da (Figure 6.11, bottom), 138 Da heavier than the as-isolated mutant. Similarly, the second most abundant peak corresponds to a His-tag modified mutant plus 139 Da. In this spectrum, the peaks arising from the as-isolated mutant are substantially reduced, indicating that 2-HMS, 141 Da, has formed a covalent adduct with the E268A mutant enzyme.

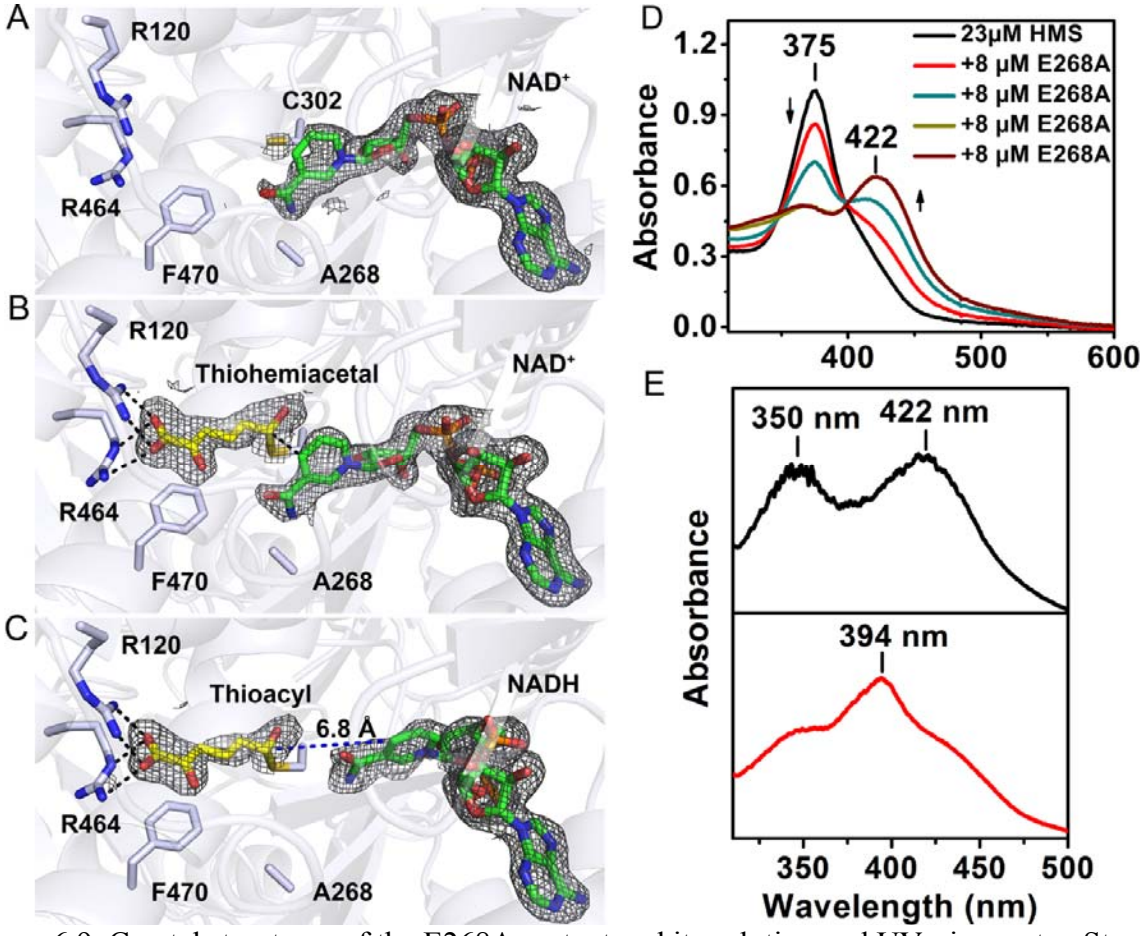


Figure 6.9. Crystal structures of the E268A mutant and its solution and UV-vis spectra. Structure of the active site of (A) E268A-NAD^+ binary complex, (B) thiohemiacetal intermediate, (C) thioacyl intermediate. (D) Solution UV-vis spectra of a titration of 2-HMS with E268A. (E) Single-crystal UV-vis spectrum of the intermediate in B (top panel) and C (bottom panel).

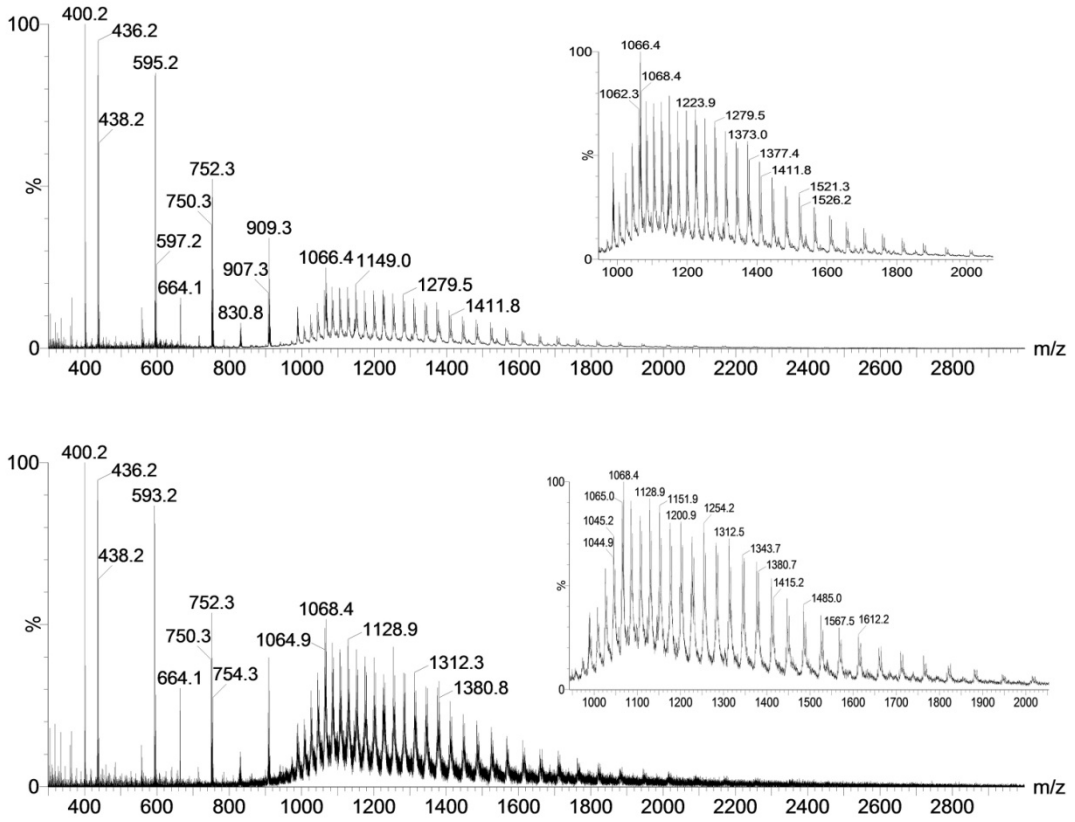


Figure 6.10. Raw ESI mass spectra of as-isolated E268A AMSDH (top) and E268A treated with 2-HMS (bottom). Narrow range scans corresponding to the proteins are shown in the insets.

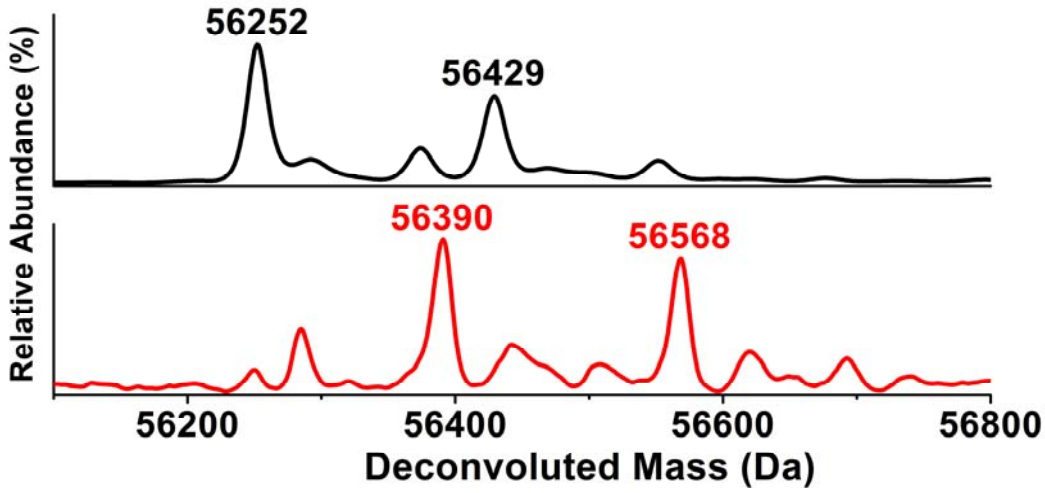


Figure 6.11. Deconvoluted positive-mode electrospray ionization (ESI) mass spectra of as-isolated E268A (top) and 2-HMS treated E268A (bottom).

We determined the crystal structure of E268A and refined it to 2.00 Å resolution. The overall structure aligns very well with the wild-type binary complex structure with a RMSD of 0.139 Å. The active site of E268A also resembles the native AMSDH structure (Figure 6.12). The nature of the absorbing species was further investigated by soaking co-crystallized E268A-NAD^+ crystals in mother liquor containing 2-HMS. By doing so, two temporally, structurally, and spectroscopically distinct intermediates were identified.

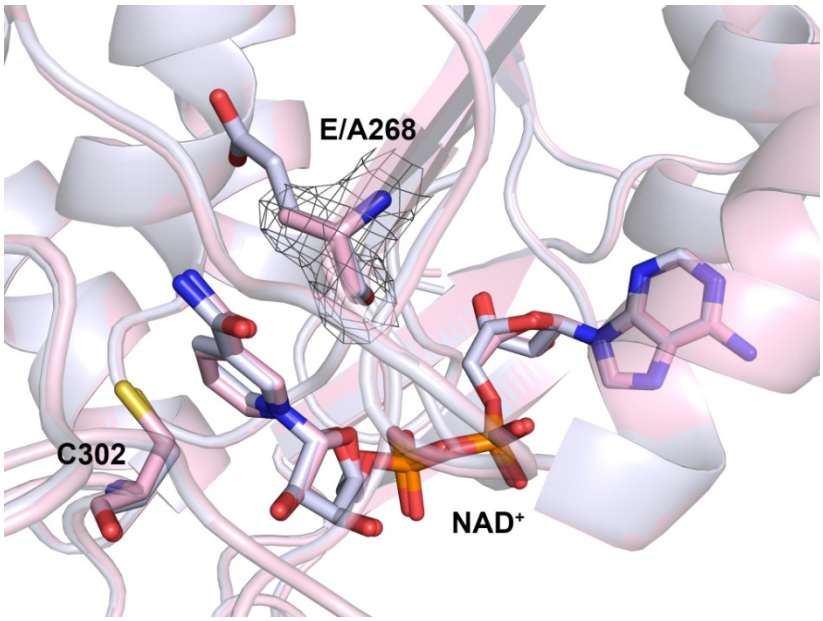


Figure 6.12. Superimposition of E268A (pink) active site with wtAMSDH (light blue).

When E268A-NAD^+ crystals are soaked with 2-HMS for 40 min or less, their single crystal electronic absorption spectra show an absorbance maximum at 422 nm (Figure 6.7B), as was observed in the solution-state titration and the stopped-flow assays. An individual electronic absorption spectrum for an E268A-NAD^+ crystal soaked with 2-HMS for 15 min can be found in Figure 6.9E. The structure of E268A-NAD^+ soaked with 2-HMS for 30 minutes before flash-cooling was solved and refined to 2.15 Å resolution. In this structure, a continuous electron density between Cys302-SG and 2-HMS-C6 is observed, similar to the thioacyl intermediate observed in the wild-type enzyme. However, in contrast to the thioacyl intermediate, the density

around C6 is not flat, indicating an sp^3- rather than sp^2-hybridized carbon (Figure 6.13A). More importantly, the C6 of 2-HMS and the C4N of NAD^+ are very close (2.4 - 2.8 Å), making it unlikely that the hydride has been transferred from the substrate. Taken together, these data allow us to assign this intermediate to a thiohemiacetal enzyme adduct (Figure 6.9B). A similar intermediate has only been trapped once previously in a crystal which contains no co-substrate (116). Hence, this is the first time for this intermediate to be trapped in the presence of NAD^+.

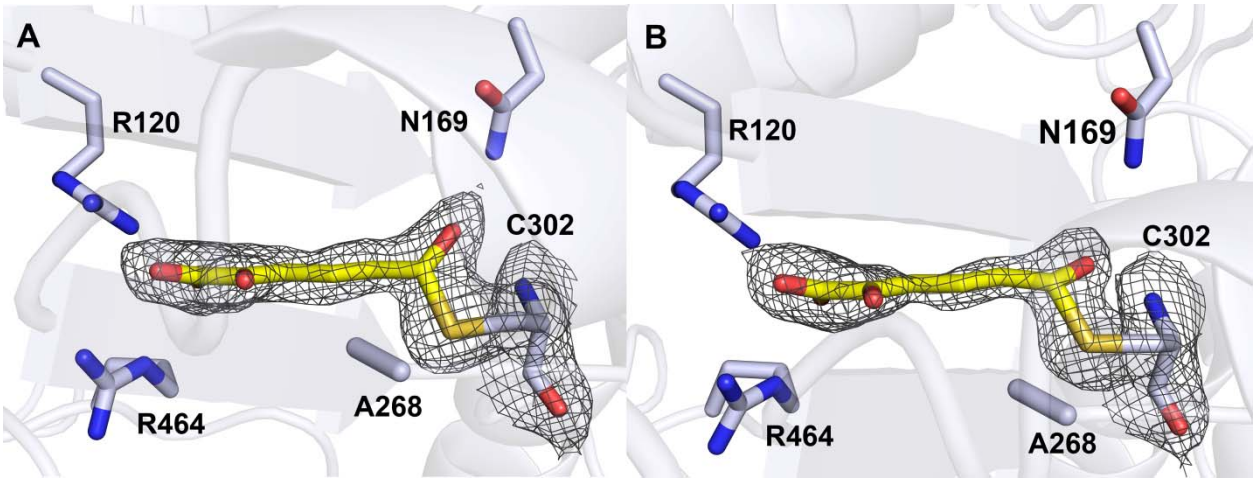


Figure 6.13. Crystal structures of two distinct catalytic intermediates. Electron density map of (A) the thiohemiacetal intermediate and (B) the thioacyl intermediate.

If the E268A-NAD^+ crystals are soaked with 2-HMS for longer than 1 hour, their single-crystal UV-Vis spectra begin to resemble that of the wild-type thioacyl intermediate with a corresponding absorbance maximum at 394 nm (Figure 6.7B), as seen in wild-type thioacyl intermediate crystals. An individual electronic absorption spectrum for an E268A-NAD^+ crystal soaked with 2-HMS for 120 min can be found in Figure 6.9E, bottom. The structure of an E268A-NAD^+ crystal soaked with 2-HMS for 180 minutes was solved and refined to 2.20 Å. The structure of this intermediate is also similar to the wild-type thioacyl enzyme adduct with NADH, rather than NAD^+ found at the active site. The distance between the C4N of NADH and C6 of 2-HMS is longer than 6.1 Å (Figure 6.9C). The electron density around C6 is flatter

(Figure 6.13B) compared with the thiohemiacetal intermediate and similar to the thioacyl intermediate trapped in the wild-type AMSDH structure (Figure 6.3F). Based on the similarities in their absorbance and structures, we conclude that this latter intermediate is equivalent to the wild-type thioacyl intermediate. It is also worth noting that the strictly conserved asparagine 169 is seen to stabilize both the thiohemiacetal and thioacyl intermediates.

6.5 Discussion

The substrate of AMSDH, 2-AMS, is an unstable aldehyde which can decay to picolinic acid and water, presumably through an electrocyclization reaction as its upstream intermediate does (193). So, to assay enzymatic activity, the upstream enzyme was utilized in the reaction mixture to generate substrate, and it was shown that AMSDH is catalytically active. Unfortunately, no kinetic parameters can be reliably determined because the concentration of 2-AMS is not well-defined in the coupled-enzyme assay. To circumvent this issue, a previously-identified, stable alternative substrate, 2-HMS (236,240), in which a hydroxyl group replaces the amino group in 2-AMS to prevent cyclization, was used to characterize the activity of AMSDH and to examine the activity of the mutants.

Substrate bound ternary complex structures were obtained by soaking co-crystallized protein and NAD^+ with 2-AMS or 2-HMS. 2-AMS is an unstable compound which decays with a $t_{1/2}$ of about 9 s at pH 7.4 and 37 °C or 35 s at pH 7.0 and 20 °C. Notably, this is its first reported structure. It appears to be stabilized in the enzyme active site by interactions with Arg120 and Arg464 so that the electrocyclization reaction cannot occur. Both arginine residues are close to the protein surface and in good positions to serve as gatekeepers, bringing the substrate into the active site. As a residue residing on a loop, Arg464 should be relatively flexible. The electron density for the side chain of Arg120 is partially missing in the binary complex structure but very

well resolved in both ternary complex structures. This observation indicates that the presence of substrate can stabilize what may be a flexible residue. It becomes evident from the coordinates that Arg120 and Arg464 play an important role in substrate recognition, stabilization, and possibly product release. Two arginine residues are rarely observed in such close proximity, stabilizing one end of the same molecule with multiple H-bonds. With the exception of the H-bonds provided by Arg120 and 464, the substrate-binding pocket is mostly composed of hydrophobic residues (Figure 6.14 B). Based on sequence alignment (Figure 6.14A), these two arginine residues are strictly conserved throughout the HMSDH family but are not found in other members of the ALDH superfamily. We propose that these dual arginines combined with the size restrictions provided by the hydrophobic pocket endow this enzyme with its specificity towards small α-substituted carboxylic acids with an aldehyde moiety, such as 2-AMS and 2-HMS.

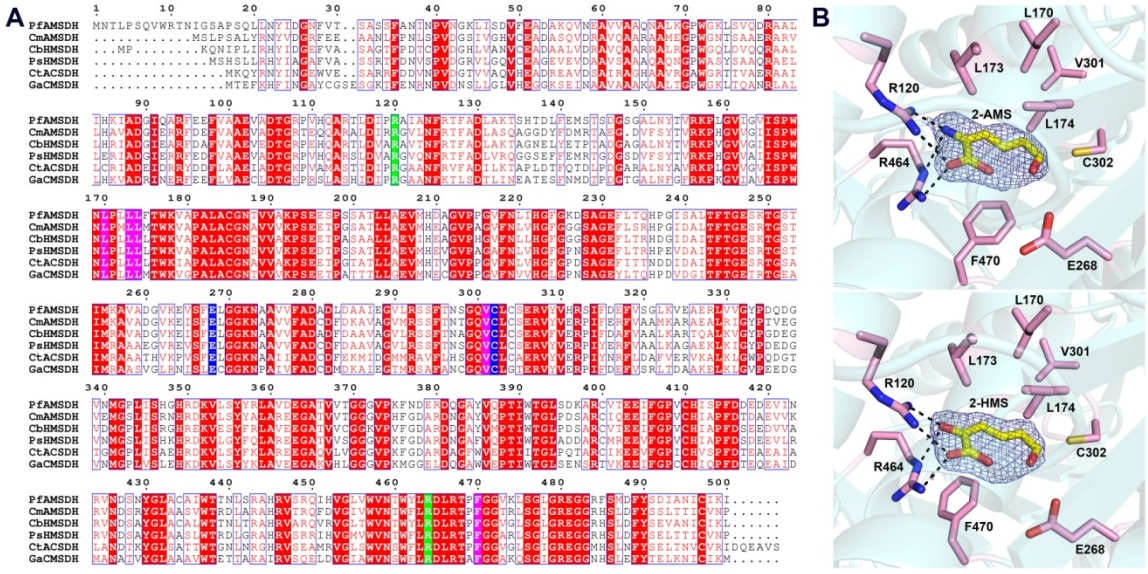


Figure 6.14. Identification of substrate binding residues for the hydroxymuconic semialdehyde dehydrogenase (HMSDH) family. (A) Sequence alignment of several enzymes from the HMSDH family. (B) Substrate binding pocked of 2-AMS, left, and 2-HMS, right, ternary complex crystal structures.

Two strictly conserved catalytic residues, Cys302 and Glu268, are present at the interior of the substrate-binding pocket. General features regarding these residues in the ALDH superfamily are (1) that the cysteine serves as a catalytic nucleophile which is anticipated to form a covalent-adduct intermediate with the substrate by a nucleophilic addition during catalysis (113,115,241), and (2) that the glutamate serves as a base to activate water for hydrolysis of the thioacyl-enzyme adduct (108,117,119,242). In the ternary complex structures, Cys302 is located at an ideal position to initiate catalysis. It is proposed to attack the aldehydic carbon (C6) of the substrate. In the two ternary complex structures, the distance between the sulfur of Cys302 and the C6 of substrate is ca. 3.3 Å. Cys302 and the aldehydic carbon form covalent bond in both thioacyl and thiohemiacetal intermediates. Mutation of Cys302 to serine led to completely inactive enzyme, further confirming its catalytic significance.

Examining the wild-type AMSDH structures shows that in the NAD^+-bound binary complex, Glu268 adopts a 'passive' conformation, pointing away from the substrate binding pocket, and forms H-bonds with both NE of Trp177 (3.2 Å distance) and the backbone oxygen of Phe470 (3.6 Å) to leave space for the reduction of NAD^+. Its electron density is very well resolved and the side chain B-factor is close to average: 28.2 $Å^2$/28.5 $Å^2$. The thiol moiety of Cys302 is 7.14 Å from Glu268 and is unlikely to form interactions. Interestingly, in both substrate bound structures, Glu268 becomes more flexible and exhibits much weaker electron density and increased side chain B-factors compared to average protein B-factors: 37.8 $Å^2$/28.5 $Å^2$ and 66.37 $Å^2$/39.7 $Å^2$. In the thioacyl intermediate structure, the electron density of Glu268 becomes very well defined again, but its side chain rotates 73° towards the bound substrate and seems to be in an 'active' position to abstract a proton from a deacylating water (Figure 6.3D). At this point in the reaction cycle, the NADH molecule needs to leave the active site to make room for the

catalytic water molecule. Movement of the nicotinamide ring of NAD^+ coupled with the rotation of an active site glutamate has previously been observed in other aldehyde dehydrogenases during catalysis (108,117,119).

Mutation of Glu268 to alanine led to the accumulation of the thiohemiacetal intermediate in both solution and crystal structure. The strictly conserved glutamate residue in the active site of ALDH enzymes has been proposed to play up to three possible roles during catalysis. It is strictly required to activate the deacylating water which allows for product release, it is in a 'passive' conformation during $NAD(P)^+$ reduction, and in some cases, it may serve to activate cysteine for nucleophilic attack (120). Based on these roles, mutation to alanine would be expected to decrease the rate of hydrolysis of the thioacyl adduct, have no effect on the rate of reduction of NADH, and possibly decrease the rate of nucleophilic attack by cysteine. Based on these understandings, deletion of the active site glutamate should cause an accumulation of the thioacyl intermediate. However, in this work the E268A mutant is shown to accumulate the preceding thiohemiacetal intermediate both in crystal and in solution. This finding suggests an additional catalytic role for this residue: rotation of Glu268 towards the active site facilitates the hydride transfer from the tetrahedral thiohemiacetal adduct to NAD^+. The rapid formation of the intermediate in solution indicates that Glu268 of AMSDH does not play a role in activating cysteine. However, it does appear necessary to complete hydride transfer from the substrate to NAD^+, and its removal turns the native substrate into a suicide inhibitor.

Based on previous studies of the ALDH mechanism and the eight high-resolution crystal structures solved, we propose a catalytic mechanism for AMSDH. As shown in Figure 6.15, NAD^+ binds to the enzyme, **1**, to form an NAD^+-bound AMSDH complex, **2**. The substrate, 2-AMS, is then recognized by Arg120 and Arg464 through multiple hydrogen-bonding

interactions, **3**. At this point, it is not yet clear whether the *E*/*Z* isomerization of the 2,3 double bond takes place before, during, or after the nucleophilic attack by Cys302 on the aldehydic carbon, which produces the tetrahedral thiohemiacetal intermediate, **4**. The isomerization drives a rotation of the substrate at the C1 end so that only the carboxylate group interacts with the two arginine residues. Next, NAD^+ is reduced to NADH by abstraction of a hydride from **4**, forming a thioacyl intermediate, **5**. Upon reduction, the nicotinamide portion of NADH moves away from the substrate as Glu268 rotates into position to activate a water molecule to perform a nucleophilic attack on the same carbon that was previously attacked by Cys302, forming a second tetrahedral intermediate, **6**. Finally, the second tetrahedral intermediate collapses, breaking the C-S bond and releasing the final products, 2-AM and NADH. In the present work, species **1** – **5** are spectroscopically and structurally characterized while intermediate **6** was not seen to accumulate.

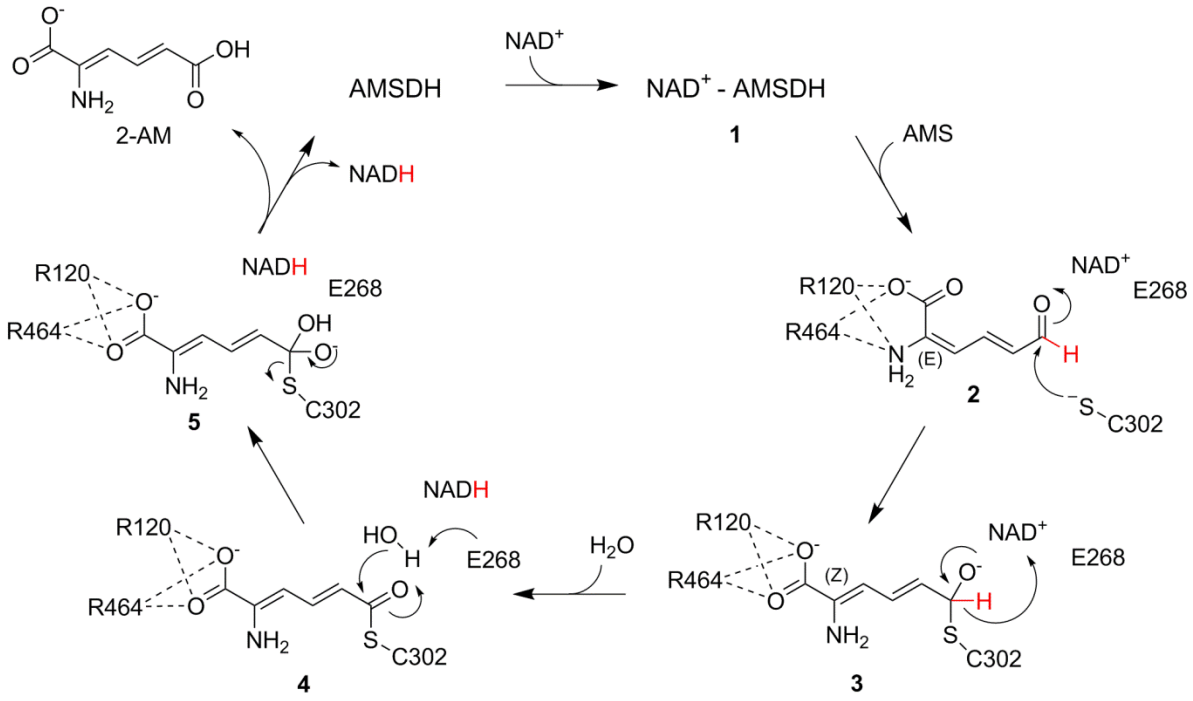


Figure 6.15. Proposed catalytic mechanism for the oxidation of 2-AMS by AMSDH.

In this work, five catalytically relevant structures of the wild-type AMSDH and three mutant structures yield a comprehensive understanding of the protein's overall structure, co-substrate binding mode, and elucidate the primary reason for substrate specificity among the HMSDH family of the ALDH superfamily. The structural and spectroscopic snapshots capture the crystal structure of a highly unstable kynurenine metabolite, 2-AMS, and two catalytic intermediates, including stabilizing a tetrahedral intermediate in a mutant protein, which was further verified by mass spectrometry. Another interesting finding revealed through solving the ternary complex and intermediate crystal structures is that an *E* to *Z* isomerization of the substrate occurs in this dehydrogenase before hydride transfer. This is the first piece of structural evidence illustrating an aldehyde dehydrogenase which proceeds via an *E*/*Z* isomerization on its substrate during catalysis.

Table 6.1. X-ray crystallography data collection and refinement statistics.

Data collection	Apo-AMSDH	NAD-AMSDH	NAD-2-AMS-AMSDH	NAD-2-HMS-AMSDH	WT-Thioacyl	E268A-AMSDH	E268A-Hemiacetal	E268A-Thioacyl
detector type	MAR300 CCD	MAR225 CCD	MAR300 CCD	MAR300 CCD	MAR225 CCD	MAR225 CCD	MAR225 CCD	MAR225 CCD
source	APS 22-ID	APS 22-BM	APS 22-ID	APS 22-ID	APS 22-BM	APS 22-BM	APS 22-BM	APS, Sector 22-BM
space group	$P2_22_12_1$	$P2_22_12_1$	$P2_22_12_1$	$P2_22_12_1$	$P2_22_12_1$	$P2_22_12_1$	$P2_22_12_1$	$P2_22_12_1$
unit cell lengths (Å)	a=88.27 b=141.89 c=172.92	a=88.58 b=142.00 c=174.38	a=88.40 b=142.12 c=174.41	a=88.57 b=142.72 c=175.01	a=88.36 b=141.75 c=174.37	a=88.53 b=141.98 c=173.80	a=88.57 b=141.56 c=174.63	a=88.33 b=141.35 c=173.53
unit cell angles (˚)	α=β=γ=90	α=β=γ=90	α=β=γ=90	α=β=γ=90	α=β=γ=90	α=β=γ=90	α=β=γ=90	α=β=γ=90
wavelength (Å)	0.8	1.0	0.8	0.8	1.0	1.0	1.0	1.0
temperature (K)	100	100	100	100	100	100	100	100
resolution (Å) [a]	45.00-2.20 (2.24-2.20)	35.00-2.00 (2.07-2.00)	35.00-1.98 (2.03-1.98)	45.00-2.15 (2.19-2.15)	45.00-1.95 (1.98-1.95)	50.00-2.00 (2.03-2.00)	50.00-2.15 (2.15-2.19)	50.00-2.20 (2.24-2.20)
completeness (%)[a]	99.8 (99.4)	99.8 (98.9)	95.2 (99.7)	99.9 (100.0)	94.9 (88.8)	97.2 (97.4)	99.9 (100)	99.6 (99.2)
R_{merge} (%)[a, b]	8.0 (53.5)	11.2 (8.9)	10.7 (78.3)	12.2 (84.1)	10.4 (84.3)	10.8 (74.5)	14.8 (58.7)	9.9 (71.7)
$I/\sigma I$ [a]	45.2 (4.4)	27.3 (2.3)	22.7 (2.2)	27.1 (2.7)	28.4 (2.2)	13.3 (3.2)	17.7 (3.9)	36.9 (3.7)
multiplicity [a]	14.0 (11.6)	13.7 (9.1)	9.8(7.8)	12.5(10.3)	10.5(6.4)	11.9(10.5)	8.8(7.6)	11.4(10.3)
Refinement								
resolution (Å)	2.20	2.00	2.00	2.15	1.95	2.00	2.15	2.20
no. reflections; working/test	109619/ 5469	149298/ 7485	140017/ 7401	114123/ 6029	145080/ 7649	144988/ 7303	119216/ 5973	110312/ 5518
R_{work} (%)[c]	18.6	17.2	16.2	18.0	17.6	19.5	19.2	18.4
R_{free} (%)[d]	23.9	21.8	20.8	23.7	21.6	23.7	23.5	22.7
no. of protein atoms	14651	14684	14684	14684	14684	14668	14668	14668
no. of ligand atoms	23	188	220	220	220	204	220	220
no. of solvent sites	702	1682	1397	802	1702	1188	1136	882
Average B-factor (Å^2)								
protein	41.1	28.5	28.5	39.7	27.9	31.7	29.2	36.2
NAD^+	N/A	28.8	40.1	56.7	37.5	37.0	41.4	56.4
NA^+	35.3	N/A	32.3	41.6	30.7	34.0	43.6	43.9
EG/GOL/AMS/HMS	38.4	42.7	32.4	45.6	23.0	42.1	25.1	28.1
solvent	42.7	37.3	37.5	43.3	35.4	37.4	34.9	39.2
PDB code	4I26	4I1W	4I25	4I2R	4NPI	4OE2	4OU2	4OUB

[a] Values in parentheses are for the highest resolution shell.
[b] Rmerge = Σi $|I_{hkl,i} - \langle I_{hkl} \rangle| / \Sigma_{hkl} \Sigma i\, I_{hkl,i}$, where $I_{hkl,i}$ is the observed intensity and $\langle I_{hkl} \rangle$ is the average intensity of multiple measurements.
[c] Rwork = $\Sigma ||F_o| - |F_c|| / \Sigma |F_o|$, where $|F_o|$ is the observed structure factor amplitude, and $|F_c|$ is the calculated structure factor amplitude.
[d] R_{free} is the R factor based on 5% of the data excluded from refinement.
[e] Based on values attained from refinement validation options in COOT.

CHAPTER 7 SUMMARY

Work in this dissertation focuses on two enzymes in the kynurenine pathway and 2-NBA biodegradation pathway: ACMSD and AMSDH. Substrates for both of the enzymes are unstable metabolic intermediates and spontaneously decay to QUIN and PIC, respectively. Structure and mechanistic studies on these two enzymes may shed light on the mechanism for regulation of side product formation in both pathways.

ACMSD is the first transition metal dependent, oxygen independent decarboxylase. The unique cofactor dependency requires a novel catalytic mechanism. Hence, a new working model has been proposed for ACMSD based on the findings described in the earlier chapters, in which a metal bound water ligand activated by His228 residue acts as a nucleophile to initiate the reaction. Investigation of the His228 residue reveals two roles: tuning metal selectivity and interacting with the metal bound water molecule. Anion inhibition experiments further demonstrated that replacing of the metal bound water ligand by bromide leads to the loss of enzyme activity. These results are consistent with the proposed ACMSD mechanism.

Since the substrate for ACMSD, ACMS, is an unstable aldehyde compound, the question of how ACMS is stabilized during catalysis is worth study. In the crystal structure of homodimer *Pf*ACMSD, we found two conserved arginine residues in the active site are in good position to bind ACMS by forming hydrogen bonding interactions with its carboxyl groups. Interestingly, one of the arginines is inserted from the neighboring subunit. Mutation of either of the arginine residues leads to the loss of enzyme activity. Simply mixing the two inactive arginine mutants can partially rescue ACMSD activity by forming hybridized heterodimers, indicating the dimeric state is the active form of ACMSD. A membrane based experiment demonstrated that these two arginine residues are involved in substrate binding. In the crystal structures of *h*ACMSD we

solved later, two arginine residues are also observed in the active site. Both of the arginines form hydrogen bonds with a competitive inhibitor in the co-crystallized structure. Hence, ACMSD harnesses its unstable substrate by stabilizing it through two active site arginine residues, one of which is contributed by a neighboring subunit, and thus, the dimeric state is required for enzyme activity. The dimer dissociation constant calculated based on our experiments is in the range for "weak" transient complexes that show a dynamic mixture of both monomer and dimer states *in vivo*. Hence, ACMSD activity could be tuned by local protein concentration and physiological environmental factors such as pH and local protein or ligand concentrations *in vivo*.

Unlike ACMSD, AMSDH has not been studied at molecular level before our work. We have cloned and expressed *Pf*AMSDH in *E.coli* and isolated it as soluble and active enzyme. We have crystallized *Pf*ACMSD in both its apo form and co-substrate NAD^+ bound binary complex form. Substrates 2-AMS and 2-HMS bound ternary complex structures are also obtained by substrate soaking and reveal two arginine residues that play a major role in substrate binding. Mutagenesis and sequence alignment further indicate that these two arginine residues are responsible for substrate selectivity in the whole HMSDH family. We have also trapped a covalent bound thioacyl intermediate in WT-AMSDH, in which the substrate molecule tautomerized to a different conformation, suggesting an innate isomerase activity of AMSDH. Mutation of a conserved active site glutamate leads to the accumulation of a new intermediate. Crystal structure of this intermediate reveals a thiohemiacetal tetrahedral covalent adduct, which is then further verified by single crystal spectroscopy and MS spectrometry. Hence, it becomes clear that this glutamate residue facilitates hydride transfer during catalysis. Based on the crystal structures we have solved and intermediates we have trapped, a better understanding of AMSDH is achieved and a working model for its mechanism of action is proposed.

REFERENCES

1. Van der Goot, A. T., and Nollen, E. A. (2013) Tryptophan metabolism: entering the field of aging and age-related pathologies. *Trends Mol. Med.* **19**, 336-344
2. Stone, T. W., and Darlington, L. G. (2002) Endogenous kynurenines as targets for drug discovery and development. *Nat. Rev. Drug Discov.* **1**, 609-620.
3. Stone, T. W., Behan, W. M., Jones, P. A., Darlington, L. G., and Smith, R. A. (2001) The role of kynurenines in the production of neuronal death, and the neuroprotective effect of purines. *J. Alzheimers. Dis.* **3**, 355-366
4. Baran, H., Jellinger, K., and Deecke, L. (1999) Kynurenine metabolism in Alzheimer's disease. *J. Neural. Transm.* **106**, 165-181
5. Schwarcz, R., Whetsell, W. O., Jr., and Mangano, R. M. (1983) Quinolinic acid: an endogenous metabolite that produces axon-sparing lesions in rat brain. *Science* **219**, 316-318
6. Beal, M. F., Ferrante, R. J., Swartz, K. J., and Kowall, N. W. (1991) Chronic quinolinic acid lesions in rats closely resemble Huntington's disease. *J. Neurosci.* **11**, 1649-1659
7. Stone, T. W., Mackay, G. M., Forrest, C. M., Clark, C. J., and Darlington, L. G. (2003) Tryptophan metabolites and brain disorders. *Clin. Chem. Lab. Med.* **41**, 852-859
8. Schwarcz, R., Guidetti, P., Sathyasaikumar, K. V., and Muchowski, P. J. (2010) Of mice, rats and men: revisiting the quinolinic acid hypothesis of Huntington's disease. *Prog. Neurobiol.* **90**, 230-245
9. Lahdou, I., Sadeghi, M., Oweira, H., Fusch, G., Daniel, V., Mehrabi, A., Jung, G., Elhadedy, H., Schmidt, J., Sandra-Petrescu, F., Iancu, M., Opelz, G., Terness, P., and

Schefold, J. C. (2013) Increased serum levels of quinolinic acid indicate enhanced severity of hepatic dysfunction in patients with liver cirrhosis. *Hum. Immunol.* **74**, 60-66

10. Li, T., Ma, J., Hosler, J. P., Davidson, V. L., and Liu, A. (2007) Detection of transient intermediates in the metal-dependent non-oxidative decarboxylation catalyzed by α-amino-β-carboxymuconic-ε-semialdehyde decarboxylase. *J. Am. Chem. Soc.* **129**, 9278-9279
11. Hasegawa, Y., Muraki, T., Tokuyama, T., Iwaki, H., Tatsuno, M., and Lau, P. C. (2000) A novel degradative pathway of 2-nitrobenzoate via 3-hydroxyanthranilate in Pseudomonas fluorescens strain KU-7. *FEMS Microbiol. Lett.* **190**, 185-190
12. Muraki, T., Taki, M., Hasegawa, Y., Iwaki, H., and Lau, P. C. (2003) Prokaryotic homologs of the eukaryotic 3-hydroxyanthranilate 3,4-dioxygenase and 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase in the 2-nitrobenzoate degradation pathway of Pseudomonas fluorescens strain KU-7. *Appl. & Environ. Microbiol.* **69**, 1564-1572
13. Fukuoka, S., Ishiguro, K., Yanagihara, K., Tanabe, A., Egashira, Y., Sanada, H., and Shibata, K. (2002) Identification and expression of a cDNA encoding human α-amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD). A key enzyme for the tryptophan-niacine pathway and "quinolinate hypothesis". *J. Biol. Chem.* **277**, 35162-35167
14. Li, T., Walker, A. L., Iwaki, H., Hasegawa, Y., and Liu, A. (2005) Kinetic and spectroscopic characterization of ACMSD from *Pseudomonas fluorescens* reveals a pentacoordinate mononuclear metallocofactor. *J. Am. Chem. Soc.* **127**, 12282-12290

15. Li, T. F., Iwaki, H., Fu, R., Hasegawa, Y., Zhang, H., and Liu, A. M. (2006) α-Amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD) is a new member of the amidohydrolase superfamily. *Biochemistry* **45**, 6628-6634

16. Holm, L., and Sander, C. (1997) An evolutionary treasure: unification of a broad set of amidohydrolases related to urease. *Proteins* **28**, 72-82

17. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic Acids Res.* **42**, D222-230

18. Liu, A., and Zhang, H. (2006) Transition metal-catalyzed nonoxidative decarboxylation reactions. *Biochemistry* **45**, 10407-10411

19. Seibert, C. M., and Raushel, F. M. (2005) Structural and catalytic diversity within the amidohydrolase superfamily. *Biochemistry* **44**, 6383-6391

20. Sterner, R., and Hocker, B. (2005) Catalytic versatility, stability, and evolution of the (betaalpha)8-barrel enzyme fold. *Chem. Rev.* **105**, 4038-4055

21. Goto, M., Hayashi, H., Miyahara, I., Hirotsu, K., Yoshida, M., and Oikawa, T. (2006) Crystal structures of nonoxidative zinc-dependent 2,6-dihydroxybenzoate (γ-resorcylate) decarboxylase from *Rhizobium* sp. strain MTP-10005. *J. Biol. Chem.* **281**, 34365-34373

22. Alhapel, A., Darley, D. J., Wagener, N., Eckel, E., Elsner, N., and Pierik, A. J. (2006) Molecular and functional analysis of nicotinate catabolism in Eubacterium barkeri. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12341-12346

23. Buchan, D. W., Ward, S. M., Lobley, A. E., Nugent, T. C., Bryson, K., and Jones, D. T. (2010) Protein annotation and modelling servers at University College London. *Nucleic Acids Res.* **38**, W563-568

24. Seffernick, J. L., de Souza, M. L., Sadowsky, M. J., and Wackett, L. P. (2001) Melamine deaminase and atrazine chlorohydrolase: 98 percent identical but functionally different. *J. Bacteriol.* **183**, 2405-2410

25. Huo, L., Fielding, A. J., Chen, Y., Li, T., Iwaki, H., Hosler, J. P., Chen, L., Hasegawa, Y., Que, L., Jr., and Liu, A. (2012) Evidence for a Dual Role of an Active Site Histidine in α-amino-β-carboxymuconic-ε-semialdehyde decarboxylase. *Biochemistry* **51**, 5811-5821

26. Huo, L., Davis, I., Chen, L., and Liu, A. (2013) The power of two: arginine 51 and arginine 239* from a neighboring subunit are essential for catalysis in α-amino-β-carboxymuconic-ε-semialdehyde decarboxylase. *J. Biol. Chem.* **288**, 30862-30871

27. Hagelueken, G., Huang, H., Mainprize, I. L., Whitfield, C., and Naismith, J. H. (2009) Crystal structures of Wzb of Escherichia coli and CpsB of Streptococcus pneumoniae, representatives of two families of tyrosine phosphatases that regulate capsule assembly. *J. Mol. Biol.* **392**, 678-688

28. Kim, H. S., Lee, S. J., Yoon, H. J., An, D. R., Kim do, J., Kim, S. J., and Suh, S. W. (2011) Crystal structures of YwqE from Bacillus subtilis and CpsB from Streptococcus pneumoniae, unique metal-dependent tyrosine phosphatases. *J. Struct. Biol.* **175**, 442-450

29. Ghodge, S. V., Fedorov, A. A., Fedorov, E. V., Hillerich, B., Seidel, R., Almo, S. C., and Raushel, F. M. (2013) Structural and mechanistic characterization of L-histidinol phosphate phosphatase from the polymerase and histidinol phosphatase family of proteins. *Biochemistry* **52**, 1101-1112

30. Hobbs, M. E., Malashkevich, V., Williams, H. J., Xu, C., Sauder, J. M., Burley, S. K., Almo, S. C., and Raushel, F. M. (2012) Structure and catalytic mechanism of LigI:

insight into the amidohydrolase enzymes of cog3618 and lignin degradation. *Biochemistry* **51**, 3497-3507

31. Hobbs, M. E., Vetting, M., Williams, H. J., Narindoshvili, T., Kebodeaux, D. M., Hillerich, B., Seidel, R. D., Almo, S. C., and Raushel, F. M. (2013) Discovery of an L-fucono-1,5-lactonase from cog3618 of the amidohydrolase superfamily. *Biochemistry* **52**, 239-253

32. Ochoa, S. (1951) Biological mechanisms of carboxylation and decarboxylation. *Physiol. Rev.* **31**, 56-106

33. Schorken, U., and Sprenger, G. A. (1998) Thiamin-dependent enzymes as catalysts in chemoenzymatic syntheses. *Biochim. Biophys. Acta* **1385**, 229-243

34. Prescott, A. G., and Lloyd, M. D. (2000) The iron(II) and 2-oxoacid-dependent dioxygenases and their role in metabolism. *Nat. Prod. Rep.* **17**, 367-383

35. Hausinger, R. P. (2004) FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39**, 21-68

36. Kivirikko, K. I., and Pihlajaniemi, T. (1998) Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases. *Adv. Enzymol. Relat. Areas. Mol. Biol.* **72**, 325-398

37. Lloyd, M. D., Lee, H. J., Harlos, K., Zhang, Z. H., Baldwin, J. E., Schofield, C. J., Charnock, J. M., Garner, C. D., Hara, T., Terwisscha van Scheltinga, A. C., Valegard, K., Viklund, J. A., Hajdu, J., Andersson, I., Danielsson, A., and Bhikhabhai, R. (1999) Studies on the active site of deacetoxycephalosporin C synthase. *J. Mol. Biol.* **287**, 943-960

38. Yin, X., O'Hare, T., Gould, S. J., and Zabriskie, T. M. (2003) Identification and cloning of genes encoding viomycin biosynthesis from Streptomyces vinaceus and evidence for involvement of a rare oxygenase. *Gene* **312**, 215-224

39. Schofield, C. J., and Ratcliffe, P. J. (2004) Oxygen sensing by HIF hydroxylases. *Nat. Rev. Mol. Cell. Biol.* **5**, 343-354

40. Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T., and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 16660-16665

41. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T., and Sedgwick, B. (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature* **419**, 174-178

42. Tsukada, Y., and Zhang, Y. (2006) Purification of histone demethylases from HeLa cells. *Methods* **40**, 318-326

43. Hegg, E. L., and Que, L., Jr. (1997) The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **250**, 625-629

44. Krebs, C., Galonic Fujimori, D., Walsh, C. T., and Bollinger, J. M., Jr. (2007) Non-heme Fe(IV)-oxo intermediates. *Acc. Chem. Res.* **40**, 484-492

45. Price, J. C., Barr, E. W., Glass, T. E., Krebs, C., and Bollinger, J. M., Jr. (2003) Evidence for hydrogen abstraction from C1 of taurine by the high-spin Fe(IV) intermediate detected during oxygen activation by taurine:alpha-ketoglutarate dioxygenase (TauD). *J. Am. Chem. Soc.* **125**, 13008-13009

46. Costas, M., Mehn, M. P., Jensen, M. P., and Que, L., Jr. (2004) Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem. Rev.* **104**, 939-986

47. Neidig, M. L., Decker, A., Choroba, O. W., Huang, F., Kavana, M., Moran, G. R., Spencer, J. B., and Solomon, E. I. (2006) Spectroscopic and electronic structure studies of aromatic electrophilic attack and hydrogen-atom abstraction by non-heme iron enzymes. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 12966-12973

48. Hoffart, L. M., Barr, E. W., Guyer, R. B., Bollinger, J. M., Jr., and Krebs, C. (2006) Direct spectroscopic detection of a C-H-cleaving high-spin Fe(IV) complex in a prolyl-4-hydroxylase. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 14738-14743

49. Solomon, E. I., Brunold, T. C., Davis, M. I., Kemsley, J. N., Lee, S. K., Lehnert, N., Neese, F., Skulan, A. J., Yang, Y. S., and Zhou, J. (2000) Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **100**, 235-350

50. Proshlyakov, D. A., Henshaw, T. F., Monterosso, G. R., Ryle, M. J., and Hausinger, R. P. (2004) Direct detection of oxygen intermediates in the non-heme Fe enzyme taurine/alpha-ketoglutarate dioxygenase. *J. Am. Chem. Soc.* **126**, 1022-1023

51. Grzyska, P. K., Appelman, E. H., Hausinger, R. P., and Proshlyakov, D. A. (2010) Insight into the mechanism of an iron dioxygenase by resolution of steps following the FeIV=HO species. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 3982-3987

52. Roach, P. L., Clifton, I. J., Hensgens, C. M., Shibata, N., Schofield, C. J., Hajdu, J., and Baldwin, J. E. (1997) Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830

53. Elkins, J. M., Ryle, M. J., Clifton, I. J., Dunning Hotopp, J. C., Lloyd, J. S., Burzlaff, N. I., Baldwin, J. E., Hausinger, R. P., and Roach, P. L. (2002) X-ray crystal structure of Escherichia coli taurine/alpha-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **41**, 5185-5192

54. Ryle, M. J., Liu, A., Muthukumaran, R. B., Ho, R. Y., Koehntop, K. D., McCracken, J., Que, L., Jr., and Hausinger, R. P. (2003) O2- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42**, 1854-1862

55. Price, J. C., Barr, E. W., Hoffart, L. M., Krebs, C., and Bollinger, J. M., Jr. (2005) Kinetic dissection of the catalytic mechanism of taurine:alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **44**, 8138-8147

56. Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr., and Krebs, C. (2003) The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from Escherichia coli. *Biochemistry* **42**, 7497-7508

57. Ryle, M. J., Koehntop, K. D., Liu, A., Que, L., Jr., and Hausinger, R. P. (2003) Interconversion of two oxidized forms of taurine/alpha-ketoglutarate dioxygenase, a non-heme iron hydroxylase: evidence for bicarbonate binding. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 3790-3795

58. Lee, H. J., Lloyd, M. D., Harlos, K., Clifton, I. J., Baldwin, J. E., and Schofield, C. J. (2001) Kinetic and crystallographic studies on deacetoxycephalosporin C synthase (DAOCS). *J. Mol. Biol.* **308**, 937-948

59. Fitzpatrick, P. F. (2003) Mechanism of aromatic amino acid hydroxylation. *Biochemistry* **42**, 14083-14091

60. Moran, G. R. (2005) 4-Hydroxyphenylpyruvate dioxygenase. *Arch. Biochem. Biophys.* **433**, 117-128

61. Stanley, T. M., Johnson, W. H., Jr., Burks, E. A., Whitman, C. P., Hwang, C. C., and Cook, P. F. (2000) Expression and stereochemical and isotope effect studies of active 4-oxalocrotonate decarboxylase. *Biochemistry* **39**, 3514

62. Harayama, S., Rekik, M., Ngai, K. L., and Ornston, L. N. (1989) Physically associated enzymes produce and metabolize 2-hydroxy-2,4-dienoate, a chemically unstable intermediate formed in catechol metabolism via meta cleavage in Pseudomonas putida. *J. Bacteriol.* **171**, 6251-6258

63. Tanner, A., Bowater, L., Fairhurst, S. A., and Bornemann, S. (2001) Oxalate decarboxylase requires manganese and dioxygen for activity. Overexpression and characterization of Bacillus subtilis YvrK and YoaN. *J. Biol. Chem.* **276**, 43627-43634

64. Svedruzic, D., Jonsson, S., Toyota, C. G., Reinhardt, L. A., Ricagno, S., Lindqvist, Y., and Richards, N. G. (2005) The enzymes of oxalate metabolism: unexpected structures and mechanisms. *Arch. Biochem. Biophys.* **433**, 176-192

65. Opaleye, O., Rose, R. S., Whittaker, M. M., Woo, E. J., Whittaker, J. W., and Pickersgill, R. W. (2006) Structural and spectroscopic studies shed light on the mechanism of oxalate oxidase. *J. Biol. Chem.* **281**, 6428-6433

66. Just, V. J., Burrell, M. R., Bowater, L., McRobbie, I., Stevenson, C. E., Lawson, D. M., and Bornemann, S. (2007) The identity of the active site of oxalate decarboxylase and the importance of the stability of active-site lid conformations. *Biochem. J.* **407**, 397-406

67. Burrell, M. R., Just, V. J., Bowater, L., Fairhurst, S. A., Requena, L., Lawson, D. M., and Bornemann, S. (2007) Oxalate decarboxylase and oxalate oxidase activities can be interchanged with a specificity switch of up to 282,000 by mutating an active site lid. *Biochemistry* **46**, 12327-12336

68. Reinhardt, L. A., Svedruzic, D., Chang, C. H., Cleland, W. W., and Richards, N. G. (2003) Heavy atom isotope effects on the reaction catalyzed by the oxalate decarboxylase from Bacillus subtilis. *J. Am. Chem. Soc.* **125**, 1244-1252

69. Anand, R., Dorrestein, P. C., Kinsland, C., Begley, T. P., and Ealick, S. E. (2002) Structure of oxalate decarboxylase from Bacillus subtilis at 1.75 A resolution. *Biochemistry* **41**, 7659-7669

70. Angerhofer, A., Moomaw, E. W., Garcia-Rubio, I., Ozarowski, A., Krzystek, J., Weber, R. T., and Richards, N. G. (2007) Multifrequency EPR studies on the Mn(II) centers of oxalate decarboxylase. *J. Phys. Chem. B.* **111**, 5043-5046

71. Imaram, W., Saylor, B. T., Centonze, C. P., Richards, N. G., and Angerhofer, A. (2011) EPR spin trapping of an oxalate-derived free radical in the oxalate decarboxylase reaction. *Free Radic. Biol. Med.* **50**, 1009-1015

72. Waldrop, G. L., Braxton, B. F., Urbauer, J. L., Cleland, W. W., and Kiick, D. M. (1994) Secondary 18O and primary 13C isotope effects as a probe of transition-state structure for enzymatic decarboxylation of oxalacetate. *Biochemistry* **33**, 5262-5267

73. Vennesland, B. (1949) The beta-carboxylases of plants; the distribution of oxalacetic carboxylase in plant tissues. *J. Biol. Chem.* **178**, 591-597

74. Speck, J. F. (1949) The effect of cations on the decarboxylation of oxalacetic acid. *J. Biol. Chem.* **178**, 315-324

75. Wise, E., Yew, W. S., Babbitt, P. C., Gerlt, J. A., and Rayment, I. (2002) Homologous (beta/alpha)8-barrel enzymes that catalyze unrelated reactions: orotidine 5'-monophosphate decarboxylase and 3-keto-L-gulonate 6-phosphate decarboxylase. *Biochemistry* **41**, 3861-3869

76. Wise, E. L., Yew, W. S., Gerlt, J. A., and Rayment, I. (2003) Structural evidence for a 1,2-enediolate intermediate in the reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase, a member of the orotidine 5'-monophosphate decarboxylase suprafamily. *Biochemistry* **42**, 12133-12142

77. Yew, W. S., Wise, E. L., Rayment, I., and Gerlt, J. A. (2004) Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: mechanistic evidence for a proton relay system in the active site of 3-keto-L-gulonate 6-phosphate decarboxylase. *Biochemistry* **43**, 6427-6437

78. Wise, E. L., Yew, W. S., Akana, J., Gerlt, J. A., and Rayment, I. (2005) Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: structural basis for catalytic promiscuity in wild-type and designed mutants of 3-keto-L-gulonate 6-phosphate decarboxylase. *Biochemistry* **44**, 1816-1823

79. Yew, W. S., Akana, J., Wise, E. L., Rayment, I., and Gerlt, J. A. (2005) Evolution of enzymatic activities in the orotidine 5'-monophosphate decarboxylase suprafamily: enhancing the promiscuous D-arabino-hex-3-ulose 6-phosphate synthase reaction catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase. *Biochemistry* **44**, 1807-1815

80. Martynowski, D., Eyobo, Y., Li, T., Yang, K., Liu, A., and Zhang, H. (2006) Crystal structure of α-Amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD):

insight into the active site and catalytic mechanism of a novel decarboxylation reaction. *Biochemistry* **45**, 10412-10421

81. Liu, A., Li, T., and Fu, R. (2007) *Amidohydrolase Superfamily*, John Wiley & Sons Ltd.
82. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies. *Curr. Opin. Chem. Biol.* **10**, 492-497
83. Bergy, J. M., Tymoczko, J. L., and Stryer, L. (2002). *Biochemistry, 5th ed.* **W. H. Freeman and Company.**
84. Racker, E. (1949) Aldehyde dehydrogenase, a diphosphopyridine nucleotide-linked enzyme. *J. Biol. Chem.* **177**, 883-892
85. Hempel, J., von Bahr-Lindstrom, H., and Jornvall, H. (1984) Aldehyde dehydrogenase from human liver. Primary structure of the cytoplasmic isoenzyme. *Eur. J. Biochem.* **141**, 21-35
86. Perozich, J., Nicholas, H., Wang, B. C., Lindahl, R., and Hempel, J. (1999) Relationships within the aldehyde dehydrogenase extended family. *Protein Sci.* **8**, 137-146
87. Zinovieva, R. D., Tomarev, S. I., and Piatigorsky, J. (1993) Aldehyde dehydrogenase-derived omega-crystallins of squid and octopus. Specialization for lens expression. *J. Biol. Chem.* **268**, 11449-11455
88. Jester, J. V., Moller-Pedersen, T., Huang, J., Sax, C. M., Kays, W. T., Cavangh, H. D., Petroll, W. M., and Piatigorsky, J. (1999) The cellular basis of corneal transparency: evidence for 'corneal crystallins'. *J. Cell. Sci.* **112 (Pt 5)**, 613-622
89. Linneberg, A., Gonzalez-Quintela, A., Vidal, C., Jorgensen, T., Fenger, M., Hansen, T., Pedersen, O., and Husemoen, L. L. (2010) Genetic determinants of both ethanol and

acetaldehyde metabolism influence alcohol hypersensitivity and drinking behaviour among Scandinavians. *Clin. Exp. Allergy.* **40**, 123-130

90. Chen, C. H., Budas, G. R., Churchill, E. N., Disatnik, M. H., Hurley, T. D., and Mochly-Rosen, D. (2008) Activation of aldehyde dehydrogenase-2 reduces ischemic damage to the heart. *Science* **321**, 1493-1495

91. Higuchi, S., Matsushita, S., Imazeki, H., Kinoshita, T., Takagi, S., and Kono, H. (1994) Aldehyde dehydrogenase genotypes in Japanese alcoholics. *Lancet* **343**, 741-742

92. Crabb, D. W., Matsumoto, M., Chang, D., and You, M. (2004) Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. *Proc. Nutr. Soc.* **63**, 49-63

93. Mills, P. B., Struys, E., Jakobs, C., Plecko, B., Baxter, P., Baumgartner, M., Willemsen, M. A., Omran, H., Tacke, U., Uhlenberg, B., Weschke, B., and Clayton, P. T. (2006) Mutations in antiquitin in individuals with pyridoxine-dependent seizures. *Nat. Med.* **12**, 307-309

94. Cheung, A. M., Wan, T. S., Leung, J. C., Chan, L. Y., Huang, H., Kwong, Y. L., Liang, R., and Leung, A. Y. (2007) Aldehyde dehydrogenase activity in leukemic blasts defines a subgroup of acute myeloid leukemia with adverse prognosis and superior NOD/SCID engrafting potential. *Leukemia* **21**, 1423-1430

95. Liu, Z. J., Sun, Y. J., Rose, J., Chung, Y. J., Hsiao, C. D., Chang, W. R., Kuo, I., Perozich, J., Lindahl, R., Hempel, J., and Wang, B. C. (1997) The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nat. Struct. Biol.* **4**, 317-326

96. Perozich, J., Nicholas, H., Lindahl, R., and Hempel, J. (1999) The big book of aldehyde dehydrogenase sequences. An overview of the extended family. *Adv. Exp. Med. Biol.* **463**, 1-7

97. Rodriguez-Zavala, J. S., and Weiner, H. (2002) Structural aspects of aldehyde dehydrogenase that influence dimer-tetramer formation. *Biochemistry* **41**, 8229-8237

98. Fan, F., Lorenzen, J. A., and Plapp, B. V. (1991) An aspartate residue in yeast alcohol dehydrogenase I determines the specificity for coenzyme. *Biochemistry* **30**, 6397-6401

99. Chen, Z., Lee, W. R., and Chang, S. H. (1991) Role of aspartic acid 38 in the cofactor specificity of Drosophila alcohol dehydrogenase. *Eur. J. Biochem.* **202**, 263-267

100. Bernard, N., Johnsen, K., Holbrook, J. J., and Delcour, J. (1995) D175 discriminates between NADH and NADPH in the coenzyme binding site of Lactobacillus delbrueckii subsp. bulgaricus D-lactate dehydrogenase. *Biochem. Biophys. Res. Commun.* **208**, 895-900

101. Didierjean, C., Rahuel-Clermont, S., Vitoux, B., Dideberg, O., Branlant, G., and Aubry, A. (1997) A crystallographic comparison between mutated glyceraldehyde-3-phosphate dehydrogenases from Bacillus stearothermophilus complexed with either NAD^+ or $NADP^+$. *J. Mol. Biol.* **268**, 739-759

102. Yaoi, T., Miyazaki, K., and Oshima, T. (1994) Roles of Arg231 and Tyr284 of Thermus thermophilus isocitrate dehydrogenase in the coenzyme specificity. *FEBS Lett.* **355**, 171-172

103. Haeffner-Gormley, L., Chen, Z., Zalkin, H., and Colman, R. F. (1992) Importance of lysine-286 at the NADP site of glutamate dehydrogenase from Salmonella typhimurium. *Biochemistry* **31**, 7807-7814

104. Huang, S., Appleman, R., Tan, X. H., Thompson, P. D., Blakley, R. L., Sheridan, R. P., Venkataraghavan, R., and Freisheim, J. H. (1990) Role of lysine-54 in determining cofactor specificity and binding in human dihydrofolate reductase. *Biochemistry* **29**, 8063-8069

105. Perez-Miller, S. J., and Hurley, T. D. (2003) Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase. *Biochemistry* **42**, 7100-7109

106. Hammen, P. K., Allali-Hassani, A., Hallenga, K., Hurley, T. D., and Weiner, H. (2002) Multiple conformations of NAD and NADH when bound to human cytosolic and mitochondrial aldehyde dehydrogenase. *Biochemistry* **41**, 7156-7168

107. Tsybovsky, Y., Donato, H., Krupenko, N. I., Davies, C., and Krupenko, S. A. (2007) Crystal structures of the carboxyl terminal domain of rat 10-formyltetrahydrofolate dehydrogenase: implications for the catalytic mechanism of aldehyde dehydrogenases. *Biochemistry* **46**, 2917-2929

108. Park, J., and Rhee, S. (2013) Structural basis for a cofactor-dependent oxidation protection and catalysis of cyanobacterial succinic semialdehyde dehydrogenase. *J. Biol. Chem.* **288**, 15760-15770

109. Kim, Y. G., Lee, S., Kwon, O. S., Park, S. Y., Lee, S. J., Park, B. J., and Kim, K. J. (2009) Redox-switch modulation of human SSADH by dynamic catalytic loop. *EMBO J.* **28**, 959-968

110. Racker, E. (1955) Mechanism of action and properties of pyridine nucleotide-linked enzymes. *Physiol. Rev.* **35**, 1-56

111. von Bahr-Lindstrom, H., Jeck, R., Woenckhaus, C., Sohn, S., Hempel, J., and Jornvall, H. (1985) Characterization of the coenzyme binding site of liver aldehyde dehydrogenase: differential reactivity of coenzyme analogues. *Biochemistry* **24**, 5847-5851

112. Hempel, J. D., and Pietruszko, R. (1981) Selective chemical modification of human liver aldehyde dehydrogenases E1 and E2 by iodoacetamide. *J. Biol. Chem.* **256**, 10889-10896

113. Farres, J., Wang, T. T., Cunningham, S. J., and Weiner, H. (1995) Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis. *Biochemistry* **34**, 2592-2598

114. Blatter, E. E., Tasayco, M. L., Prestwich, G., and Pietruszko, R. (1990) Chemical modification of aldehyde dehydrogenase by a vinyl ketone analogue of an insect pheromone. *Biochem. J.* **272**, 351-358

115. Kitson, T. M., Hill, J. P., and Midwinter, G. G. (1991) Identification of a catalytically essential nucleophilic residue in sheep liver cytoplasmic aldehyde dehydrogenase. *Biochem. J.***275 (Pt 1)**, 207-210

116. Blanco, J., Moore, R. A., and Viola, R. E. (2003) Capture of an intermediate in the catalytic cycle of L-aspartate-beta-semialdehyde dehydrogenase. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 12613-12617

117. D'Ambrosio, K., Pailot, A., Talfournier, F., Didierjean, C., Benedetti, E., Aubry, A., Branlant, G., and Corbier, C. (2006) The first crystal structure of a thioacylenzyme intermediate in the ALDH family: new coenzyme conformation and relevance to catalysis. *Biochemistry* **45**, 2978-2986

118. Moniot, S., Bruno, S., Vonrhein, C., Didierjean, C., Boschi-Muller, S., Vas, M., Bricogne, G., Branlant, G., Mozzarelli, A., and Corbier, C. (2008) Trapping of the

thioacylglyceraldehyde-3-phosphate dehydrogenase intermediate from Bacillus stearothermophilus. Direct evidence for a flip-flop mechanism. *J. Biol. Chem.* **283**, 21693-21702

119. Moore, S. A., Baker, H. M., Blythe, T. J., Kitson, K. E., Kitson, T. M., and Baker, E. N. (1998) Sheep liver cytosolic aldehyde dehydrogenase: the structure reveals the basis for the retinal specificity of class 1 aldehyde dehydrogenases. *Structure* **6**, 1541-1551

120. Wang, X., and Weiner, H. (1995) Involvement of glutamate 268 in the active site of human liver mitochondrial (class 2) aldehyde dehydrogenase as probed by site-directed mutagenesis. *Biochemistry* **34**, 237-243

121. Rebek, J., Jr. (1990) On the structure of histidine and its role in enzyme active sites. *Struct. Chem.* **1**, 129-131

122. Nishizuka, Y., Ichiyama, A., and Hayaishi, O. (1970) Metabolism of the benzene ring of tryptophan (mammals). *Methods Enzymol.* **17**, 463-466

123. Kurnasov, O., Goral, V., Colabroy, K., Gerdes, S., Anantha, S., Osterman, A., and Begley, T. P. (2003) NAD biosynthesis: identification of the tryptophan to quinolinate pathway in bacteria. *Chem. Biol.* **10**, 1195-1204.

124. Colabroy, K. L., and Begley, T. P. (2005) Tryptophan catabolism: identification and characterization of a new degradative pathway. *J. Bacteriol.* **187**, 7866-7869

125. Mehler, A. H. (1964) Nicotinic acid biosynthesis: control by an enzyme that competes with a spontaneous reaction. *Science* **145**, 817-819

126. Reinhard, J. F., Jr. (2004) *Pharmacological manipulation of brain kynurenine metabolism*. *Ann. N. Y. Acad. Sci.* **1035**, 335-349

127. Guillemin, G. J., Williams, K. R., Smith, D. G., Smythe, G. A., Croitoru-Lamoury, J., and Brew, B. J. (2003) Quinolinic acid in the pathogenesis of alzheimer's disease. in *Adv. Exp. Med. Biol.* pp 167-176

128. Guidetti, P., and Schwarcz, R. (2003) 3-Hydroxykynurenine and quinolinate: pathogenic synergism in early grade huntington's disease? in *Adv. Exp. Med. Biol.* pp 137-145

129. Schwarcz, R. (2004) The kynurenine pathway of tryptophan degradation as a drug target. *Curr. Opin. Pharmacol.* **4**, 12-17

130. Mehler, A. H. (1956) Formation of picolinic and quinolinic acids following enzymatic oxidation of 3-hydroxyanthranilic acid. *J. Biol. Chem.* **218**, 241-254

131. Garavaglia, S., Perozzi, S., Galeazzi, L., Raffaelli, N., and Rizzi, M. (2009) The crystal structure of human α-amino-β-carboxymuconic-ε-semialdehyde decarboxylase (ACMSD) in complex with 1,3-dihydroxyacetonephosphate suggests a regulatory link between NAD synthesis and glycolysis. *FEBS J.* **276**, 6615-6623

132. Horton, R. M. (1995) PCR-mediated recombination and mutagenesis. SOEing together tailor-made genes. *Mol. Biotechnol.* **3**, 93-99

133. Otto, J. J. (1993) Immunoblotting. *Methods in Cell Biology* **37**, 105-117

134. Paz, M. A., Fluckiger, R., Boak, A., Kagan, H. M., and Gallop, P. M. (1991) Specific detection of quinoproteins by redox-cycling staining. *J. Biol. Chem.* **266**, 689-692

135. Yan, F., Li, T., Lipscomb, J. D., Liu, A., and Liu, H. W. (2005) Site-directed mutagenesis and spectroscopic studies of the iron-binding site of (S)-2-hydroxypropylphosphonic acid epoxidase. *Arch. Biochem. Biophys.* **442**, 82-91

136. Chen, L. Y., Doi, N., Durley, R. C. E., Chistoserdov, A. Y., Lidstrom, M. E., Davidson, V. L., and Mathews, F. S. (1998) Refined crystal structure of methylamine

dehydrogenase from paracoccus denitrificans at 1.75 angstrom resolution. *J. of Mol. Biol.* **276**, 131-149

137. Otwinowski, Z., and Minor, W. . (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 303-326

138. Vagin, A., and Teplyakov, A. (2010) Molecular replacement with MOLREP. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 22-25

139. Bailey, S. (1994) The Ccp4 Suite - Programs for Protein Crystallography. *Acta Crystallogr. D.* **50**, 760-763

140. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. . (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D.* **53**, 240-255

141. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D.* **60**

142. Ryle, M. J., Koehntop, K. D., Liu, A. M., Que, L., and Hausinger, R. P. (2003) Interconversion of two oxidized forms of taurine/alpha-ketoglutarate dioxygenase, a non-heme iron hydroxylase: Evidence for bicarbonate binding. *Pro. Natl. Acad. Sci. U. S. A.* **100**, 3790-3795

143. Muller, I., Stuckl, C., Wakely, J., Kertesz, M., and Uson, I. (2005) Succinate complex crystal structures of the α-ketoglutarate-dependent dioxygenase AtsK - Steric aspects of enzyme self-hydroxlation. *J. Biol. Chem.* **280**, 5716-5723

144. Davis, J. C., Lin, S. S., and Averill, B. A. (1981) Kinetics and optical spectroscopic studies on the purple acid phosphatase from beef spleen. *Biochemistry* **20**, 4062-4067

145. Zambonelli, C., and Roberts, M. F. (2003) An iron-dependent bacterial phospholipase D reminiscent of purple acid phosphatases. *J. Biol. Chem.* **278**, 13706-13711

146. Vainshtein, B. K., Melik-Adamyan, W. R., Barynin, V. V., Vagin, A. A., and Grebenko, A. I. (1981) Three-dimensional structure of the enzyme catalase. *Nature* **293**, 411-412

147. Pyrz, J., Roe, A., Stern, L., and Que, L. J. (1985) Model studies of iron-tyrosinate proteins. *J. Am. Chem. Soc.* **107**, 614-620

148. Bradley, F. C., Lindstedt, S., Lipscomb, J. D., Que, L., Jr., Roe, A. L., and Rundgren, M. (1986) 4-Hydroxyphenylpyruvate dioxygenase is an iron-tyrosinate protein. *J. Biol. Chem.* **261**, 11693-11696

149. Davis, M., Orville, A., Neese, F., Zaleski, J., Lipscomb, J., and Solomon, E. (2002) Spectroscopic and electronic structure studies of protocatechuate 3,4-dioxygenase: Nature of tyrosinate-Fe(III) Bonds and their contribution to reactivity. *J. Am. Chem. Soc.* **124**, 602-614

150. Que, L. J. (1988) Metal-tyrosinate proteins in biological applications of Raman spectroscopy: Resonance Raman spectra of heme and metalloproteins. *Vol. Resonance Raman spectra of heme and metalloproteins, Wiley-Interscience, New York*

151. Tomimatsu, Y., Klint, S., and Scherer, J. (1976) Resonance Raman spectra of iron(III)-, copper(II)-, cobalt(III)-, and manganese(III)-transferrins and of bis (2,4,6-trichlorophenolato) diimidazolecopper(II) monohydrate, a possible model for copper(II) binding to transferrins. *Biochemistry* **15**, 4918-4924

152. Siu, D.-T., Orville, A., Lipscomb, J., D., O., and Que, L. J. (1992) Resonance Raman studies of Protocatechuate 3,4-dioxygenase from *Brevibacterium fuscum*. *Biochemistry* **31**, 10443-10448

153. Que, L. J., Heistand, R. I., Mayer, R., and Roe, A. L. (1980) Resoannce Raman studies of pyrocatechase-inhibitor complexes. *Biochemistry* **19**, 2588-2593

154. Que, L. J., and Heistand, R. H. I. (1979) Resonance Raman Studies of Pyrocatechase. *J. Am. Chem. Soc.* **101**, 2219-2221

155. Heistand, R. I., Lauffer, R., Fikrig, E., and Que, L. J. (1982) Catecholate and phenolate iron complexes as models for the dioxygenases. *J. Am. Chem. Soc.* **104**, 2789-2796

156. Gaber, B., Miskowski, V., and Spiro, T. (1974) Resonance Raman scattering from iron(III)- and copper(II)- transferrin and an iron(III) model compound. A spectroscopic interpretation of the transferrin binding site. *J. Am. Chem. Soc.* **96**, 6868-6873

157. Aisen, P., and Leibman, A. (1972) Lactoferrin and transferrin: A comparative study. *Biochim. Biophys. Acta* **75**, 203-222

158. Ainscough, E., Brodie, A., Plowman, J., Bloor, S., Loehr, J., and Loehr, T. (1980) Studies on human lactoferrin by electron paramagnetic resonance, fluorescence, and resonance Raman spectroscopy. *Biochemistry* **19**, 4072-4079

159. Aasa, R., Mamlstrom, A., Saltman, P., and Vanngard, T. (1963) The specific binding of iron(III) and copper(II) to transferrin and conalbumin. *Biochim. Biophys. Acta* **75**, 203-222

160. Cox, D. D., Benkovic, S. J., Bloom, L. M., Bradley, F. C., Nelson, M. J., Que, L., Jr., and Wallick, D. E. (1988) Catecholate LMCT bands as probes for the active sites of nonheme iron oxygenases. *J. Am. Chem. Soc.* **110**, 2026-2032

161. Smith, J., Thomson, A., Proudfoot, A., and Wells, T. (1997) Identification of an Fe(III)-dihydroxyphenylalanine site in recombinant phosphomannose isomerase from *Candida albicans*. *Eur. J. Biochem.* **42**, 1854-1862

162. Ryle, M., Liu, A., Muthukumaran, R., Ho, R., Koehntop, K., McCracken, J., Que, L. J., and Hausinger, R. (2003) O_2- and α-ketoglutarate-dependent tyrosyl radical formation in TauD, an α-Keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42**, 1854-1862

163. Michaud-Soret, I., Anderson, K., Que, L. J., and Haavik, J. (1995) Resonance Raman studies of catecholate and phenolate complexes of recombinate human tyrosine hydroxylase. *Biochemistry* **34**, 5504-5510

164. Farquhar, E. R., Koehntop, K. D., Emerson, J. P., and Que, L. J. (2005) Post-translational self-hydroxylation: A probe for oxygen activation mechanisms in non-heme iron enzymes. *Biochem. Biophys. Res. Commun.* **338**, 230-239

165. Antanaitis, B., Strekas, T., and Aisen, P. (1982) Characterization of pink and purple uteroferrin by resonance Raman and CD spectroscopy. *J. Biol. Chem.* **257**, 3766-3770

166. Siamwiza, M., Lord, R., Chen, M., Takamatsu, T., Narada, I., Matsuura, H., and Shimanauchi, T. (1975) Interpretation of the doublet at 850 and 830 cm^{-1} in the Raman spectra of tyrosyl residues in proteins and certain model compounds. *Biochemistry* **14**, 4870-4876

167. Salma, S., Strong, J. D., Neilands, J. B., and Spiro, T. G. (1978). *Biochemistry* **17**, 3781-3785

168. Gaber, B., Sheridian, F., Bazer, F., and Roberts, R. (1979) Resonance Raman scattering from uteroferrin, the purple glycoprotein of porcine uterus. *J. Biol. Chem.* **254**, 8340-8342

169. Lindstedt, S., and Rundgren, M. (1982) Blue color, metal content, and substrate binding in 4-hydroxyphenylpyruvate dioxygenase from *Pseudomonas sp*. strain p. J. 874. *J. Biol. Chem.* **257**, 11922-11931

170. Farquhar, E. R., Emerson, J. P., Koehntop, K. D., Reynolds, M. F., Trmčić, M., and Que, L. (2011) *In vivo* self-hydroxylation of an iron-substituted manganese-dependent extradiol cleaving catechol dioxygenase. *J. Biol. Inorg. Chem.* **16**, 589-597

171. Koehntop Kevin, D., Marimanikkuppam, S., Ryle Matthew, J., Hausinger Robert, P., and Que, L., Jr. (2006) Self-hydroxylation of taurine/alpha-ketoglutarate dioxygenase: evidence for more than one oxygen activation mechanism. *J. Biol. Inorg. Chem.* **11**, 63-72

172. Serre, L., Sailland, A., Sy, D., Boudec, P., Rolland, A., Pebay-Peyroula, E., and Cohen-Addad, C. (1999) Crystal structure of *Pseudomonas fluorescens* 4-hydroxyphenylpyruvate dioxygenase: An enzyme involved in the tyrosine degradation pathway. *Structure.* **7**, 977-988

173. Horrocks, W. D. J., Ishley, J. N., Holmquist, B., and Thompson, J. S. (1980) Structural and electronic mimics of the active site of cobalt(II)-substituted zinc metalloenzymes. *J. Inorg. Biochem.* **12**, 131-141

174. Sellin, S., Eriksson, L. E., Aronsson, A. C., and Mannervik, B. (1983) Octahedral metal coordination in the active site of glyoxalase I as evidenced by the properties of Co(II)-Glyoxalase I. *J. Biol. Chem.* **258**, 2091–2093

175. Bertini, I., and Luchinat, C. (1984) High spin cobalt(II) as a probe for the investigation of metalloproteins. *Adv. Inorg. Biochem.* **6**, 71-111

176. Seffernick, J. L., McTavish, H., Osborne, J. P., de Souza, M. L., Sadowsky, M. J., and Wackett, L. P. (2002) Atrazine chlorohydrolase from Pseudomonas sp strain ADP is a metalloenzyme. *Biochemistry* **41**, 14430-14437

177. Bencini, A., Bertini, I., Canti, G., Gatteschi, D., and Luchinat, C. (1981) The EPR Spectra of the Inhibitor Derivatives of Cobalt Carbonic Anhydrase. *J. Inorg. Biochem.* **14**, 81-93

178. Holm, R. H., Kennepohl, P., and Solomon, E. I. . (1996) Structural and functional aspects of metal sites in biology. *Chem. Rev.* **96**, 2239-2314

179. Whittaker, M. M., and Whittaker, J. W. (1998) A glutamate bridge is essential for dimer stability and metal selectivity in manganese superoxide dismutase. *J. Biol. Chem.* **273**, 22188-22193

180. Forouhar, F., Abashidze, M., Jayaraman, S., Cunningham, K., Ciao, M., Ma, L., Xiao, R., Acton, T. B., Montelione, G. T., Hunt, J. F., and Tong, L. . (2007) Crystal structure of 4-oxalomesaconate hydratase (LigJ) from Rhodopseudomonas palustris. *RCSB Protein Databank*

181. Silverman, D. N., and Lindskog, S. (1988) The catalytic mechanism of carbonic anhydrase: implications of a rate-limiting protolysis of water. *Acc. Chem. Res.* **21**, 30-36

182. Sly, W. S., and Hu, P. Y. (1995) Human carbonic anhydrases and carbonic anhydrase deficiencies. *Annu. Rev. Biochem.* **64**, 375-401

183. Woolley, P. (1975) Models for metal ion function in carbonic anhydrase. *Nature* **258**, 677-682

184. Donald, J. E., Kulp, D. W., and DeGrado, W. F. (2011) Salt bridges: geometrically specific, designable interactions. *Proteins* **79**, 898-915

185. Prell, J. S., O'Brien, J. T., Steill, J. D., Oomens, J., and Williams, E. R. (2009) Structures of protonated dipeptides: the role of arginine in stabilizing salt bridges. *J. Am. Chem. Soc.* **131**, 11442-11449

186. Ghosh, M., and Datta, A. K. (1994) Probing the Function(S) of Active-Site Arginine Residue in Leishmania-Donovani Adenosine Kinase. *Biochem J* **298**, 295-301

187. Reczkowski, R. S., Taylor, J. C., and Markham, G. D. (1998) The active-site arginine of S-adenosylmethionine synthetase orients the reaction intermediate. *Biochemistry* **37**, 13499-13506

188. Chaidaroglou, A., Brezinski, D. J., Middleton, S. A., and Kantrowitz, E. R. (1988) Function of arginine-166 in the active site of Escherichia coli alkaline phosphatase. *Biochemistry* **27**, 8338-8343

189. Blachnio, K., and Przykorska, A. (2005) [Specific arginine mediated RNA recognition]. *Postepy. Biochem.* **51**, 339-344

190. Charnock, S. J., Brown, I. E., Turkenburg, J. P., Black, G. W., and Davies, G. J. (2002) Convergent evolution sheds light on the anti-beta-elimination mechanism common to family 1 and 10 polysaccharide lyases. *Pro. Natl. Acad. Sci. U.S.A.* **99**, 12067-12072

191. Schlippe, Y. V. G., and Hedstrom, L. (2005) Is Arg418 the catalytic base required for the hydrolysis step of the IMP dehydrogenase reaction? *Biochemistry* **44**, 11700-11707

192. Doherty, M. K., Pealing, S. L., Miles, C. S., Moysey, R., Taylor, P., Walkinshaw, M. D., Reid, G. A., and Chapman, S. K. (2000) Identification of the active site acid/base catalyst in a bacterial fumarate reductase: a kinetic and crystallographic study. *Biochemistry* **39**, 10695-10701

193. Colabroy, K. L., and Begley, T. P. (2005) The pyridine ring of NAD is formed by a nonenzymatic pericyclic reaction. *J. Am. Chem. Soc.* **127**, 840-841

194. Pucci, L., Perozzi, S., Cimadamore, F., Orsomando, G., and Raffaelli, N. (2007) Tissue expression and biochemical characterization of human 2-amino 3-carboxymuconate 6-semialdehyde decarboxylase, a key enzyme in tryptophan catabolism. *FEBS J.* **274**, 827-840

195. Koontz, W. A., and Shiman, R. (1976) Beef kidney 3-hydroxyanthranilic acid oxygenase. Purification, characterization, and analysis of the assay. *J. Biol. Chem.* **251**, 368-377

196. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D.* **66**, 213-221

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

198. Fukuwatari, T., Ohsaki, S., Fukuoka, S., Sasaki, R., and Shibata, K. (2004) Phthalate esters enhance quinolinate production by inhibiting alpha-amino-beta-carboxymuconate-epsilon-semialdehyde decarboxylase (ACMSD), a key enzyme of the tryptophan pathway. *Toxicol. Sci.* **81**, 302-308

199. Fukuwatari, T., Ohsaki, S., Suzuki, Y., Fukuoka, S., Sasaki, R., and Shibata, K. (2003) The effects of phthalate esters on the tryptophan-niacin metabolism. *Adv. Exp. Med. Biol.* **527**, 659-664

200. Ellis, K. E., Frato, K. E., and Elliott, S. J. (2012) Impact of quaternary structure upon bacterial cytochrome c peroxidases: does homodimerization matter? *Biochemistry* **51**, 10008-10016

201. Renatus, M., Stennicke, H. R., Scott, F. L., Liddington, R. C., and Salvesen, G. S. (2001) Dimer formation drives the activation of the cell death protease caspase 9. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 14250-14255

202. Datta, D., McClendon, C. L., Jacobson, M. P., and Wells, J. A. (2013) Substrate and inhibitor-induced dimerization and cooperativity in caspase-1 but not caspase-3. *J. Biol. Chem.* **288**, 9971-9981

203. Hor, L., Dobson, R. C., Downton, M. T., Wagner, J., Hutton, C. A., and Perugini, M. A. (2013) Dimerization of bacterial diaminopimelate epimerase is essential for catalysis. *J. Biol. Chem.* **288**, 9238-9248

204. Swanson, R. V., Bourret, R. B., and Simon, M. I. (1993) Intermolecular complementation of the kinase activity of CheA. *Mol. Microbiol.* **8**, 435-441

205. Marianayagam, N. J., Sunde, M., and Matthews, J. M. (2004) The power of two: protein dimerization in biology. *Trends Biochem. Sci.* **29**, 618-625

206. Nooren, I. M., and Thornton, J. M. (2003) Structural characterisation and functional significance of transient protein-protein interactions. *J. Mol. Biol.* **325**, 991-1018

207. Unkles, S. E., Rouch, D. A., Wang, Y., Siddiqi, M. Y., Glass, A. D., and Kinghorn, J. R. (2004) Two perfectly conserved arginine residues are required for substrate binding in a high-affinity nitrate transporter. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 17549-17554

208. Ogura, T., Whiteheart, S. W., and Wilkinson, A. J. (2004) Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA^{+} ATPases. *J. Struct. Biol.* **146**, 106-112

209. Parkinson, J. S., and Kofoid, E. C. (1992) Communication modules in bacterial signaling proteins. *Annu. Rev. Genet.* **26**, 71-112

210. Hess, J. F., Bourret, R. B., and Simon, M. I. (1988) Histidine phosphorylation and phosphoryl group transfer in bacterial chemotaxis. *Nature* **336**, 139-143

211. Hess, J. F., Oosawa, K., Matsumura, P., and Simon, M. I. (1987) Protein phosphorylation is involved in bacterial chemotaxis. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 7609-7613

212. Saunders, A. H., Griffiths, A. E., Lee, K. H., Cicchillo, R. M., Tu, L., Stromberg, J. A., Krebs, C., and Booker, S. J. (2008) Characterization of quinolinate synthases from Escherichia coli, Mycobacterium tuberculosis, and Pyrococcus horikoshii indicates that [4Fe-4S] clusters are common cofactors throughout this class of enzymes. *Biochemistry* **47**, 10999-11012

213. Owe-Young, R., Webster, N. L., Mukhtar, M., Pomerantz, R. J., Smythe, G., Walker, D., Armati, P. J., Crowe, S. M., and Brew, B. J. (2008) Kynurenine pathway metabolism in human blood-brain-barrier cells: implications for immune tolerance and neurotoxicity. *J. Neurochem.* **105**, 1346-1357

214. Heyes, M. P., Saito, K., Lackner, A., Wiley, C. A., Achim, C. L., and Markey, S. P. (1998) Sources of the neurotoxin quinolinic acid in the brain of HIV-1-infected patients and retrovirus-infected macaques. *FASEB J.* **12**, 881-896

215. Guillemin, G. J., Kerr, S. J., and Brew, B. J. (2005) Involvement of quinolinic acid in AIDS dementia complex. *Neurotox. Res.* **7**, 103-123

216. Guillemin, G. J., Wang, L., and Brew, B. J. (2005) Quinolinic acid selectively induces apoptosis of human astrocytes: potential role in AIDS dementia complex. *J. Neuroinflammation.* **2**, 16

217. Widner, B., Leblhuber, F., Walli, J., Tilz, G. P., Demel, U., and Fuchs, D. (2000) Tryptophan degradation and immune activation in Alzheimer's disease. *J. Neural. Transm.* **107**, 343-353

218. *The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC (www.pymol.org).*

219. Howes, B. D., Abraham, Z. H., Lowe, D. J., Bruser, T., Eady, R. R., and Smith, B. E. (1994) EPR and electron nuclear double resonance (ENDOR) studies show nitrite binding to the type 2 copper centers of the dissimilatory nitrite reductase of Alcaligenes xylosoxidans (NCIMB 11015). *Biochemistry* **33**, 3171-3177

220. Nguyen, T. T., Fedorov, A. A., Williams, L., Fedorov, E. V., Li, Y., Xu, C., Almo, S. C., and Raushel, F. M. (2009) The mechanism of the reaction catalyzed by uronate isomerase illustrates how an isomerase may have evolved from a hydrolase within the amidohydrolase superfamily. *Biochemistry* **48**, 8879-8890

221. Xu, S., Li, W., Zhu, J., Wang, R., Li, Z., Xu, G. L., and Ding, J. (2013) Crystal structures of isoorotate decarboxylases reveal a novel catalytic mechanism of 5-carboxyl-uracil decarboxylation and shed light on the search for DNA decarboxylase. *Cell Res.* **23**, 1296-1309

222. Page, M. I., and Badarau, A. (2008) The mechanisms of catalysis by metallo beta-lactamases. *Bioinorg. Chem. Appl.*, 576297

223. Pingoud, A., Fuxreiter, M., Pingoud, V., and Wende, W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol. Life Sci.* **62**, 685-707

224. Jonsson, B. M., Hakansson, K., and Liljas, A. (1993) The structure of human carbonic anhydrase II in complex with bromide and azide. *FEBS Lett.* **322**, 186-190

225. Aubert, S. D., Li, Y., and Raushel, F. M. (2004) Mechanism for the hydrolysis of organophosphates by the bacterial phosphotriesterase. *Biochemistry* **43**, 5707-5715

226. Wilson, D. K., and Quiocho, F. A. (1993) A pre-transition-state mimic of an enzyme: X-ray structure of adenosine deaminase with bound 1-deazaadenosine and zinc-activated water. *Biochemistry* **32**, 1689-1694

227. Liljas, A., Hakansson, K., Jonsson, B. H., and Xue, Y. F. (1994) Inhibition and Catalysis of Carbonic-Anhydrase - Recent Crystallographic Analyses. *Eur. J. of Biochem.* **219**, 1-10

228. Keszthelyi, D., Troost, F. J., and Masclee, A. A. (2009) Understanding the role of tryptophan and serotonin metabolism in gastrointestinal function. *Neurogastroenterol. Motil.* **21**, 1239-1249

229. Myint, A. M., Kim, Y. K., Verkerk, R., Scharpe, S., Steinbusch, H., and Leonard, B. (2007) Kynurenine pathway in major depression: evidence of impaired neuroprotection. *J. Affect. Disord.* **98**, 143-151

230. Ogawa, T., Matson, W. R., Beal, M. F., Myers, R. H., Bird, E. D., Milbury, P., and Saso, S. (1992) Kynurenine pathway abnormalities in Parkinson's disease. *Neurology* **42**, 1702-1706

231. Kerr, S. J., Armati, P. J., Guillemin, G. J., and Brew, B. J. (1998) Chronic exposure of human neurons to quinolinic acid results in neuronal changes consistent with AIDS dementia complex. *AIDS* **12**, 355-363

232. Heyes, M. P., Brew, B. J., Martin, A., Price, R. W., Salazar, A. M., Sidtis, J. J., Yergey, J. A., Mouradian, M. M., Sadler, A. E., Keilp, J., and et al. (1991) Quinolinic acid in cerebrospinal fluid and serum in HIV-1 infection: relationship to clinical and neurological status. *Ann. Neurol.* **29**, 202-209

233. Unden, G., and Bongaerts, J. (1997) Alternative respiratory pathways of Escherichia coli: energetics and transcriptional regulation in response to electron acceptors. *Biochim. Biophys. Acta* **1320**, 217-234

234. Nicholls, D. G. (2002) Mitochondrial function and dysfunction in the cell: its relevance to aging and aging-related disease. *Int. J. Biochem. Cell Biol.* **34**, 1372-1381

235. Hempel, J., Nicholas, H., and Lindahl, R. (1993) Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework. *Protein Sci.* **2**, 1890-1900

236. Ichiyama, A., Nakamura, S., Kawai, H., Honjo, T., Nishizuka, Y., Hayaishi, O., and Senoh, S. (1965) Studies on the Metabolism of the Benzene Ring of Tryptophan in Mammalian Tissues. Ii. Enzymic Formation of Alpha-Aminomuconic Acid from 3-Hydroxyanthranilic Acid. *J. Biol. Chem.* **240**, 740-749

237. Orville, A. M., Buono, R., Cowan, M., Heroux, A., Shea-McCarthy, G., Schneider, D. K., Skinner, J. M., Skinner, M. J., Stoner-Ma, D., and Sweet, R. M. (2011) Correlated single-crystal electronic absorption spectroscopy and X-ray crystallography at NSLS beamline X26-C. *J. Synchrotron Radiat.* **18**, 358-366

238. Pothineni, S. B., Strutz, T., and Lamzin, V. S. (2006) Automated detection and centring of cryocooled protein crystals. *Acta Crystallogr. D* **62**, 1358-1368

239. Geoghegan, K. F., Dixon, H. B., Rosner, P. J., Hoth, L. R., Lanzetti, A. J., Borzilleri, K. A., Marr, E. S., Pezzullo, L. H., Martin, L. B., LeMotte, P. K., McColl, A. S., Kamath, A. V., and Stroh, J. G. (1999) Spontaneous alpha-N-6-phosphogluconoylation of a "His tag" in Escherichia coli: the cause of extra mass of 258 or 178 Da in fusion proteins. *Anal. Biochem.* **267**, 169-184

240. He, Z., Davis, J. K., and Spain, J. C. (1998) Purification, characterization, and sequence analysis of 2-aminomuconic 6-semialdehyde dehydrogenase from Pseudomonas pseudoalcaligenes JS45. *J. Bacteriol.* **180**, 4591-4595

241. Abriola, D. P., Fields, R., Stein, S., MacKerell, A. D., Jr., and Pietruszko, R. (1987) Active site of human liver aldehyde dehydrogenase. *Biochemistry* **26**, 5679-5684

242. Steinmetz, C. G., Xie, P., Weiner, H., and Hurley, T. D. (1997) Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure* **5**, 701-711