5-7-2016

Investigation Of A Novel Mammalian Thiol Dioxygenase Structure: Human Cysteamine Dioxygenase

Tseng Xiong
txiong1@gsu.edu

Tseng Xiong
Georgia State University, txiong9@gmail.com

Recommended Citation

INVESTIGATION OF A NOVEL MAMMALIAN THIOL DIOXYGENASE STRUCTURE:

HUMAN CYSTEAMINE DIOXYGENASE

by

TSENG XIONG

Under the Direction of Aimin Liu, PhD

ABSTRACT

In 2007, a gene homolog of CDO encoded by the gene Gm237 in the DUF164 family was identified as cysteamine dioxygenase (ADO). ADO is one of the only known thiol dioxygenases found in mammals. Both ADO and its partner cysteine dioxygenase (CDO) are non-heme iron dependent enzymes that play a crucial role in the biosynthesis of taruine/hypotaurine by insertion of a dioxygen molecule. However, ADO has been overshadowed by CDO as heavy research focus on CDO over the past decade has led to the elucidation of its structure and possible mechanistic properties. In an effort to further understand ADO's mechanism and regulating role in vivo, this work will be focused on the mammalian hADO and trying to gain further insight on hADO's structural features via crystallography work.

Investigation of the crystallization parameters for hADO has elucidated several potential

conditions. Detailed work on these crystallization parameters will be presented.

INDEX WORDS: Cysteamine dioxygenase, Cysteine dioxygenase, Protein crystallography,

Cysteamine, Fe-reconstitution, Non-heme dioxygenase

INVESTIGATION OF A NOVEL MAMMALIAN THIOL DIOXYGENASE STRUCTURE:

HUMAN CYSTEAMINE DIOXYGENASE

by

TSENG XIONG

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

Master of Science

in the College of Arts and Sciences

Georgia State University

2016

2016

INVESTIGATION OF A NOVEL MAMMALIAN THIOL DIOXYGENASE STRUCTURE:

HUMAN CYSTEAMINE DIOXYGENASE

by

TSENG XIONG

Committee Chair: Aimin Liu

Committee: Giovanni Gadda

Donald Hamelberg

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2016

DEDICATION

I would like to dedicate this dissertation to my family and friends. I would like to thank my parents Ge and Xai Xiong for their contributions and support through the years. You guys have always supported me no matter what I wanted to pursue and have always been there for me. I would like to thank my brothers Luke, Noah, and Kevin along with my sister Elizabeth for being there for me and always encouraging me to strive to be the best that I can be. Without you guys I wouldn't be where I am today. I would like to thank my girlfriend Maikou for her help, love, and support. You are the staple in my life and have helped me become who I am today. Thank you for your blessings, encouragement, and support. Due to all of your support, motivation, and guidance I am where I am today as a person and scholar.

Lastly, I would like to thank the Lord for his guidance, compassion, and love. He has made my success possible by blessing me with the family and friends that I have and by putting me in the right situations at the right time. Thank you for always watching over me and walking with me even when I become astray.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1   INTRODUCTION

Enzymes are proteins which functions as biological catalysts. Enzymes are produced in all living organisms and are necessary in almost all metabolic processes to sustain life[1]Enzyme properties are versatile and to date enzymes are used in various assets of applications ranging from the food and beverage industry, starch and fuel production, organic synthesis, pharmaceutical etc. Due to evolution of modern biotechnology and protein engineering the possibilities of utilizing enzymes in multiple facets of life have become possible.[2] However, all of these possibilities would not be available if not for the understanding of enzymes at the molecular and biological levels.

Enzymes are composed of the building blocks called amino acids. There are 20 key amino acids utilized by the body's proteins and they are either obtained from food or synthesized by the body. When amino acids form polypeptides, they can fold to form three-dimensional structures due to hydrophobic forces and intermolecular interactions between the amino acid residues. The structure formed is significant to the amino acid sequence allowing the enzyme to have different folds and reactive chemical nature. Enzymes can bind specific substrates via different enzyme binding affinities such as the lock and key model. Many enzymes tend to require cofactors such as metals, flavin, NAD, etc to perform these enzymatic activities. One key class of enzyme that incorporates metals is called metalloenzymes. Metalloenzymes are all protein associated with a metal ion to form a coordinative active site for catalytic activity. Some mettalloenzyme classes fall under heme and non-heme enzymes where heme enzymes are associated with presence of an iron porphryin which bonds via covalently or non-covalently interactions and non-heme enzymes which lack presence of a heme.

Enzymes have been studied since the 1800s and remain to be the focal point of many research investigations to date. The biological identification of enzymes remained elusive until the early 1900s when James B. Sumner showed that the enzyme urease was in a fact a protein and crystallized it..[3] Sumner's crystallization work opened up the doors for new investigative premises of enzymes as the discovery that proteins can be crystallized allowed for the structure analysis of enzymes by x-ray crystallography. Structural biology or protein crystallography has allowed for the study of enzymes at the atomic level and remains to be a key tool for understanding the structure and function of an enzyme.

1.1 Protein Crystallography

The discovery of X-rays in the late 1800s along with subsequent developments by William Brag and his son revealed that X-ray diffraction could be used to determine the atomic structure of atoms.[4] Their work proved to be the foundation for X-ray crystallography as years later, the first protein structure of myoglobin was determined and published. Since then, macromolecule crystallography has become a practical tool utilized in various fields and applications because studies at the atomic level of molecules have proved to be of great value in revealing structure/function relationship; which is important for understanding how enzymes, nucleic acids, or other macromolecules will interact in a biological system. Crystallography studies have become of considerable interest to many different fields including academia, pharmaceutical, biotechnological, chemical industries, etc as a promising tool in protein engineering, drug design, and various other biological system studies such as understanding certain diseases. The elaborate information obtained from the 3-dimensonal structure of a

protein or other biological macromolecule provides potential information on understanding

mechanisms by which enzymes, receptor, hormones, lipids, etc. function in a biological setting. 5

One key application of protein crystallography is utilizing the information obtained from

the crystal structure to understand enzyme function and mechanism. This involves

understanding the chemical reaction taking place between the substrates, cofactors, etc. involved

for the formation of a product and enzyme kinetics. Understanding these aspects along with the

cellular location of the protein allows insight on cellular regulation which can prove to be useful

information for understanding the nature of particular diseases or abnormal malfunctions.

Information from the structure of proteins can also identify mutational consequences which can

affect enzyme function. There are various types of mutations: substitution, deletion, inversion,

etc. causing the enzyme to be improperly folded or lose/gain enzyme function. Thus, mutations

can greatly affect enzyme activity and cause imbalance in cellular regulation leading to diseases.

Structural crystallography relies mainly on the scattering of X-rays by the electrons in the

molecules of the investigated sample. The crystal is exposed to an X-ray beam and exposed at

various orientations resulting in the scattering of X-ray radiation in selected directions.

Scattering is governed by the geometry of the crystal unit cell while the intensity of the

diffractions depends on the grouping of all atoms within the unit cell. This property allows for 3-

D diffraction gratings unique to the geometry of the protein. Thus, in general the crystal

structure is encoded in the diffracted X-rays and its intensities which is defined by the shape and

symmetry of the cell. The position of each atom in the unit cell influences the intensities of the

reflections. The primary result from the X-ray diffractions is a map of the electron density

within the crystal and after proper refinement the electron density map can be then be processed.

Several programs today such as Mosflm, XDS, and HKL-2000 were developed for processing

electron density maps. From the electron density map, a structural model can be built, often with

the assistance of the amino acid sequence information (i.e., the primary structure). The clarity of

the electron maps depends on the resolution of the diffraction data. At low resolutions, it is not

possible to accurately locate or distinguish between individual atoms and their identity.

However, with increasing resolution, it becomes possible to identify the stereochemistry of

individual amino acids and peptide groups to give a clear layout of the protein structure.

Macromolecule crystallography has come a long way as today there is establishment of a Protein

Data Base (PDB) as a repository site for all solved crystal structures. As of today, there are at

least 37,914 deposited crystal structures.

1.2   Protein Dynamic Growth

Protein crystallization is a complex process with various parameters that must be taken

into consideration. One key parameter is the nature and the properties of the protein being

crystallized. Proteins tend to be large, complex molecules which can compose of several

subunits. This can cause proteins to have varying molecular and dynamic properties depending

on the protein's amino acid sequence which can affect protein crystallization. This can include

but is not limited to relatively flexible structures, chemically or physically instable

conformations, unique three-dimensional folds, or charge and size of the protein. For a crystal to

form, it does not depend on the properties of the protein alone. The interactions between protein

molecules and its surrounding solution environment must align properly to from a crystal with

specific geometric arrangement. Surrounding environmental factors such as pH, ionic strength,

temperature, concentration of protein, etc as shown in Figure 1 are all factors that must be taken

into consideration. Even today, it is still not possible to predict the crystallization conditions of a

protein as change in a single experimental parameter can influence the outcome of a crystallization experiment.

Crystallization is an intricate process that involves multiple equilibria between different states of the species being crystallized. There are three stages commonly associated with crystallization: 1. nucleation 2. crystal growth 3. cessation of growth. Nucleation occurs when enough molecules associate in three dimensional forms, resulting in a thermodynamically stable aggregate which provides a suitable platform for crystal growth. Crystal growth and nucleation occur in supersaturated solutions when the concentrations of the crystallized species exceed equilibrium solubility. However, the ideal conditions for suitable crystallization conditions are finite due to many different factors. [6] Therefore, it is important to guide the process toward a state of reduced solubility by modifying various parameters to ensure that solubility is decreased at a slow rate with respect to the nature of the species being crystallized in order to ensure that orderly crystals are formed instead of amorphous precipitate. The ideal approach would be to begin at the lowest supersaturation and slowly increase saturation due to depletion of solute from the solution as this allows the nuclei to grow properly and allow formation of define and stable crystals which will diffract. [7]

A wide variety of chemical, physical, and biochemical parameters affect the protein crystallization process. The purity of the protein being analyzed, the properties of the solution being used for crystallization (e.g. salts, buffers, pH, temperature), and the various techniques used for sampling handling are all important factors that must be taken into consideration. Table 1.2 lists some common parameters that are taken into consideration for protein crystallization.

Table 1. Environmental factors for protein crystallography. Adapted from ref 8&10.

Physical Factors	Chemical Factors	Biochemical Factors
Temperature	Precipitant type	Sample purity
Time	Precipitant concentration	Macromolecular impurities
Pressure	pH and buffer	Aggregation
Methodology	Ionic strength	Posttranslational Modifications
Gravity/Sedimentation	Reducing/oxidizing environment	Sample source
Vibrations/Sounds	Sample concentration	Chemical modifications
Magnetic/Electric Fields	Metal ions	Sequence modifications
Viscosity	Detergents	Sample symmetry
Equilibration rate	Polyions	Ligands, co-factors, inhibitors
Volume	Cross-linkers	Sample pI
Surface of crystallization	Reagent source	Purification methods
Sample handling		Microbial contamination

With so many parameters to be considered, there is no exact approach to determine the proper conditions that must be used to crystallize a protein. A few parameters will be briefly

discussed here, however more explicit reviews on the different parameters are explained in sources. [8,10]

1.2.1 Ligand/Substrate/Cofactors Parameter

It is often possible to increase the stability of a protein via addition of a ligand, substrate, nucleic acid, cofactor, or small molecule. [11] The ligand or stabilizing molecule can be added during various stages to the protein whether if it's during certain protein purification steps, during the entire purification process, or during the final concentration before crystallization. The addition of a ligand may help to stabilize the protein and lessened or elimination issues such as protein denaturation. It is also applicable to soak the ligands into pre-existing crystals to form and trap key intermediates to help elucidate the enzymatic pathway. It has been shown that certain proteins when expressed with a ligand of interest can help protein expression and increase protein solubility. [12,13] Addition of ligand during the various purification steps of protein purification such as cell-lysis was the key for the success of obtaining pure protein as seen for the crystallization of a recombinant enzyme (kinase 1) from a baculovirus in insect cells.

If no ligand was present, purification of the enzyme resulted in a mixture of protein, DNA, and lipids. If the ligand is soluble, the protein-ligand complex can be formed by adding the ligand to the protein during the final steps of concentration. It has also been noted that cofactors such as metals can be a key parameter for crystallization conditions. Many enzymes requires metals such as Fe, Mg, Ni, etc to be biologically active. For this reason, it is reasonably expected that incorporating metals, particularly divalent metals to metal dependent enzymes would help with crystal formation. This could be due to the fact that metal ions can stabilize the protein and the intermolecular interactions in the crystals.[14,15]

1.2.2   Precipitants

Protein solubility can be varied by changing the composition of the solution. This can be by changing solubility-influencing additives such as alcohols, detergents, hydrophilic polymers, etc. which are commonly referred to as precipitants. Listed in table 2 are common precipitants. There are four broad categories for protein precipitants: salts, organic solvents, polymers, and non-volatile organic compounds.

Table 2. Lists of additives that can affect protein solubility. Adapted from ref 8 & 10.

Salts	Organic Solvent	Polymers
Ammonium or sodiumn sulfate	Ethanol	Poly(ethylene) glycol 1000, 3350, 6000, 20000, etc.
Lithium sulfate	Isopropanol	Jeffamine T
Lithium chloride	1,3- propandediol	Polyamine
Sodium/ammonium/potassium chloride	Acetone	
Sodium/ammonium acetate	Butanol	
Magnesium/calcium sulfate	Acetonitrile	
Cetyltrimethyl ammonium salts	Dimethyl sulfoxide	
Calcium chloride	2,5-hexanediol	
Sodium formate	Methanol	
Ammonium nitrate	1,3-butyrolactone	

The ionic strength of the solution can be increased to decrease the solubility of a protein by a method referred to as "salting out". As salt concentration increases, water molecules are monopolized and form bonds with the small salt ions, causing the protein to aggregate. This interaction could result in the ordered arrangement of proteins to form crystals.[16]

Addition of organic solvents also produces similar effects by competing for water. This is due to the hydrogen bonding capabilities of organic compounds and leads to the aggregation and increase elastic forces between protein molecules. The mechanisms of polymers to facilitate protein crystallization are still not yet clearly understand, however it has been observed that polymers also induce protein crystallization by competing with proteins for water and by lowering the dielectric constant of the solution.

1.2.3 Other Parameters

Various other parameters are important in crystallizing a protein by gradually forcing the protein to become less soluble. Protein solubility is sensitive and typically relies greatly on parameters such as pH and temperature. [17] Most protein have been successfully been crystallized at their respective physiological pH. It was once believed that proteins would be more readily crystallized at their respective pI, however further research has proven that at physiological pH, there is less risk of denaturation. Temperature dependence of protein solubility is due to the protonation and deprotonation reaction constants of amino acid side chains.[18] Protein concentration is also a key parameter to allow for proper crystal nucleation. Common concentration ranges from 5 to 30 mg/mL but can range to even lower and higher concentrations depending on the solubility of the protein. Protein crystals will form and can form at various concentrations but it has been identified that only at optimal concentrations will proteins crystallize with reproducibility and consistency. The previously described parameters are only a few of many parameters that can influence protein crystallization. However, as of today there are still no clear definitive set of established parameters that can be associated as the key to successful protein crystallization as crystallization parameters will differ from protein to protein.

1.3   Cupin Family

The Cupin superfamily is a diverse family of proteins with a conserved 6 β-barrel domain with a high diversity and range of biological functions as this family is associated with Archaea, Eubacteria, and Eukarota.[19] The identity of the Cupin family originated from identification and recognition of a conserved motif found within proteins such as the wheat protein germin and germin-like proteins (GLPS). Most proteins in this family shows low overall sequence similarities however, they each contain two distinct short conserved cupin sequence motifs $G_{x5}H_xH_{x3-6}E_{x6}G$ (motif 1) and $G_{x5-7}PxG_{x2}H_{x3}N$ (motif 2), separated by a less conserved intermotif region (15-50 residues). The patented "jelly roll" barrel is formed with the cupin motif 1 and 2 along with the intermotif region where each forms two of the core six β-strands found in the folded structure. These degree of high hydrophobic interactions and short loops contribute to the thermal stability and protease resistant properties of proteins in this family.[20]

The Cupin superfamily consist of proteins with a wide range of biological functions including enzymatic activities such as decarboxylases, dioxygenases, hydrolases, isomerases, epimerases, and non-enzymatic functions such as seed storage or transcription factors.[21] Studies have shown that activity and function of cupins probably originated with ancestral cupins as simple sugar and cyclic nucleotide non-enzymatic binding domains which eventually evolved to metal-coordinating enzymatic domains. [22] This was based on the fact that a set of conserved histidine residues can be traced back to the ancestral cupins. A majority of the metal-dependent enzymes have been associated with metals such iron, nickel, zince, manganese, cobalt, or copper as each cofactor allows for a unique chemistry. Well known enzymes in the cupin family such oxalate oxidase and oxalate decarboxylase have been determined to be able to coordinate with a

Mn ion. Other enzymes in this family like acireductone dioxygenase binds Ni and praline-3-

hyroxylase utilizes Fe as a cofactor. Purposed mechanisms of these enzymes generally involve

the sequential binding of the substrate and a dioxygen molecule to the metal cation. Initially

metal binding enzymes in this family were thought to possess ligands consisting of 3 His-1-

carboxylate. It is now however apparent that there are many variations in the sequence as

deletions or substitutions of the ligands can result in alternate binding sites. As of today, there

are many notable variations of these ligands in the binding site due to structural characterization

work. For instance, the ligands of various enzymes can consist of 2-His, 2-His-1-Gln, 3-His, 3-

His-1-Glu, etc. [23] Efforts toward elucidating the structure and role of these metal binding

enzymes has been beneficial for understanding the nature of metal binding enzymes and their

mechanistic properties in vivo.

1.3.1   Thiol Dioxygenase

Our interest in the Cupin family revolves around a sub group of enzymes in this family

called thiol dioxygenase. In the mammalian cell there are many processes that can result in the

oxidation of thiol groups due to the reactive free sulfhydryl groups to form disulfides, sulfinates,

sulfenates, and sulfonates. Many of these oxidations occur nonenzymatically yet, there are still a

number of thiol reactions that occur by enzymatic reactions. [24] It has been shown that there are

enzymes involved in these reactions which can oxidize free sulfhydryl groups with high degree

and specificity. One interesting group of enzymes that can oxidize free sulfhydryl groups

includes thiol dioxygenases cysteine dioxygenase (CDO) and cysteamine dioxgyenase (ADO).

To date, CDO and ADO are the only two reported mammalian thiol dioxygenases. CDO and

ADO are unique iron-dependent enzymes that adds molecular oxygen to sulfhydryl group to

form a sulfinic acid in cysteine metabolism. These two thiol dioxygenases are essential for

hypotaurine/taurine biosynthesis.[25] Cysteine and dioxygen are used as substrates by CDO to

form cysteinesulfinic acid which can then (1) decarboxylate to hypotaurine or (2) transaminate to

3-sulfinylpyruvate which can decomposes to pyruvate. ADO on the other hand uses the

substrate cysteamine, a decarboxylated cysteine derived during the coenzyme A synthesis

pathway. Once cysteamine is released during coenzyme A degradation, ADO converts

cysteamine and dioxygen to form hypotaurine which is further oxidized to taurine (Figure 1).

Thus, both CDO and ADO play key roles in thiol metabolism by regulating cysteine and

cysteamine levels in the body for the production of hypotaurine/taurine from two different

cysteine metabolites. Thiol metabolism for the production of hypotaurine/taurine has become of

importance because of the physiological relevance of taurine for cardiovascular function, the
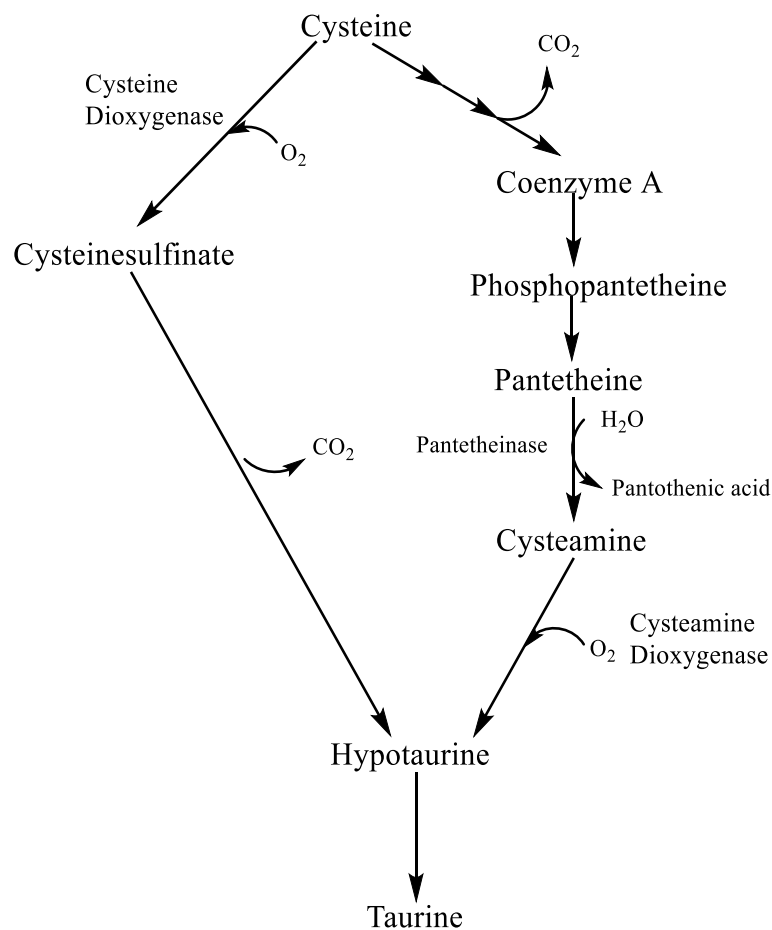
retina, and the central nervous system.[26]

Figure 1. Taurine biosynthesis pathway from cysteine. Image adapted from ref. 27.

1.3.1.1   CDO

Yamaguchi and coworkers were able to successfully purify and characterize CDO from rat liver in the 1970s. They discovered that CDO had the capability to bind one atom of iron per molecule and also went on to solve the amino acid sequence.[28,29] Future contributions from other groups went on to discover that the 200 amino acid residue protein was identical for rat and murine CDOs while human CDO only differed by 16 amino acid residues.[30] Influx of interest in CDO started when clinical evidence indicated that malfunction in cysteine metabolism due to CDO can lead to altered cysteine to sulfate ratio which causes sulfate depletion along with other adverse effects.[31,32] Other reports have shown that impaired cysteine catabolism was present in

patients afflicted with rheumatoid arthritis, liver diseases, Parkinson disease, Alzheimer disease, and systemic lupus erythematosus. Depletion of cysteine catabolism would either cause a depletion of sulfate and taurine while leading to an accumulation of substrate cysteine which has been speculated to impair sulfation reactions in vivo.[33,34,35]
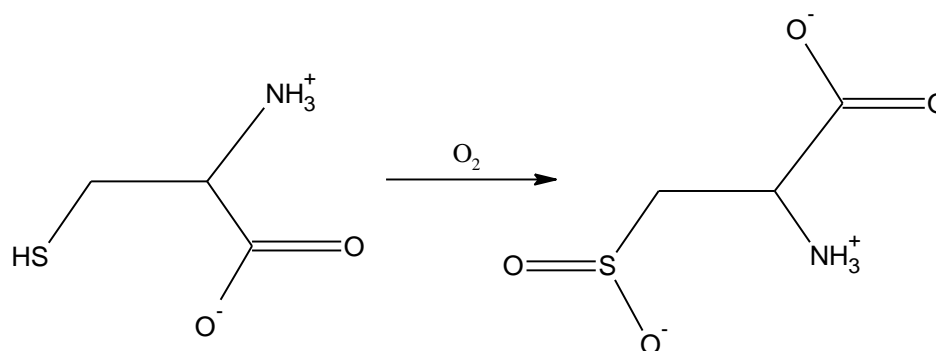

Figure 2. Reaction catalyzed by CDO

The elucidation of the crystal structure of CDO from rat, mouse, and human has revealed that it is indeed a member of the cupin family for it contained the patented cupin fold. CDO contains a β-sandwhich of seven anti-parallel β-strands on the lower side (β1, β2, β4, β7, β9, β12, and β13) and six anti-parallel β-strands (β3, β5, β6, β8, β10, β11) on the upper region.[31,36,37] The cupin core is present in CDO as β-strands 3 and 4 corresponds to βC and βD (cupin motif 1), β-strands 5 and 6 corresponds to βE and βF (intermotif region), and β-strands 7 and 8 corresponds to βG and βH (cupin motif 2). This illustration can be seen on Figure 3.

Another interesting discovery is that CDO contained a novel metal center with Fe as its cofactor. Instead of a 3-His-1-Glu variation, it was revealed that CDO coordinated Fe^{2+} by a facial triad of 3 histidine side chains without any carboxylate ligand present near the iron center. It was evident that no other residue replaced the conserved glutamate as a metal-coordinating ligand, thus indicating that CDO is a unique discovery as most monnuclear non-heme Fe^{2+} enzymes normally use the 2-His-1-Glu or 3-His-1-Glu coordination. Therefore, the tetrahedral

coordination of mononuclear iron of CDO is unique for most cupin superfamily typically utilizes a penta- or hexa- coordination metal center.   Further work on CDO indicated that the novel center of CDO would catalyze reactions differently than other classes of dioxygenases. Rather than cleavage of a C-C bond or hydroxylation, CDO involves oxidation of a sulfyhydryl group. Crystal structure of human CDO in complex with substrate L-cysteine[36] along with crystal structure of rate CDO with Fe^{2+} bound persulfenate intermediate suggest that Fe^{2+} directly activates both cysteine and O_2, causing a transfer of both oxygen atoms from the iron atom to a single sulfur atom.[31] This is possible because of the electron rich substrate compound. However, studying of the mechanism of CDO has proved to be quite challenging because both the substrate and resulting products are not detectable by UV-Vis absorption.

Crystal evidence from mammalian CDO also indicated that there is a unique crosslink in the active site of the enzyme (Figure 3). A Cys-Tyr intramolecular crosslink between Cys93 and Tyr157 is indicative of a rare thioether cofactor that has been identified to possibly play a role in regulating CDO function via a substrate-regulation from cysteine to modulate the Cys-Tyr cofactor formation as form of feed-forward activation.
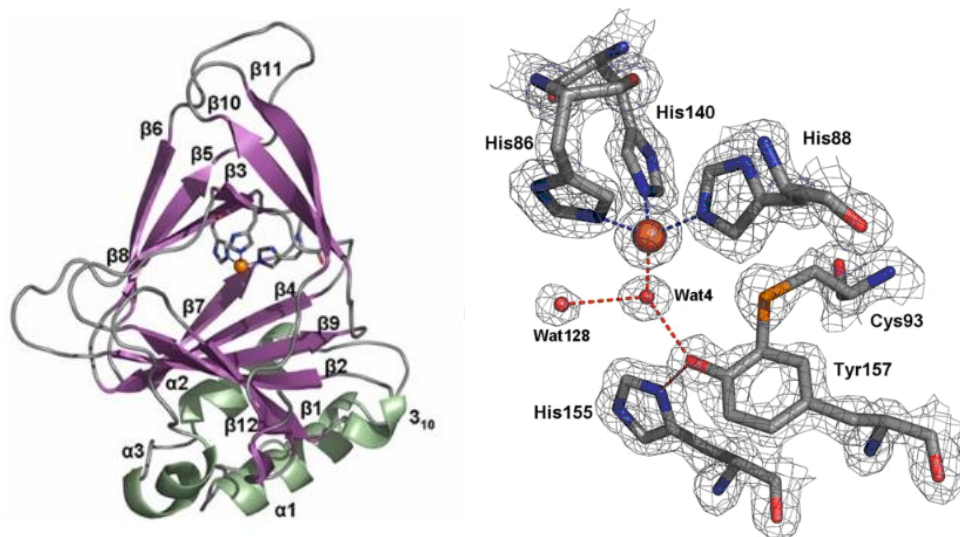
Figure 3. 1.5 Å Resolution Crystal Structure of Recombinant R. Norvegicus Cysteine Dioxygenase illustrating the purin domain and a Cys-Tyr crosslink. Image taken from ref. 31. (PDB 2B5H)

1.3.1.2 ADO

In 2007,studies were done on a hypothetical murine protein homolog of CDO encoded by the gene Gm237 in the DUF1647 family which was identified to show significant cysteamine dioxygenase activity in vitro.[24] Sequence alignment of this murine protein homolog which was named ADO when compared to CDO only resulted in 14%-19% similarity. However the alignment indicated that ADO belongs to the cupin family while also potentially exhibiting the same 3-His metal binding motif as CDO. Upon its first discovery from horse kidney in 1963, ADO has been reported to require sulfide, elemental sulfur, selenium, or hydroxylamine to function.[38] The recombinant ADO however, reported that the use of ferrous iron as a cofactor was necessary while requiring none of the exogenous elements mentioned previously. To date, little information about ADO is known. For several decades, a vast majority of the research effort was focused more on CDO because of CDO and its association as a major regulator of taurine synthesis. However, association of cysteamine as a potential intermediate in taurine

biosynthesis for the cause of Hallervorden-Spatz syndrome has renewed interest. Lack of panthothenate kinase (PANK2) which is necessary for the synthesis of coenzyme A and release of cysteamine during coenzyme A degradation has been identified to be a cause of Hallervorden-Spatz syndrome. [39] Furthermore, evidence also indicates that cysteamine could serve as an endogenous regulator of the immune system activity [40,41] and a potential therapeutic agent of Huntington disease. [42,43]

ADO like many other cupins possess the two conserved cupin motifs, however like CDO, it is missing the highly conserved glutamate in motif 1 as this residue is replaced by valine or glycine (compared to cysteine or glycine in CDO). Motif 2 for the cupin family is less conserved among the DUF1637 family but all contain the third metal-binding histidine residue (His 193 in human ADO). The alignment would suggest that this third histidine residue would be involved in iron ligand binding, however no direct evidence have suggested that. Even though both CDO and ADO utilizes a thiol substrate to catalyze a similar reaction, it has been reported that ADO could potentially contain an iron binding site that is different from CDO due to ADO's ability to retain bound iron while iron can dissociate rather easily from CDO.[24,31]

For these reasons, it is necessary to pursue higher levels of understanding of ADO. ADO is non-heme mononuclear iron dependent enzyme that catalyzes the oxidation of cysteamine to hypotaruine in the presence of oxygen. Sequence alignment of various ADO strands indicates that ADO consist of multiple intriguing conserved and missing residues as shown by Figure 5.

Figure 4. Sequence alignment of various strands of ADO.

Compasrison of hADO (1) with various strands of ADO. Thee conserved glutamate is replaced by ether glycine or valine for ADO. A potential cys-tyr crosslink similar to CDO can be also be seen. Note: black highlights are from conserved residues similar to the cupin family and lighter highlights are for conserved residues within ADOs.[27]

To date, no structure of ADO has been reported. The structure of an enzyme is a crucial factor for indebt mechanistic study and understanding the regulation roles of the enzyme in vivo. For these reasons, we are highly interested in trying to pursue the structure of ADO.
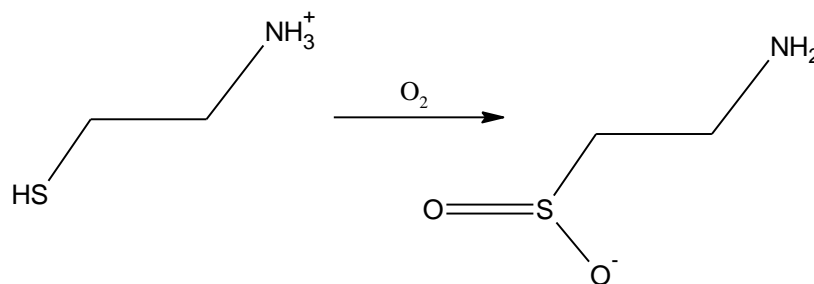

Figure 5. Reaction catalyzed by ADO

2 MATERIALS/METHODS

2.1 hADO Purification

The gene human ADO (hADO) was purchased from DNASU and was cloned to pET28 plasmid with a His-tag and transferred into B L21 (DE3) competent cells. The *E*-coli cells with hADO were grown in Luria Broth at 37°C with shaking at 200 rpm at first in 10 mL and then transferred to 50 mL medium size containing 10 mg/mL and 50 mg/mL kanamycin respective to the increasing Luria Broth size. hADO cells from the medium scale was then transferred to a 500 mL size Luria Broth with 100 mg/mL kanamycin. The large scale culture was grown at

37°C with shaking at 200 rpm until O.D. was approximately 0.6. At this point, the temperature was decreased to 25°C and induced with 0.6mM isopropyl-β-thiogalactopyranoside and allowed to grow overnight. The cells were then harvested by centrifugation at 4°C, 8000 x g for 20 min. Cell pellets were obtained and dissolved in a lysis buffer (20 mM Tris pH 8.5, 300 mM NaCl, 5% glycerol) and the re-suspended mixture was disrupted using a cell disruptor. The cell debris was centrifuged at 4°C and 27000 x g for 1 hour. The supernant was obtained and subjected to an affinity column (Ni-NTA resin, 26 x 200 mm) and eluted at a flow rate of 4 mL/min by using a linear imidazole gradient generated from buffer A (20 mM Tris-base, pH 8.0, 300 mM NaCl, 5mM imidazole, 5% glycerol) and 0-80% of buffer B (20 mM Tris-base, pH 8.0, 300 mM NaCl, 500 mM imidazole, 5 % glycerol). The hADO fraction was then concentrated and loaded onto a size exclusion column HiLoad Superdex 200 and eluted at a flow rate of 2 mL/min with a desalting buffer (10 mM Tris-base, pH 8.0, 5% glycerol). 12% SDS-PAGE was used to monitor and confirm that the peak had a consisted weight to that of the theoretical molecular weight of hADO mono (MW= 29 KDa).
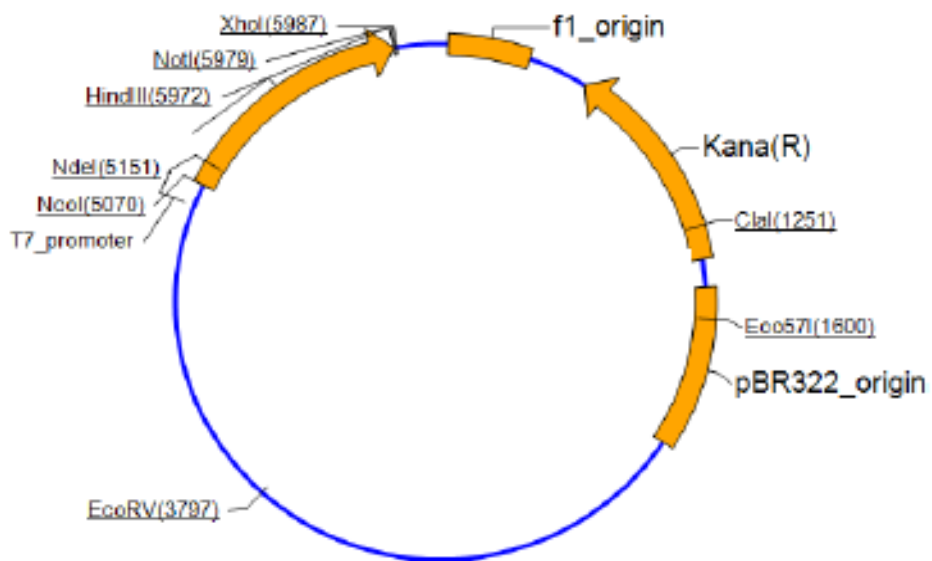
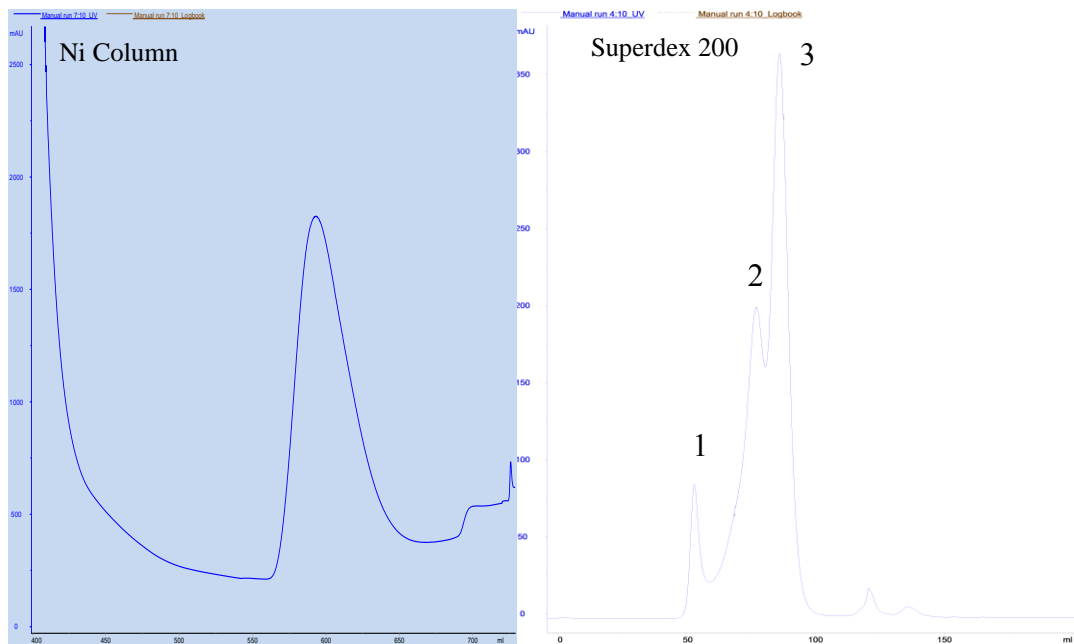*Figure 6. Plasmid map of ADO.*

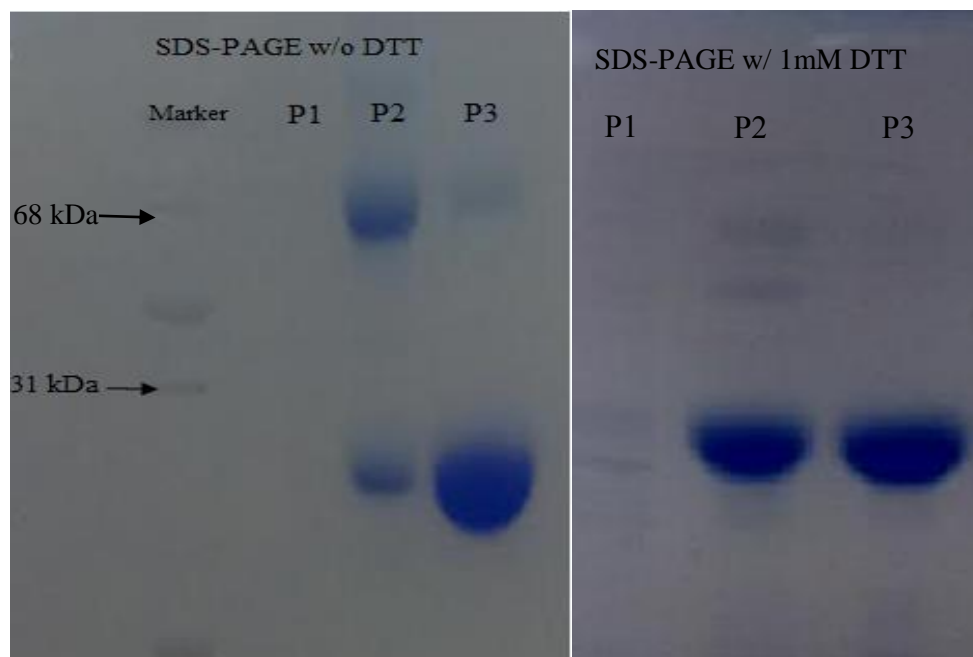*Figure 7. FPLC elution profile from Ni-NTA column and Superdex 75 on hADO.*

Figure 8. SDS-PAGE of hADO w/o DTT and with DTT after Superdex 75.
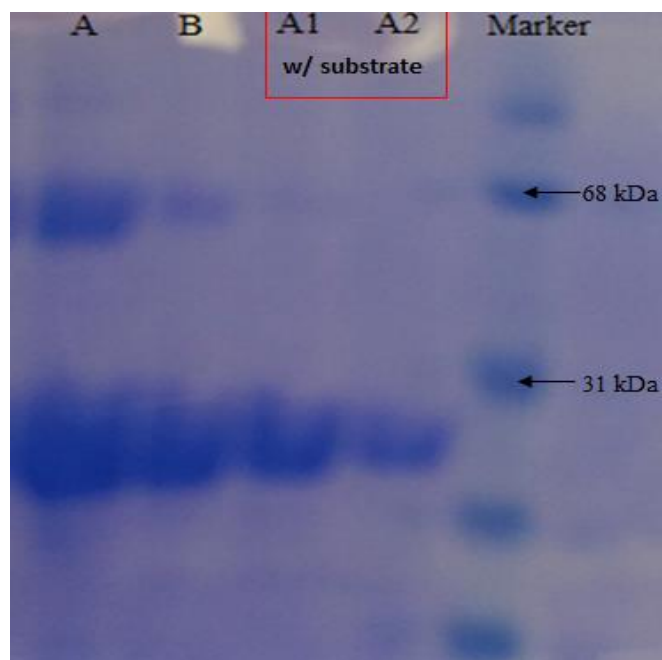

Figure 9. SDS-PAGE of hADO before and after treatment with 3mM substrate. A and B corresponds to hADO before treatment with substrate and A1 and B1 is after treatment with substrate.

2.2   hADO Crystallization Parameters

Crystals were observed for hADO protein prepared by two different methods.

2.2.1   *As Isolated hADO without His-tag*

hADO purified from HiLoad Superdex 75 was incubated with TEV protease (1:10 ratio) with stirring overnight at 4ºC. The TEV treated hADO was then subjected to a 5 mL Ni-NTA resin and eluted at a flow rate of 1.5 mL with a linear imidazole gradient generated from buffer A and buffer B. The hADO fraction was concentrated and loaded onto a desalting column and eluted with the desalting buffer without any glycerol.

2.2.2   *hADO without His-tag reconstituted with Fe*

hADO without His6-tag were incubated with 20mM EDTA while stirring overnight at 4ºC. The EDTA-treated hADO was then loaded onto a desalting column and eluted with the desalting buffer to remove any EDTA. The collected EDTA treated protein was then made anaerobic via a shrink line under constant N_2 gas. In an anaerobic glove box, iron ammonium sulfate was dissolved in anaerobic water and added to the anaerobically prepared hADO. The Fe-reconstituted hADO was then subjected to another desalting via a desalting column with desalting buffer without glycerol.

2.3   Screening by Hampton Crystal Kits

hADO prepared as previously mentioned were screened by approximately 100 different conditions each experiment using Hampton research crystal screen kits at varying protein concentrations (10 mg/mL, 15 mg/mL, 20 mg/mL, 25 mg/mL, 30 mg/mL). Fresh DTT or cysteamine substrate was prepared and 1mM or 10 mM of substrate were added to each protein concentration prior to screening. 1 uL of each protein concentration was mixed with 1 uL of

mother liquor via a Crystal Screen Robot via micro-batch diffusion method and stored and monitored at 16°C until crystals were observed and harvested.

The hADO crystals were harvested from the drops and flash frozen with liquid nitrogen and sent to SER-CAT for remote analysis. The diffraction data were collected at SER-CAT beam ID-22 or BM-22 line with 180 frames.

3   RESULTS

Evidence of potential crystals was observed at multiple conditions. Screening with Fe-reconstituted hADO yielded in unstable and un-harvestable seed shaped crystals at various conditions. The crystals from Figure 8 were the result of 15mg/mL hADO protein, 0.2M ammonium acetate, 0.1M Hepes pH 7.5, 45%v/v-2-methyl-2,4, pentanediol where 1 mM substrate was added to hADO resulting in A and 10 mM substrate for the crystal formed in B.
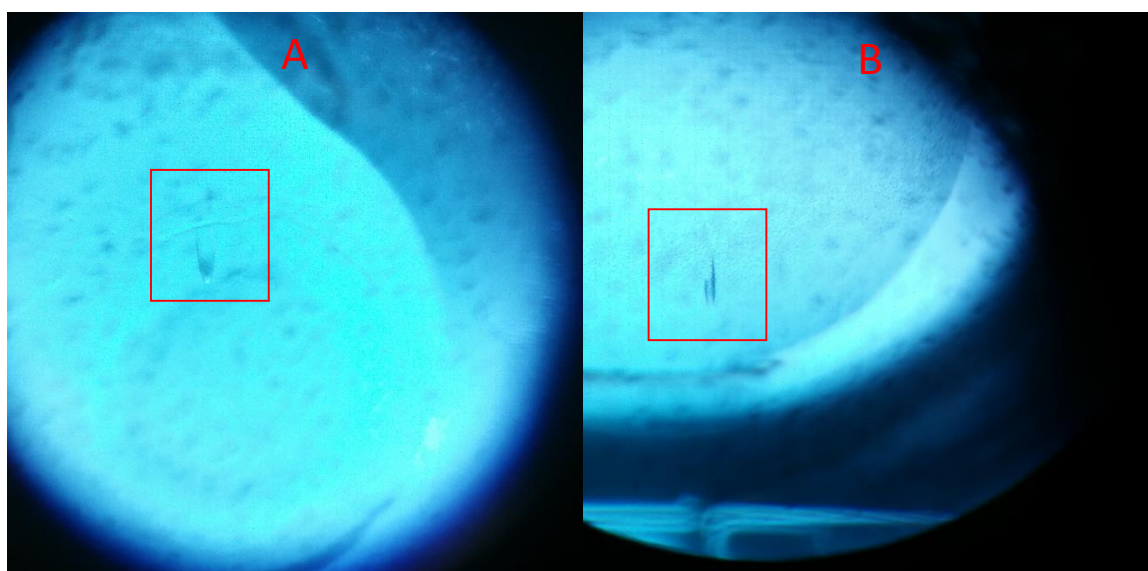
Figure 10. Seed-like hADO crystals were observed from the same condition at two different substrate concentration.

Crystals from C in Figure 9 resulted from 20 mg/mL of hADO, 10 mM substrate, 0.2M potassium tartrate tetrahydrate, 0.1M sodium citrate tribasic dehydrate pH 5.6, 2.0 M ammonium sulfate. Crystal D was the result of 15 mg/mL hADO protein, 10mM substrate, 0.2M lithium sulfate monohydrate, 0.1M Bis-Tris pH 5.5, and 25% w/v PEG 3350.
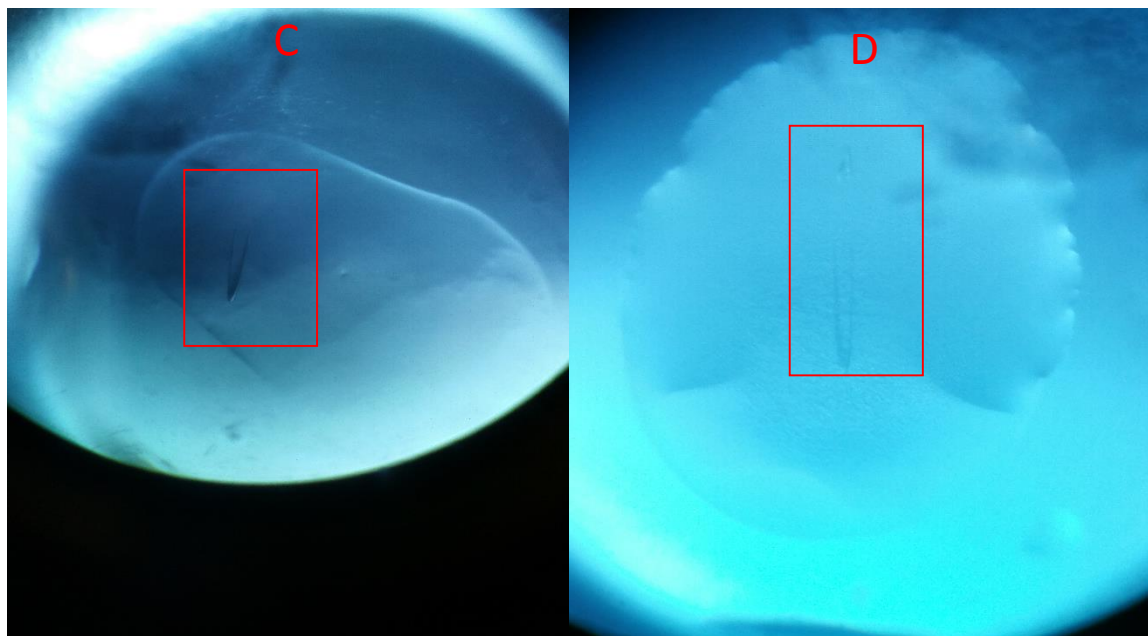
Figure 11. Seed-like hADO crystals observed with Fe-reconstituted hADO.

A different hADO crystal structure was observed at 15 mg/mL (E) and 20 mg/mL hADO (F), 0.2 Sodium fluoride, 20% w/v PEG 3350 from Fe-reconstituted hADO shown on Figure 10. The crystals were collected stable and robust enough to be collected.
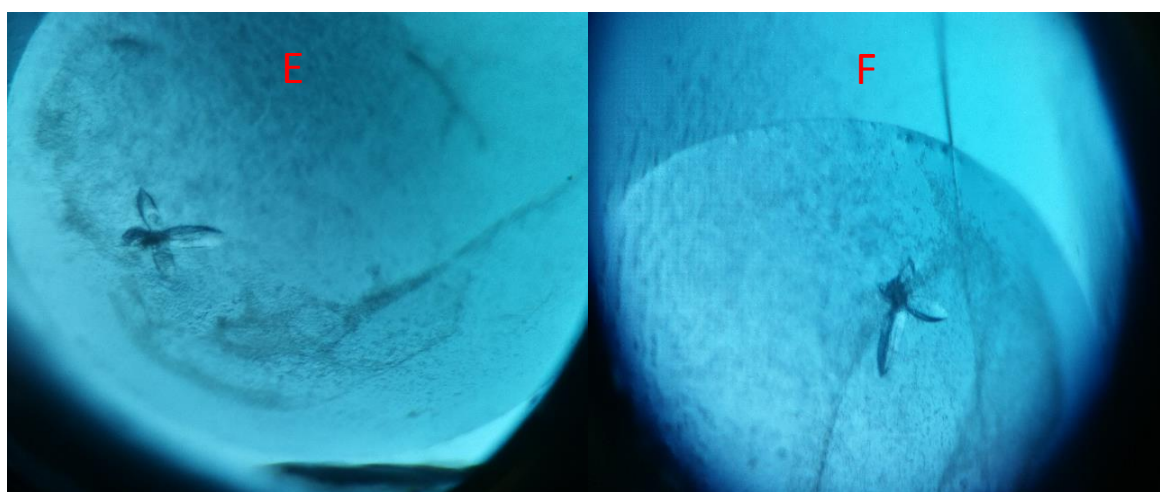
Figure 12. Flower-like hADO crystals observed with Fe-reconstituted hado.

hADO crystals without Fe reconstitution yielded in a long rod crystal rod at the condition of 20 mg/mL hADO, 0.1 M Sodium cacoldylate trihydrate pH 6.5, 0.1 M Sodium acetate trihydrate (Figure 11).

Figure 13. Rod-shaped hADO crystals

Figure 14. SER-CAT beam analysis of hADO crystals from Figure 9.

Figure 15. SER-CAT beam analysis of hADO crystals from Figure 9.

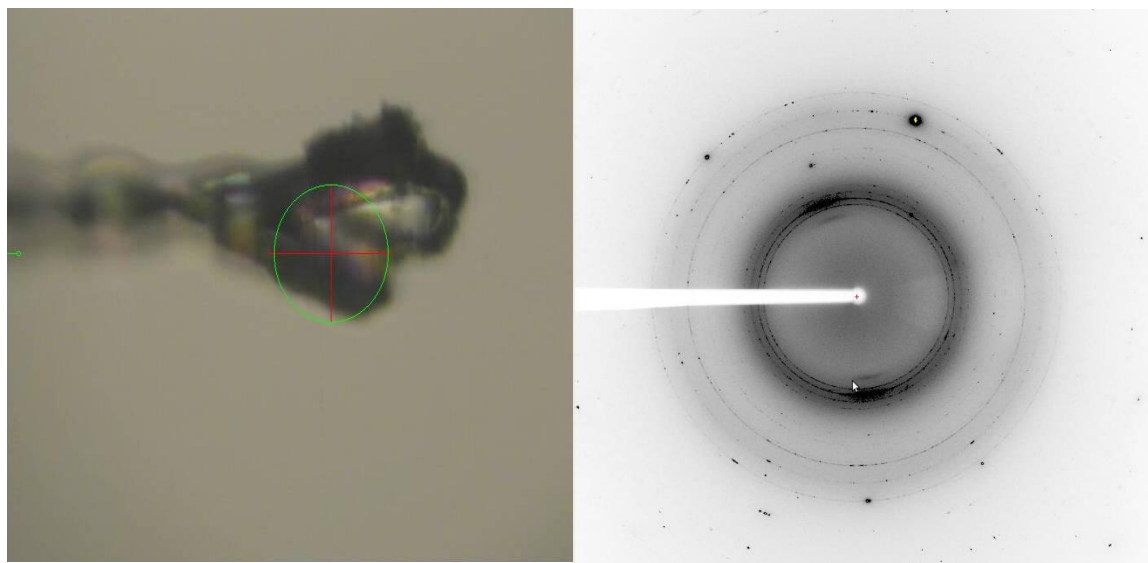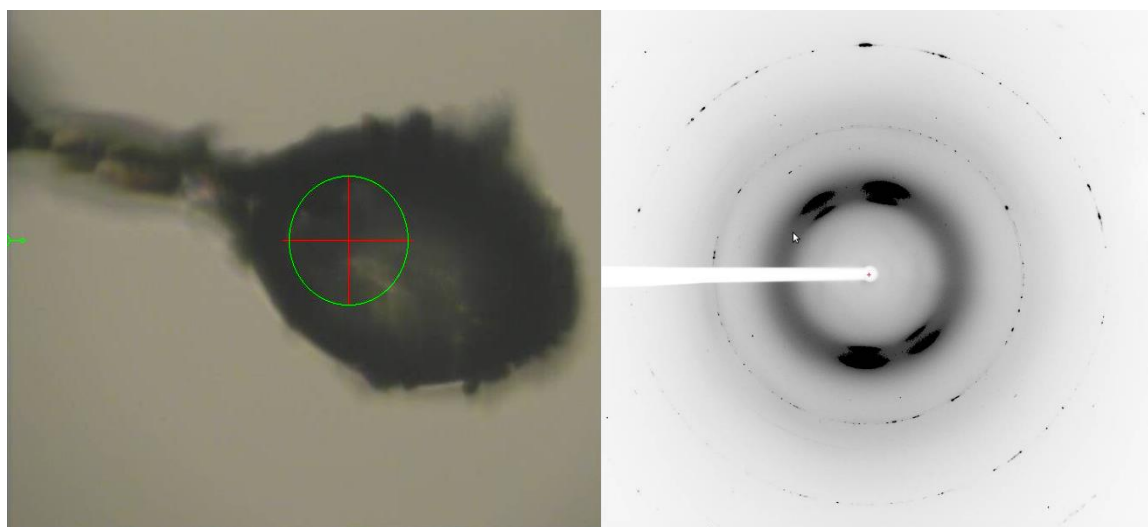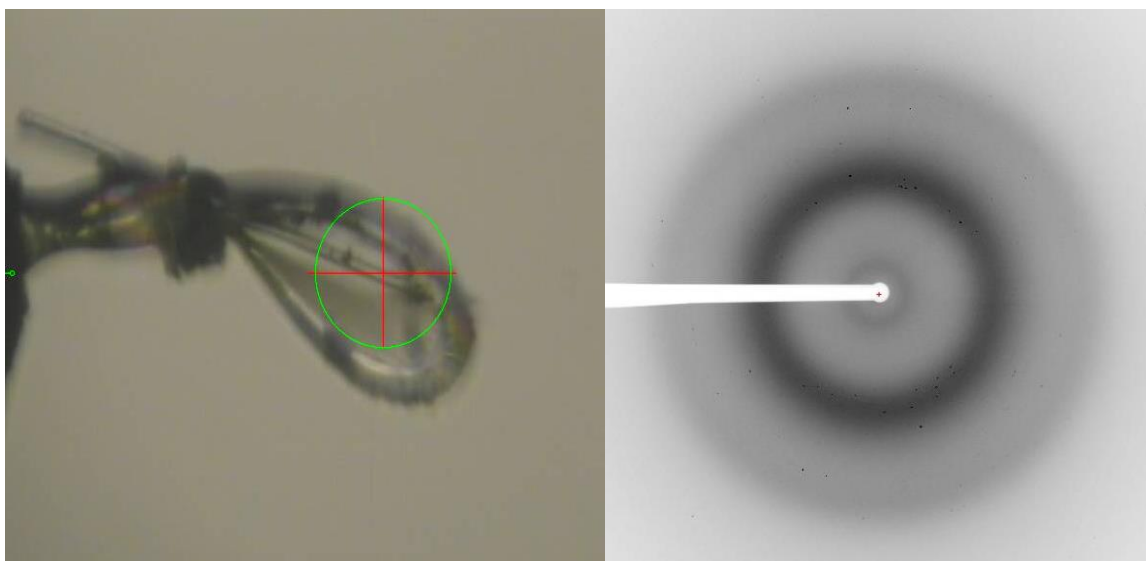*Figure 16. SER-CAT imaging analysis of the rod-shaped hADO.*

4   DISCUSSION

ADO is non-heme mononuclear iron dependent enzyme that catalyzes the oxidation of cysteamine to hypotaruine in the presence of oxygen. To date, no structure of ADO has been reported. Blasts study indicates that CDO is the closest relative of ADO but only with approximately 19% sequence similarity. Findings from the structure of CDO indicate that CDO utilizes a 3-His metal binding motif. Sequence study also indicates that ADO also has this conserved 3-His metal binding motif. The structure of an enzyme is a crucial factor for indebt mechanistic study and understanding the regulation roles of the enzyme in vivo.

Recombinant hADO cloned to pET28 plasmid with a His-tag was successfully cloned and purified as previously mentioned above. The purification conditions allowed for optimal purification of the protein normally yielding on average of 1-3 mg/mL. Ni-NTA purification of hADO showed an elution profile with normally one distinctive peak (Figure 5). SDS-PAGE of the protein after Ni-NTA shows two distinctive major bands approximately at 29 KDa and 58 KDa. Further purification of hADO with Superdex 75 will

result in typically 3 peaks (Figure 5). Peak 1does not show absorbance at 280 nm and does

not show any bands when analyzed by SDS-PAGE, thus indicating that it is not the desired

protein hADO. It does absorb strongly at 260 nm, suggesting that it could be from DNA.

Peak 2 and peak 3 both show strong protein absorbance at 280nm and SDS-PAGE further

validated the expected two bands of hADO (Figure 6). The protein proved to be rather

stable, making it optimal to handle the protein without constant freezing of protein as long

as the protein was kept under ice.

Crystallization screenings of hADO yielded in various potential conditions from the

tested parameters. Initial screenings of hADO with His-tag intact resulted in lack of protein

crystal formation. His-tag removal via TEV protease helped increase potential crystal

formation for the protein. It has previously been reported for CDO and other proteins that

His-tag removal may help in stabilizing the protein. The flexible tail end from His-tag may

disrupt proper protein nucleation. However, this is not true for all scenarios as it has been

shown that some proteins are easier and able to be crystallize with the his-tagg still intact.

When removed and purified from bacterial cell, hADO tends to for a mixture of monomer

and dimer as indicated from SDS-PAGE.   When DTT is added to the loading buffer gel, the

normally seen two bands converge into just one single band (Figure 6). This suggests that

the ADO eluted from the size exclusion column are in different multimer forms. This

resulted in initially screenings with DTT as a reducing agent for hADO at various

concentrations in order to keep hADO in a monomeric form. Crystal formation was

observed after these modifications, however the crystals observed did not have a definitive

shape and were highly unstable during crystal harvest. One key finding was that the

substrate cysteamine was able to reduce hADO to a monomer in the same manner that DTT

can as indicated via SDS-PAGE (Figure 7). It has been shown previously that substrate can

help to stabilize and support protein during crystallization[11,13]. This finding facilitated the

usage of the substrate cysteamine as a substitute for DTT as a reducing agent.

CDO has been previously reported to incorporate various divalent metals as a cofactor as

Mccoy and Stipanuk's was able to crystallize CDO with nickel and iron respectively.

Based on this notion, crystallization of hADO was explored with different metals. No

results were obtained from crystallization of hADO with nickel however, hADO with Fe-

reconstituted did result in formation of low diffracting crystals. Even though Fe has been

noted to be the preferred metal for catalysis, it has been proven by various groups that the

native metal is not always the desired metal needed for crystallization conditions. Test

revolving around as purified hADO without metal reconstitution or EDTA treatment yielded

in a crystal structure of a rod, very different from hADO reconstituted with Fe. hADO

grown in Luria broth will be incorporated with various metals ranging due to trace amounts

of metals present in the Luria broth. Thus, the identity of the metal resulting in the

formation of the rod crystal is unknown. The rod crystal collected also resulted in slightly

diffracting crystals but proved to be promising as it has been noted by Stipanuk that

crystallization of hCDO also resulted in rod-shaped crystals.

1   CONCLUSION

These results have helped to reveal crucial conditions needed for the structure elucidation

of hADO. From the nature of the crystal formation and diffraction patterns of the analyzed

crystals, it is evident that the crystals obtained are not from salt formation. The results for

hADO prove to be promising with various parameters needing to be refined in order to solve

the structure. For future work, the range of protein concentration of hADO should be explored in higher concentrations as the protein has shown promise to crystallize in concentration ranges from 15-25 mg/mL. More emphasis could also be focused on the metal cofactor which seems to play a role in the types of hADO crystals formed. Apo protein along with other divalent metals other than Fe can be explored as crystallization parameters. The role of the substrate cysteamine as a natural reducing agent has proven to be a crucial factor for crystal stability for harvesting; however various other reducing agents such as previously tested DTT should be revisited and tested at various concentrations. Once the optimal conditions for hADO crystals are determined, then procedures for phasing for a *de novo* structure needs to be implemented. This could be through various methods such as heavy metal soaking by incubating the protein crystals with metals such as Pt, Au, etc for a short time. Another approach is growing the protein with M9 media that is substituted with Set-Met.

REFERENCES

1. 1Stryer L, Berg JM, Tymoczko JL. *Biochemistry* (5th ed.). San Francisco: W.H. Freeman. 2002. Print.

2. 2Ole Kirk etc al. Industrial enzyme applications. *Current Opinion in Biotechnology 2002*.14(4): 345-351

3. *3*Carey, Charles W. American Scientists. *Infobase Publishing*, 354–355. *Retrieved 20 January 2016*.

4. "The Nobel Prize in Physics 1915". Nobel Foundation. *Retrieved 2016-02-10*

5. Wlodawer, A., Minor, W., Dauter, Z. and Jaskolski, M. Protein crystallography for non-crystallographers. *FEBS Journal* **2008**, 275: 1–21

6. S.D. Durbin and G. Feher, Protein crystallization, Annu. Rev. *Phys. Chem,* **1996**, 47:171

7. C. Weber, Physical principles of protein crystallization, *Adv. Prot. Chem*, **1991**. 41:1

8. McPherson, Review: Current approaches to macromolecular crystallization, *Eur. J. Biochem* **1990.** 189:1

9. Crystallization Research Tools, Hampton Research Corp, **1999**, 9:1.

10. McPherson. Preparation and analysis of protein crystals. J. Wiley & Sons, New York. 1982. Print.

11. Hassell, Anne M. et al. "Crystallization of Protein–ligand Complexes." *Acta Crystallographica Section D: Biological Crystallography* 63.Pt 1 **2007**: 72–79

12. He B, Gampe RT Jr, et al. Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. *Mol Cell*. 2004, 16(3):425-38.

13. Sack JS etc al. Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Prox Natl Acad Sci USA* 2001, **98(9): 425-38.**

14. McPherson, Crystallization of biological macromolecules (Spring Harbor Laboratory Press, New York, 1999)

15. S. Trakhanov and F. Quiocho,Influence of divalent cations in protein crystallization. *Protein Sci* **1995.** 4:1914.

16. J.K. Baird, S.C. Scott and Y.W. Kim, Theory of the effect of pH and ionic strength on the nucleation of protein crystals, *J. Cryst. Growth* **2001.** 232:50

17. J.P. Guilloteau, M.M. Ries-Kautt and A.F. Ducruix, Variation of lysozyme solubility as a function of temperature in the presence of organic and inorganic salts, *J. Cryst. Growth* **1992,** 122:223

18. A.A. Chernov and H. Komatsu: *Principles of crystal growth in protein crystallization, in Science and Technology of Crystal Growth*. Kluwer Academic Publishers, Dordrecht. 1995. Print.

19. Dunwell JM, Khuri S, Gane PJ. Microbial relatives of the seed storage proteins of higher plants: conservation of structure and diversification of function during evolution of the cupin super-family. *Microbiol Mol Biol Rev*. **2000**. 64:153–179

20. Dunwell JM, Culham A, Carter CE, Sosa-Aguirre CR, Goodenough PW. Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* **2001,** 26: 740–746

21. Stipanuk, M. H. Annu. Rev. *Nutr* **2004,** .24,539–577

22. Uberto R, Moomaw EW. Protein Similarity Networks Reveal Relationships among Sequence, Structure, and Function within the Cupin Superfamily. *PLoS ONE* **2013,** 8(9)

23. Ferman A. Chavez, Atanu Banerjee and Bojan Sljivic. Modeling the Metal Binding Site in Cupin Proteins. *Biomimetics*. **2011**. Web 25, Feb 2016.

24. Dominy, J. E., Jr.; Simmons, C. R.; Hirschberger, L. L.; Hwang, J.; Coloso, R. M.; Stipanuk, M. H., Discovery and characterization of a second mammalian thiol dioxygenase, cysteamine dioxygenase. *The Journal of biological chemistry* **2007**, 282 (35), 25189-98.

25. Coloso, R. M., Hirschberger, L. L., Dominy, J. E., Lee, J. I., and Stipanuk, M.H. Cysteamine dioxygenase: evidence for the physiological conversion of cysteamine to hypotaurine in rat and mouse tissues. *Adv. Exp. Med. Biol*. **2006,** 583, 25–36

26. Huxtable, RJ. Physiological actions of taurine. *Physiol Rev* **1992,** 72(1): 101–63

27. Stipanuk, M.H. etc al. Thiol dioxygenases: unique families of cupin proteins. *Amino Acids*, **2011**, 41: 91-102

28. Yamaguchi K, Hosokawa Y, Kohashi N, Kori Y, Sakakibara S, Ueda I. Rat liver cysteine dioxygenase (cysteine oxidase). Further purification, characterization, and analysis of the activation and inactivation. *J Biochem (Tokyo)* **1978**, 83:479–491.

29. Sakakibara S, Yamaguchi K, Hosokawa Y, Kohashi N, Ueda I. Purification and some properties of rat liver cysteine oxidase (cysteine dioxygenase) *Biochim Biophys Acta*. **1976**; 422:273–279.

30. Hirschberger LL, Daval S, Stover PJ, Stipanuk MH. Murine cysteine dioxygenase gene: structural organization, tissue-specific expression and promoter identification. *Gene.* **2001**; 277:153–161.

31. Simmons, C. R.; Liu, Q.; Huang, Q.; Hao, Q.; Begley, T. P.; Karplus,P. A.; Stipanuk, M. H., Crystal structure of mammalian cysteine dioxygenase. A novel mononuclear iron center for cysteine thiol oxidation. *The Journal of biological chemistry* **2006**, 281(27), 18723-33.

32. Knopf, K., Sturman, J. A., Armstrong, M., and Hayes, K. C., Taurine: an essential nutrient for the cat, *J. Nutr*. **1978** 108:773-77

33. Heafield, M. T., and Williams, A. C. Curr. Opin. Neurol. *Neurosurg*. **1992,** 5, 288–294

34. Bradley, H., Gough, A., Sokhi, R. S., Hassell, A., Waring, R., and Emery, P. Sulfate metabolism is abnormal in patients with rheumatoid arthritis. Confirmation by in vivo biochemical findings. *J. Rheumatol*. **1994,** 21, 1192–1196

35. Davies, M. H., Ngong, J. M., Pean, A., Vickers, C. R., Waring, R. H., and Elias, E. Sulphoxidation and sulphation capacity in patients with primary biliary cirrhosis. *J. Hepatol*. **1995,** 22, 551–560

36. Ye S, Wu X, Wei L, Tang D, Sun P, Bartlam M, Rao Z. An insight into the mechanism of human cysteine dioxygenase. Key roles of the thioether-bonded tyrosine-cysteine cofactor. *J Biol Chem*. **2007**. 282:3391–3402

37. McCoy, J. G.; Bailey, L. J.; Bitto, E.; Bingman, C. A.; Aceti, D. J.; Fox, B. G.; Phillips, G. N., Jr., Structure and mechanism of mouse cysteine dioxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, 103(9), 3084-9.

38. Cavallini, D., Scandurra, R., and Demarco, C. The enzymatic oxidation of cysteamine to hypotaurine in the presence of sulfide. *J. Biol. Chem*. **1963.** 238, 2999–3005

39. Zhou, B., Westaway, S. K., Levinson, B., Johnson, M. A., Gitschier, J., and Hayflick, S. J., A novel pantothenate kinase gene (PANK2) is defective in Hallervorden-Spatz syndrome.*Nat. Genet*.**2001,** 28:345-349.

40. Berruyer, C., Martin, F. M., Castellano, R., Macone, A., Malergue, F., Garrido-Urbani, S., Millet, V., Imbert, J., Dupre, S., Pitari, G., Naquet, P., and Galland, F. (2004) *Mol. Cell. Biol.* **24**, 7214–7224

41. Berruyer, C., Pouyet, L., Millet, V., Martin, F. M., LeGoffic, A., Canonici, A., Garcia, S., Bagnis, C., Naquet, P., and Galland, F. *J. Exp. Med.* **2006**, 203, 2817–2827

42. Dedeoglu, A., Kubilus, J. K., Jeitner, T. M., Matson, S. A., Bogdanov, M., Kowall, N. W., Matson, W. R., Cooper, A. J., Ratan, R. R., Beal, M. F., Hersch, S. M., and Ferrante, R. J. *J. Neurosci.* **2002,** 22, 8942–8950

43. Lesort, M., Lee, M., Tucholski, J., and Johnson, G. V. (2003) *J. Biol. Chem.* **2003,** 278, 3825–3830